Implicit Neural Representation for Spherical Harmonics Reconstruction of Motion-Corrupted Fetal Diffusion MRI

Wenxuan Wu[1(✉)], Irina Grigorescu[1], Ruowen Qu[2], Joseph V. Hajnal[1], J-Donald Tournier[1], and Maria Deprez[1]

[1] Research Department of Early Life Imaging, School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK
wenxuan.wu@kcl.ac.uk

[2] School of Electronic Engineering and Computer Science, Queen Mary University of London, London, UK

Abstract. Diffusion MRI of the fetal brain has potential to provide important insights into early development of white matter fibers. However, in vivo fetal acquisitions are severely corrupted by unpredictable motion, making the data unusable for downstream analysis. To address this challenge, we propose an implicit neural representation framework for fetal diffusion MRI reconstruction. The proposed method models the diffusion signal as a spatially and angularly continuous field by predicting direction-independent spherical harmonics (SH) coefficients from continuous spatial coordinates. Based on this representation, we formulate a physics-based forward slice acquisition model that links the underlying clean diffusion volume with the observed motion-corrupted slices, enabling joint optimization of the diffusion signal, fetal motion and acquisition artefacts. Furthermore, the single slice orientation acquisition typically used for each diffusion gradient introduces an ill-posed reconstruction problem along the slice-select direction. To mitigate this issue, we introduce a novel Through-Plane Continuity regularization loss that enforces smoothness in through-plane direction to suppress stripe artifacts, while preserving genuine anatomical details. Experiments on one in vivo fetal dataset and two simulated neonatal datasets demonstrate that the proposed framework achieves improved reconstruction fidelity, sharper anatomical features, and stronger robustness to acquisition artifacts. The code is available at: https://github.com/baby-MedIA/INR-dMRI-Recon

Keywords: Diffusion MRI · Slice-To-Volume Reconstruction· Fetal Brain

1 Introduction

Diffusion Magnetic Resonance Imaging (dMRI) is a non-invasive modality for detailed investigation of developing white matter tracts in fetal brain by probing the restricted random walk of water molecules within tissues [12–14,20,21,26,41] However, acquiring high-fidelity 3D fetal MRI is profoundly challenged by unpredictable fetal motion. Although single-shot fast imaging sequences such as

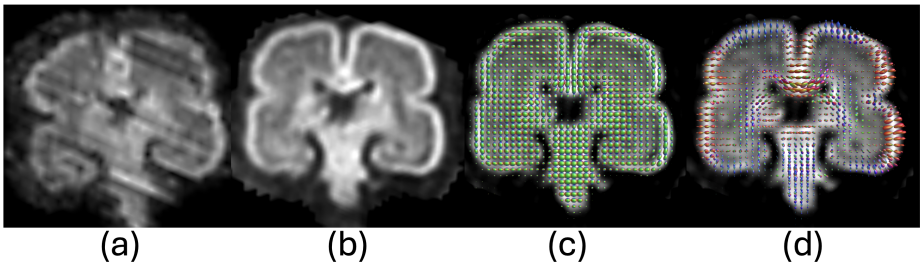

Fig. 1: Visualization of fetal diffusion MRI reconstruction. (a) Motion-corrupted diffusion-weighted images. (b) Reconstructed diffusion-weighted images. (c) SH-based diffusion representation where each voxel corresponds to a spherical function. (d) Fiber orientation density estimated from the reconstructed SH representation.

Echo Planar Imaging [18, 33] effectively freeze motion during acquisition of each 2D slice, the inter-slice motion remains present (Fig. 1a). This misalignment results in corrupted 3D volumes that are fundamentally unsuitable for quantitative microstructural analysis (Fig. 4 (a)).

To recover high-resolution continuous volumes from these corrupted stacks, Slice-to-Volume Reconstruction (SVR) algorithms [5, 8, 27, 29, 35–38, 40] have been extensively developed. Traditional model-based methods [8, 16, 27] employ an iterative two-stage optimization process alternating between slice-to-volume registration and scattered data interpolation. Initial learning-based paradigms emulate this two-stage strategy [35–37, 39] by training deep neural networks [9, 15, 17, 34] to predict slice motion parameters and subsequently apply super-resolution algorithms. While these supervised architectures significantly reduce processing time and deliver strong performance, they critically depend on the availability of high-quality motion-free data to serve as pseudo-ground truth during the training phase. To bypass this severe data dependency constraint, self-supervised approaches [29, 38] leveraging Implicit Neural Representations (INR) [19, 28] have demonstrated great promise. By mapping continuous spatial coordinates to image intensities and simulating the physical slice acquisition process, these implicit frameworks jointly optimize the underlying 3D volume alongside rigid motion and bias fields to achieve rapid super-resolution without requiring external training data.

However, current learning-based SVR frameworks are designed for structural MRI, which models a simple scalar intensity field. Diffusion MRI, on the other hand, embodies voxels with complex spherical functions where the observed signal depends heavily on the applied diffusion gradient vector [6, 7], require correction of the relative orientation of the tissue microstructure disrupted by fetal motion, and require sampling diverse gradient directions within a strictly limited scan time. Consequently, data for any specific gradient is typically acquired along a single slice orientation rather than the mutually orthogonal stacks common in structural imaging [11, 32]. This lack of intersecting geometric constraints causes the network to overfit through-plane motion artifacts as high-frequency signals

which visually manifests as severe stripe artifacts aligned with the acquisition plane (see Fig. 5).

Current model-based SVR frameworks for fetal diffusion imaging [2, 5] combine traditional slice-to-volume registration pipelines with per-voxel spherical harmonics (SH) fitting. Operating on a pre-defined discrete voxel grid inherently imposes a computational complexity bottleneck that restricts high-resolution outputs. Crucially, these classical pipelines lack explicit geometric regularization to counteract the single-orientation acquisition constraints which introduce the stripe artifacts directly into the reconstructed volume.

The main contributions of this work are summarized as follows:

- To the best of our knowledge, we propose the first learning-based reconstruction approach for motion-corrupted fetal diffusion MRI (Fig. 1a). This self-supervised INR framework estimates the underlying diffusion signal (Fig. 1b) as a continuous SH-parameterized 3D field (Fig. 1c), while correcting fetal motion, gradient orientation, and acquisition-specific nuisance parameters within a unified end-to-end optimization scheme.
- We introduce a novel Through-Plane Continuity regularization that explicitly targets stripe artifacts that are difficult to correct in single-orientation acquisitions (Fig. 1a). By penalizing signal discontinuities along the slice-selection axis, we suppress motion-induced striping while preserving genuine anatomical boundaries.
- The reconstructed diffusion signal (Fig. 1c) is represented in a continuous SH basis, which naturally supports estimation of Fiber Orientation Distributions (Fig. 1d) estimation via Constrained Spherical Deconvolution [30, 31] and facilitates tractography without additional resampling or interpolation.
- Extensive experiments demonstrate that the proposed method achieves superior reconstruction fidelity, sharper structural detail, and improved artifact suppression compared to classical iterative reconstruction methods.

2 Methods

2.1 Continuous Representation of Diffusion MRI Signal

Diffusion MRI (dMRI) is sensitive to the diffusion of water molecules within a certain diffusion time along different gradient directions $\mathbf{g}$. There is significant diffusion along WM tracts, while diffusion of water across the tracts is impeded due to microstructural barriers. If water molecules can only diffuse by a short distance in the direction of the gradient, the dMRI signal remains high; conversely, the further the water molecules travel along $\mathbf{g}$ (the gradient direction), the larger the reduction in the observed diffusion signal [6, 7].

Our aim is to reconstruct the underlying theoretical artefact-free continuous diffusion signal $V(\mathbf{x}, \mathbf{g})$ from the acquired discrete data corrupted by motion and intensity artefacts. For each spatial coordinate $\mathbf{x}$, the diffusion signal is a function on the sphere, parametrised by the direction of the diffusion gradient $\mathbf{g}$, represented as a 3D unit vector, or alternatively using spherical coordinates (θ, ϕ).

We model this spatially varying spherical signal using the spherical harmonics (SH) representation

$$V(\mathbf{x}, \mathbf{g}) \approx \sum_{l=0}^{L} \sum_{m=-l}^{l} c_l^m(\mathbf{x})\, Y_l^m(\mathbf{g}), \tag{1}$$

where $Y_l^m(\mathbf{g})$ are orthonormal SH basis of degree $0 \leq l \leq L$ and order $|m| \leq l$

$$Y_l^m(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}}\, P_l^m(\cos\theta)\, e^{im\phi}. \tag{2}$$

These basis are eigenfunctions of the spherical Laplacian where the degree l controls the angular frequency, while the order m specifies the azimuthal variation within that frequency band. To represent the diffusion signal we adopt a real-valued SH basis where only even degrees are retained due to antipodal symmetry. The spherical signal at each spatial location $\mathbf{x}$ is thus encoded by a coefficient vector $\mathbf{c}(\mathbf{x}) \in \mathbb{R}^M$ with $M = (L+1)(L+2)/2$ [30,31].

We propose to define the spherical harmonic respresentation of the diffusion signal on the spatially continuous domain (as opposed to an image grid) using Implicit Neural Representations

$$\mathbf{c}(\mathbf{x}) = \mathrm{MLP}\Big(\gamma(\mathbf{x})\Big), \tag{3}$$

where γ is a multi-resolution hash grid encoding [22] designed to efficiently capture high-frequency anatomical details [19, 28, 38].

2.2 Slice Acquisition Model

High Angular Resolution Diffusion Imaging (HARDI) [6,7] samples the diffusion-weighted signal across a discrete set of spatial locations and gradient directions. The acquired data is organized into multiple stacks $\{\mathcal{K}_k\}$. Each stack $\mathcal{K}_k$ consists of a sequence of 2D slices $\{S_{k,j}\}$, where the data is sampled on a regular grid with voxel coordinates $\mathbf{x}_{k,j,i}$ in the scanner coordinate system. All slices within a single stack share a common diffusion gradient direction $\mathbf{g}_k$. Because of the unpredictable and rapid movement of the fetus during the scan [5, 8, 16, 27] each acquired slice $S_{k,j}$ is associated with a unique rigid transformation

$$T_{k,j}(\mathbf{x}') = \mathbf{R}_{k,j}\mathbf{x}' + \mathbf{t}_{k,j} \tag{4}$$

that maps the anatomical coordinate $\mathbf{x}'$ to the scanner coordinate $\mathbf{x}$, where $\mathbf{R}_{k,j} \in SO(3)$ and $\mathbf{t}_{k,j} \in \mathbb{R}^3$ denote the rotation and translation parameters respectively. This transformation describes the position of the fetal head in scanner space during the acquisition of the slice $S_{k,j}$ through the relationship $\mathbf{x} = T_{k,j}(\mathbf{x}')$. The diffusion gradient direction for each slice is also rotated to $\mathbf{g}'_{k,j} = \mathbf{R}_{k,j}^{-1}\mathbf{g}_k$ to maintain physical consistency with the fetal head orientation [5].

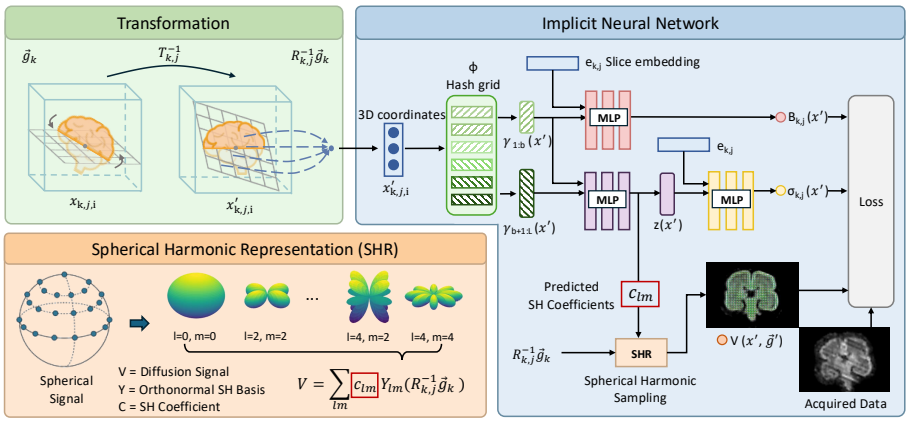

Fig. 2: The architecture of the proposed framework.

The signal observed at a slice location $\mathbf{x}_{k,j,i}$ is therefore sampled from the theoretical 3D diffusion signal $I'_{k,j}(\mathbf{x}', \mathbf{g}'_{k,j})$ through a point spread function (PSF) $P_{k,j}$ centered on the anatomical position $\mathbf{x}'_{k,j,i} = T^{-1}_{k,j}(\mathbf{x}_{k,j,i})$ [5,8,16]. The PSF $P_{k,j}$ accounts for the finite resolution of the imaging system, and is typically approximated as a 3D Gaussian [16]. In the scanner space, this Gaussian is centered on the acquired voxel position $\mathbf{x}_{k,j,i}$ and has a diagonal covariance matrix where variances on the diagonal reflect the resolution of the acquired slices. The acquired signal $S_{k,j}$ at voxel $\mathbf{x}_{k,j,i}$ is obtained by convolving $I'_{k,j}$ with $P_{k,j}$ and adding acquisition noise $\epsilon_{k,j,i}$ expressed in scanner space as follows:

$$S_{k,j}(\mathbf{x}_{k,j,i}) = I'_{k,j}(T^{-1}_{k,j}(\mathbf{x}_{k,j,i}), \mathbf{R}^{-1}_{k,j}\mathbf{g}_k) * P_{k,j}(\mathbf{x}_{k,j,i}) + \epsilon_{k,j,i}. \tag{5}$$

Assuming motion- and artefact-free continuous diffusion volume $V(\mathbf{x}', \mathbf{g})$ in a canonical anatomical space, the intermediate motion-free diffusion signal $I'_{k,j}$ can be modelled by first sampling the motion-corrected gradient-sensitised volume $V(\mathbf{x}', \mathbf{g}'_{k,j})$ from the SH representation 1, followed by multiplication with slice-wise bias field $B_{k,j}(\mathbf{x}')$, a slow-varying modulation of slice intensity due to MRI acquisition artefacts, such as inhomogeneity of the magnetic fields and interaction of magnetisation with fetal motion (so-called spin history effects):

$$I'_{k,j}(\mathbf{x}', \mathbf{g}'_{k,j}) = B_{k,j}(\mathbf{x}')V(\mathbf{x}', \mathbf{g}'_{k,j}). \tag{6}$$

The full forward acquisition model is obtained by combining equations 1, 5 and 6.

2.3 Model architecture

Instead of directly regressing the diffusion signal $V(\mathbf{x}', \mathbf{g}')$, our primary volume network MLP_V is designed to predict direction-independent SH coefficients

(eqn. 3) as a function of anatomical coordinate $\mathbf{x}'$. Specifically, the hash-encoded spatial features $\gamma(\mathbf{x}')$ are fed into this network to yield a set of even-degree SH coefficients $c_l^m(\mathbf{x}')$ alongside a spatial latent vector $\mathbf{z}(\mathbf{x}') \in \mathbb{R}^d$

$$\mathbf{c}(\mathbf{x}'),\, \mathbf{z}(\mathbf{x}') = \mathrm{MLP}_V\Big(\gamma(\mathbf{x}')\Big). \tag{7}$$

Using these predicted spatial SH coefficients, the clean diffusion signal is reconstructed for any given gradient (eqn. 1). The predicted latent vector $\mathbf{z}(\mathbf{x}')$ encodes the anatomical context and is subsequently concatenated with the learnable slice embedding vector $\mathbf{e}_{k,j}$. This combined vector is processed by an uncertainty network MLP_U to predict the slice-dependent spatially varying standard deviation $\sigma_{k,j}(\mathbf{x}')$ of the diffusion signal. This predicted uncertainty quantifies the localized noise level where higher values indicate a greater probability of corruption by motion artifacts or sensor noise.

In parallel, a dedicated bias branch MLP_B estimates the slice-dependent multiplicative bias fields $B_{k,j}$ from the lower-resolution hash encoding feature components γ_{low} and the slice embedding $\mathbf{e}_{k,j}$

$$B_{k,j}(\mathbf{x}') = \mathrm{MLP}_B\Big(\gamma_{\mathrm{low}}(\mathbf{x}'),\, \mathbf{e}_{k,j}\Big).$$

This specific architectural bottleneck restricts the high-frequency information capacity and forces the network to learn the slice-dependent intensity variations relying strictly on low-frequency spatial context which aligns with the established physical nature of the MRI bias field.

Per-slice transformations $T_{k,j}$ are encoded by axis-angle representation with six learnable parameters per transformation. Additionally, we learn per-slice uncertainty $\nu_{k,j}$ to exclude corrupted and misaligned slices.

2.4 Loss functions

We formulate the reconstruction of the diffusion signal as an inverse problem, where the goal is to recover a continuous volumetric representation $V(\mathbf{x}', \mathbf{g}')$, so that slices $S_{k,j}$ simulated using the forward acquisition model (eqn. 1,5, 6) are consistent with the observed measurements $\hat{S}_{k,j}$. The primary data fidelity term $\mathcal{L}_D$ is therefore formulated as the negative log-likelihood of a Gaussian distribution, which utilises the predicted uncertainty $\sigma_{k,j}(\mathbf{x}') + \nu_{k,j}$ to down-weight the reconstruction penalty for outlier pixels and slices heavily corrupted by motion artifacts or sensor noise

$$\mathcal{L}_D = \frac{1}{N} \sum_{k,j,i} \left(\frac{\Big(\hat{S}_{k,j,i} - S_{k,j}(\mathbf{x}_{k,j,i})\Big)^2}{2\,\sigma_{k,j}^2(\mathbf{x}'_{k,j,i})} + \log(\sigma_{k,j}(\mathbf{x}'_{k,j,i}) + \nu_{k,j}) \right), \tag{8}$$

where N is the number of acquired voxels. In addition to uncertainty, this loss allows for joint optimisation of the SH coefficients $\mathbf{c}$, transformations $T_{k,j}$ and

bias fields $B_{j,k}$, as the simulated slices $S_{k,j}$ depend on all of these learnable variables. To further suppress magnification of the background noise while solving the inverse problem, we also incorporate a total variation spatial regularisation loss $\mathcal{L}_R$ [38] on the diffusion signal V. Comprehensive mathematical definitions and implementation details of $\mathcal{L}_R$ are provided in the supplementary material.

Using this approach we reconstruct the volumetric diffusion signal, while also jointly correcting for degradation during data acquisitions. Specifically, we correct fetal motion by aligning acquired slices to the anatomical space using transformation $T_{k,j}$, compensate for blurring due to limited and potentially anisotropic resolution of acquired data by incorporating PSF $P_{k,j}$, correct MR acquisition artefacts through estimation of bias fields $B_{k,j}$; and perform outlier rejection through slice- and voxel-wise uncertainty estimation.

2.5 Through-Plane Continuity Regularization

Unlike structural fetal MRI which routinely acquires multiple orthogonal stacks of the same contrast to constrain the reconstruction of 3D volume, diffusion MRI requires acquisition of multiple stacks to sample numerous gradient directions. Consequently, data is typically acquired along a single slice orientation. This makes correction of intensity artefacts (i.e. estimation of bias fields $B_{k,j}$) highly ill-posed, resulting in stripe artefacts in the reconstructed volume (see Fig. 6 (a) and (b)).

To mitigate the stripe artifacts, we introduce a novel Through-Plane Continuity (TPC) regularization term $\mathcal{L}_T$ to penalize unnatural signal variations along the through-plane direction, since abrupt high-frequency changes along this specific axis are predominantly caused by acquisition artifacts rather than genuine anatomy. For each acquired voxel $x_{k,j,i}$ with stack-specific unit through-plane direction $\mathbf{n}_k$, we calculate the through-plane gradient magnitude of corresponding diffusion weighted volume $V(\mathbf{x}', \mathbf{g}'_{k,j})$ at transformed location $\mathbf{x}'_{k,j,i} = T_{k,j}^{-1}(\mathbf{x}_{k,j,i})$ in the transformed through-plane direction $\mathbf{n}_{k,j} = R_{k,j}^{-1}(\mathbf{n}_k)$ as

$$G_z = \frac{1}{|\mathcal{N}_z|} \sum_{\boldsymbol{\delta}_z \in \mathcal{N}_z} \frac{\left\| V\left(\mathbf{x}'_{k,j,i} + \boldsymbol{\delta}_z,\, \mathbf{g}'_{k,j}\right) - V\left(\mathbf{x}'_{k,j,i},\, \mathbf{g}'_{k,j}\right) \right\|}{\|\boldsymbol{\delta}_z\|_2}. \tag{9}$$

where $\mathcal{N}_z$ is a 1D through-plane neighborhood obtained by randomly sampling offsets $\boldsymbol{\delta}_z = u\,\mathbf{n}_{k,j}$, where u is a continuous scalar drawn from a uniform distribution $\mathcal{U}(-s, s)$ bounded by the physical slice thickness s.

To ensure we do not over-smooth genuine sharp diffusion signals, we also compute an average spatial gradient magnitude over all directions G_{all} as

$$G_{\text{all}} = \frac{1}{|\mathcal{N}_{\text{PSF}}|} \sum_{\boldsymbol{\delta} \in \mathcal{N}_{\text{PSF}}} \frac{\left\| V\left(\mathbf{x}'_{k,j,i} + \boldsymbol{\delta},\, \mathbf{g}'_{k,j}\right) - V\left(\mathbf{x}'_{k,j,i},\, \mathbf{g}'_{k,j}\right) \right\|}{\|\boldsymbol{\delta}\|_2}. \tag{10}$$

where $\mathcal{N}_{\text{PSF}}$ is set of offsets obtained by sampling the PSF centered around the observed voxel. An artifact is identified when the through-plane variation is

disproportionately large compared to the local 3D spatial context. We quantify this anomaly using a ratio defined as

$$R_{k,j,i} = \frac{G_{\text{all}} + \epsilon}{G_z}, \tag{11}$$

where ϵ is a small constant ensuring numerical stability. A high ratio indicates an abrupt intensity change primarily along the acquisition axis which is characteristic of a stripe artifact rather than true signal. Thus, the Through-Plane Continuity regularization loss is formulated as

$$\mathcal{L}_T = \frac{1}{N} \sum_{k,j,i} \text{ReLU}(\log(R_{k,j,i}) - \tau), \tag{12}$$

where N is the number of acquired voxels and τ is a margin threshold that determines the onset of penalization. This formulation smoothly discourages excessively high through-plane gradients relative to the local context and effectively suppresses stripe artifacts while preserving high-fidelity anatomical details.

3 Experiments

3.1 Datasets

Fetal Brain dMRI Dataset We conduct experiments on a fetal brain diffusion MRI dataset of nine healthy subjects, acquired on a 3T Philips Achieva scanner under two protocols with different spatial resolution and angular sampling. Three subjects were scanned using a standard clinical protocol with 32 diffusion-weighted volumes ($b = 750\,\text{s/mm}^2$), with an anisotropic resolution of $2 \times 2 \times 3.5\,\text{mm}^3$ and 1.75 mm slice overlap. Six subjects were scanned as a part of the developing Human Connectome Project (dHCP) [1, 3, 10, 11, 32]. We utilize the spin-echo volumes and the highest shell ($b = 1000\,\text{s/mm}^2$), yielding 80 diffusion directions with isotropic $2 \times 2 \times 2\,\text{mm}^3$ resolution.

Simulated Neonatal Brain dMRI Dataset Neonatal dMRI data were obtained from dHCP dataset. We selected 20 subjects with high-quality reconstructions [2], covering a post-menstrual age range of 26–44 weeks. For quantitative benchmarking with known ground truth, we used the fully reconstructed $b = 1000$ s/mm^2 shell at $1.5 \times 1.5 \times 1.5$ mm^3 as reference and simulated fetal-like acquisitions by downsampling it to $2 \times 2 \times 2$ mm^3, and applying inter-slice rigid motion with translations sampled from $\mathcal{U}(-3, 3)$ mm and rotations from $\mathcal{U}(-6, 6)^\circ$, together with simulated bias field, Rician noise, and motion-induced signal void artifacts.

Simulated Neonatal Brain dMRI Template Dataset To provide a cleaner ground truth for controlled benchmarking, we additionally perform simulation

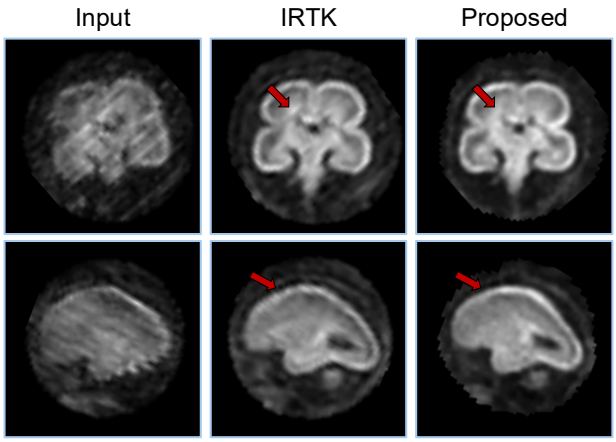

Fig. 3: Qualitative reconstruction results on fetal dataset. Left: motion corrupted input data. Middle: IRTK reconstruction shows residual striping artefacts (see red arrows). Right: Proposed method shows cleaner structure with no striping artefacts.

on a neonatal diffusion template constructed from dHCP data over 33–44 weeks PMA [25]. We use the template's $b = 1000$ s/mm^2 component as the reference signal, which is available at $1 \times 1 \times 1$ mm^3 isotropic resolution. To match the fetal-like setting, we apply the same corruption pipeline as in the subject-based simulation, which includes downsampling the template to $2 \times 2 \times 2$ mm^3, inter-slice rigid motion, a simulated bias field, Rician noise, and motion-induced signal void artifacts.

3.2 Implementation Details

We compare our proposed method against IRTK[3] [5], a classical iterative (non-deep-learning) approach for fetal diffusion MRI reconstruction. IRTK is executed on an Intel Xeon E5-2634 v3 CPU with multi-threaded acceleration (24 threads). Our framework is implemented in PyTorch [24], leveraging custom CUDA kernels via the `tiny-cuda-nn` library [22,23] to accelerate implicit neural representation learning. All neural network experiments are conducted on an NVIDIA Tesla V100 GPU with 32GB of memory. The model is trained for 6000 iterations by minimizing the total loss $\mathcal{L} = \mathcal{L}_D + \mathcal{L}_R + 0.5\,\mathcal{L}_T$, where the threshold parameter τ used in $\mathcal{L}_T$ is set to 80.

3.3 Comparative Evaluation and Results

We compare our method with IRTK on three datasets using both quantitative and qualitative evaluation. Reconstruction quality is measured using peak

[3] https://github.com/mariadeprez/irtk-simple

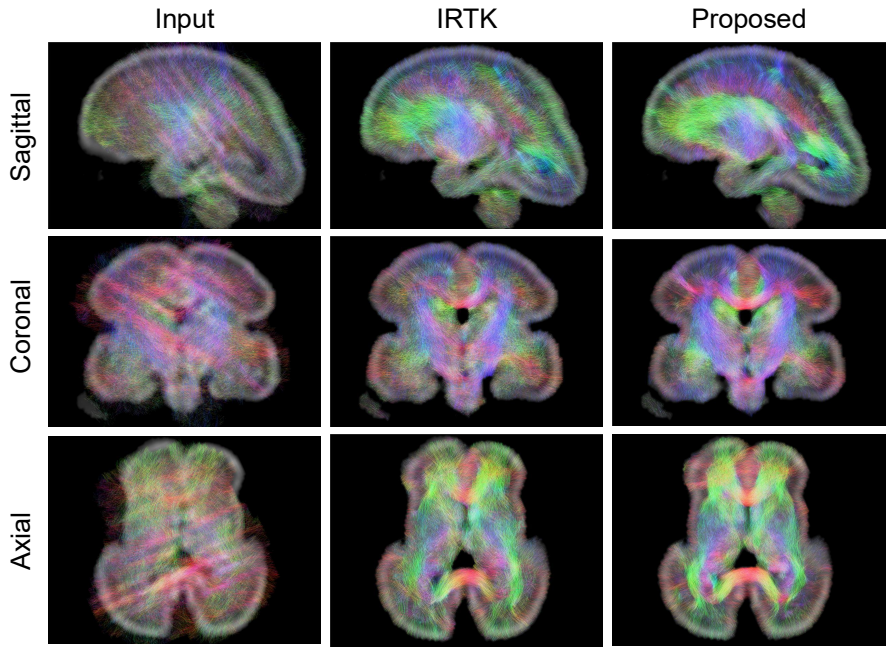

Fig. 4: Whole brain tractography of the fetus at 26 weeks GA. Left: motion corrupted input data result is tractography severly corrupted by motion artefacts. Middle: IRTK reconstruction shows less clear features and residual striping artefacts. Right: Proposed method shows cleaner structure with no striping artefacts.

signal-to-noise ratio (PSNR), structural similarity index measure (SSIM), normalized root mean square error (NRMSE), and normalized cross-correlation (NCC), which quantify signal fidelity, structural similarity, normalized error, and intensity correlation.

Results on Fetal Data Since in-vivo fetal datasets lack a 3D ground truth, we adopt a self-consistency evaluation protocol [5, 16] to assess reconstruction performance. For each subject, five stacks with minimal corruption are held out as a validation set, while the remaining stacks are used for reconstruction. Because the reconstructed volume and the held-out stacks still exhibit discrepancies caused by motion and bias field effects, both methods must estimate these acquisition-specific nuisance factors before quantitative comparison. First, the diffusion volume $V(\mathbf{x}, \mathbf{g})$ is reconstructed using training stacks (27 or 75). Next, the latent volume $V(\mathbf{x}, \mathbf{g})$ is fixed, the five held-out stacks are introduced, and their motion and bias field parameters are estimated. Using the frozen $V(\mathbf{x}, \mathbf{g})$ and the estimated nuisance variables, the held-out stacks are simulated using the forward model (eqn. 1,5,6) and compared with the corresponding acquisitions to evaluate reconstruction fidelity.

Table 1: Performance comparison with IRTK on three datasets. We show mean and standard deviation. The best performace is highlighted in bold.

Method	PSNR ↑	SSIM ↑	NRMSE ↓	NCC ↑	Run Time ↓
Fetal Data					
IRTK	23.34 (±0.98)	0.7516 (±0.1263)	0.1828 (±0.0274)	0.8085 (±0.0350)	8.77 (±4.93)
Proposed	**23.54 (±1.23)**	**0.9014 (±0.0246)**	**0.1788 (±0.0284)**	**0.8194 (±0.0374)**	**5.52 (±0.16)**
Simulated Neonatal Data					
IRTK	23.99 (±0.78)	0.9639 (±0.1000)	0.1520 (±0.0133)	0.7654 (±0.0540)	47.80 (±3.74)
Proposed	**25.20 (±0.72)**	**0.9828 (±0.0055)**	**0.1325 (±0.0105)**	**0.8412 (±0.0312)**	**6.75 (±0.11)**
Simulated Neonatal Template Data					
IRTK	24.74 (±1.24)	0.9554 (±0.0048)	0.1393 (±0.0055)	0.9469 (±0.0039)	102.15 (±2.86)
Proposed	**24.88 (±0.90)**	**0.9688 (±0.0040)**	**0.1379 (±0.0115)**	**0.9504 (±0.0090)**	**6.03 (±0.09)**

Table 1 summarizes the quantitative comparative results for the fetal dataset where our proposed method consistently outperforms IRTK across all evaluated metrics. Specifically, the proposed framework yields improvements across all metrics by achieving gains of 0.86%, 19.21%, and 1.34% in PSNR, SSIM, and NCC respectively while simultaneously reducing the NRMSE by 2.18%. The improvement in SSIM indicates better preservation of structural detail and sharper anatomical boundaries. Beyond reconstruction quality, our approach demonstrates superior computational efficiency and reduces the average reconstruction time by 37.05% per subject. The qualitative results presented in Fig. 3 compare the diffusion-weighted MRI signals $V(\mathbf{x}, \mathbf{g})$ synthesized from a randomly sampled diffusion gradient vector $\mathbf{g}$ and the reconstructed spherical harmonic volumes. These visualizations demonstrate that compared to IRTK, our method preserves more intricate anatomical structures while removing striping artefacts. Additionally, we present the effect of the reconstructed diffusion signal quality on downstream tasks by performing constrained spherical deconvolution [30, 31] followed by probabilistic tractography performed in MRTrix[4], see Fig. 4. We see that motion artefacts propagate into tractography. Both IRTK and our proposed method provide more consistent anatomically plausible white matter tracts, with the proposed methods showing cleaner and sharper features.

[4] https://www.mrtrix.org/

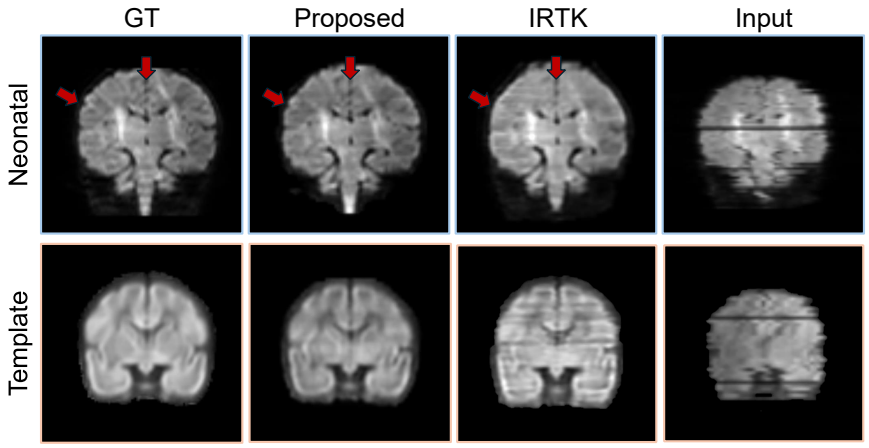

Fig. 5: Qualitative reconstruction results on two simulated neonatal datasets. Red arrows point to cortical folding

Results on Simulated Neonatal Data Simulated neonatal datasets were generated from clean ground truth neonatal datasets corrupted with fetal like motion and bias artefacts. The performance metrics were calculated by comparing the reconstructed and the ground truth volumes $V(\mathbf{x}, \mathbf{g})$ projected on the original gradient sampling scheme.

As reported in Table 1, our proposed method demonstrates superior performance on the neonatal dataset by achieving improvements of 5.04%, 1.96%, and 9.90% in PSNR, SSIM, and NCC respectively while yielding a 12.8% reduction in NRMSE. In addition, the proposed approach reduces computational cost, decreasing the average processing time by 85.87% relative to IRTK. All improvements were statistically significant.

Qualitative comparison of proposed and IRTK reconstructions is shown in Fig. 5 (top row) together with motion-corrupted stacks and reference volumes. The proposed method produces clearer anatomical structures without striping artefacts, whereas IRTK loses fine cortical folding patterns and exhibits stripe artifacts.

Results on Simulated Neonatal Template Data As an average across multiple subjects, the neonatal templates provide more stable image quality and serves as a cleaner reference for validation. Similarly to neonatal dataset, the performance metrics are calculated by comparison of reconstructed diffusion volumes to the ground truth templates at isotropic 1.0mm resolution.

The quantitative results presented in Table 1 demonstrate that our proposed method consistently outperforms IRTK across all metrics by achieving gains of 0.56%, 1.4%, and 0.36% in PSNR, SSIM, and NCC respectively while simulta-

Table 2: Quantitative ablation study on Spherical Harmonics (SH) order using fetal dMRI dataset. SH order 2 and 4 produces best results, with nearly identical performance with no statistically significant difference.

	Fetal Data			
SH Order	PSNR ↑	SSIM ↑	NRMSE ↓	NCC ↑
2	**23.59** (±**1.03**)	**0.9013** (±**0.0291**)	**0.1779** (±**0.0287**)	**0.8205** (±**0.0319**)
4	**23.54** (±**1.23**)	**0.9014** (±**0.0246**)	**0.1788** (±**0.0284**)	**0.8194** (±**0.0374**)
6	23.14 (±1.27)	0.8926 (±0.0277)	0.1871 (±0.0290)	0.7989 (±0.0454)
8	22.41 (±1.94)	0.8776 (±0.0359)	0.2045 (±0.0421)	0.7572 (±0.0900)

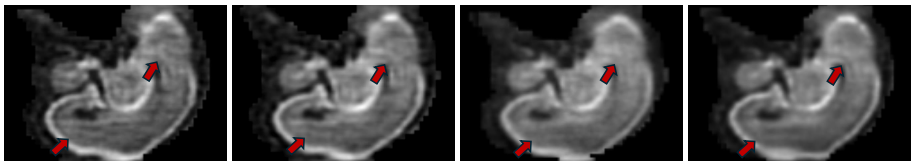

Fig. 6: Qualitative ablation study of the proposed artifact mitigation modules. (a) Without bias field network or TPC regularization. (b) Bias field network only. (c) TPC regularization only. (d) Both components. We can observe that our novel TPC regularisation is both essential and effective at removing the through-plane stripe artefacts.

neously reducing the NRMSE by 1.00%. Runtime is also reduced fifteen-fold, indicating higher scalability.

Qualitative comparisons in Fig. 5 (bottom row) shows that while both methods recover the major anatomical structures, IRTK exhibits noticeable stripe artifacts. The proposed method produces more uniform intensity and better preserves anatomical structures, resulting in improved reconstruction quality.

3.4 Ablation Study

Ablation experiments are conducted on the fetal dataset to evaluate the contribution of each component of the proposed framework.

Effect of Spherical Harmonic Order. We examine the influence of the spherical harmonic (SH) order L using $L \in \{2, 4, 6, 8\}$. As shown in Table 2, low-order models provide better performance, with $L = 2$ and $L = 4$ achieving nearly identical results across all metrics, with no statistically significant difference.

Table 3: Quantitative ablation study of the two artifact mitigation modules.

Compoments		Fetal Data			
Bias Field	TPC	PSNR ↑	SSIM ↑	NRMSE ↓	NCC ↑
✗	✗	22.77 (±1.31)	0.8898 (±0.0293)	0.1954 (±0.0297)	0.7841 (±0.0466)
✓	✗	23.25 (±1.24)	0.8964 (±0.0273)	0.1848 (±0.0286)	0.8071 (±0.0410)
✗	✓	23.07 (±1.29)	0.8949 (±0.0263)	0.1889 (±0.0301)	0.7972 (±0.0434)
✓	✓	**23.54** (±**1.23**)	**0.9014** (±**0.0246**)	**0.1788** (±**0.0284**)	**0.8194** (±**0.0374**)

Increasing the order to $L = 6$ and $L = 8$ degrades reconstruction consistency, reflecting overfitting to the noise and artefacts in fetal diffusion MRI, rather than fitting meaningful anatomical patterns. In practice we select 4th order SH because of its ability to encode fibre crossing.

Effect of Artifact-Handling Modules. We further analyze the roles of the bias-field network (Sec. 2.3) and the proposed TPC regularization (Sec. 2.5). Visual results in Fig. 6 show that removing both components leads to pronounced stripe artifacts and intensity inconsistency. Adding only the bias-field network reduces but does not remove striping artefacts. Using TPC regularization alone substantially suppresses through-plane stripe artifacts. The best visual quality and structural consistency are obtained when both components are enabled. These results highlight that our novel TPC regularisation is essential for correcting the through-plane stripe artefacts.

Quantitative results in Table 3 support these observations. Each module individually improves performance across all metrics, and their combination provides the largest gains. These findings indicate that modeling bias fields and enforcing through-plane consistency are both important for robust fetal dMRI reconstruction under severe motion.

4 Conclusion

In this paper, we present an implicit neural representation framework for spatially and angularly continuous reconstruction of fetal diffusion MRI. To handle severe motion and acquisition artifacts in in-vivo scans, we formulate a physics-based slice acquisition model that directly links the clean SH-based volume with the acquired motion-corrupted slices. The proposed framework jointly estimates the high-resolution diffusion volume, slice-dependent bias fields, and predictive uncertainty using three neural networks, while jointly optimizing rigid motion parameters with dynamic reorientation of the diffusion gradients. To further address the single acquisition direction used in fetal diffusion MRI, we introduce a

novel Through-Plane Continuity regularization that suppresses stripe artifacts by enforcing smoothness across the acquisition plane direction. Experiments on one real fetal dataset and two simulated neonatal datasets demonstrate that the proposed method consistently outperforms existing baselines in reconstruction fidelity and artifact suppression, while preserving fine anatomical structures relevant for downstream tractography and microstructural analysis.

It is worth noting that our framework acts as a natural extension of existing INR-based slice-to-volume reconstruction for structural MRI. By adapting our forward modeling strategy, we expect to extend our approach to recent paradigms like physics-informed networks [36] and 3D Gaussian splatting [4]. Future work will aim to rigorously evaluate the performance of these methods, and to find novel ways to boost both the fidelity and efficiency of dMRI reconstruction.

Acknowledgements

This research was supported by the European Research Council under the European Union's Seventh Framework Programme [FP7/ 20072013]/ERC grant agreement no. 319456 Dhcp project and the National Institute for Health and Care Research (NIHR) Clinical Research Facility at Guy's and St Thomas' NHS Foundation Trust. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. The authors gratefully acknowledge financial support from the China Scholarship Council.

References

1. Bastiani, M., Andersson, J.L., Cordero-Grande, L., Murgasova, M., Hutter, J., Price, A.N., Makropoulos, A., Fitzgibbon, S.P., Hughes, E., Rueckert, D., et al.: Automated processing pipeline for neonatal diffusion mri in the developing human connectome project. Neuroimage **185**, 750–763 (2019)
2. Christiaens, D., Cordero-Grande, L., Pietsch, M., Hutter, J., Price, A.N., Hughes, E.J., Vecchiato, K., Deprez, M., Edwards, A.D., Hajnal, J.V., et al.: Scattered slice shard reconstruction for motion correction in multi-shell diffusion mri. Neuroimage **225**, 117437 (2021)
3. Cordero-Grande, L., Christiaens, D., Hutter, J., Price, A.N., Hajnal, J.V.: Complex diffusion-weighted image estimation via matrix recovery under general noise models. Neuroimage **200**, 391–404 (2019)
4. Dannecker, M., Jia, S., Stolt-Ansó, N., Girard, N., Auzias, G., Rousseau, F., Rueckert, D.: Fast and explicit: Slice-to-volume reconstruction via 3d gaussian primitives with analytic point spread function modeling. In: Medical Imaging with Deep Learning
5. Deprez, M., Price, A., Christiaens, D., Estrin, G.L., Cordero-Grande, L., Hutter, J., Daducci, A., Tournier, J.D., Rutherford, M., Counsell, S.J., et al.: Higher order spherical harmonics reconstruction of fetal diffusion mri with intensity correction. IEEE transactions on medical imaging **39**(4), 1104–1113 (2020)

6. Descoteaux, M., Angelino, E., Fitzgibbons, S., Deriche, R.: Regularized, fast, and robust analytical q-ball imaging. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **58**(3), 497–510 (2007)
7. Frank, L.R.: Characterization of anisotropy in high angular resolution diffusion-weighted mri. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **47**(6), 1083–1099 (2002)
8. Gholipour, A., Estroff, J.A., Warfield, S.K.: Robust super-resolution volume reconstruction from slice acquisitions: application to fetal brain mri. IEEE transactions on medical imaging **29**(10), 1739–1758 (2010)
9. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR). pp. 770–778 (2016)
10. Hughes, E.J., Winchman, T., Padormo, F., Teixeira, R., Wurie, J., Sharma, M., Fox, M., Hutter, J., Cordero-Grande, L., Price, A.N., et al.: A dedicated neonatal brain imaging system. Magnetic resonance in medicine **78**(2), 794–804 (2017)
11. Hutter, J., Tournier, J.D., Price, A.N., Cordero-Grande, L., Hughes, E.J., Malik, S., Steinweg, J., Bastiani, M., Sotiropoulos, S.N., Jbabdi, S., et al.: Time-efficient and flexible design of optimized multishell hardi diffusion. Magnetic resonance in medicine **79**(3), 1276–1292 (2018)
12. Jakab, A., Kasprian, G., Schwartz, E., Gruber, G.M., Mitter, C., Prayer, D., Schöpf, V., Langs, G.: Disrupted developmental organization of the structural connectome in fetuses with corpus callosum agenesis. Neuroimage **111**, 277–288 (2015)
13. Kasprian, G., Brugger, P.C., Schöpf, V., Mitter, C., Weber, M., Hainfellner, J.A., Prayer, D.: Assessing prenatal white matter connectivity in commissural agenesis. Brain **136**(1), 168–179 (2013)
14. Kasprian, G., Brugger, P.C., Weber, M., Krssák, M., Krampl, E., Herold, C., Prayer, D.: In utero tractography of fetal white matter development. Neuroimage **43**(2), 213–224 (2008)
15. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Commun. ACM **60**(6), 84–90 (2017)
16. Kuklisova-Murgasova, M., Quaghebeur, G., Rutherford, M.A., Hajnal, J.V., Schnabel, J.A.: Reconstruction of fetal brain mri with intensity matching and complete outlier removal. Medical image analysis **16**(8), 1550–1564 (2012)
17. LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D.: Backpropagation applied to handwritten zip code recognition. Neural computation **1**(4), 541–551 (1989)
18. Mansfield, P.: Multi-planar image formation using nmr spin echoes. Journal of Physics C: Solid State Physics **10**(3), L55–L58 (1977)
19. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: Nerf: Representing scenes as neural radiance fields for view synthesis. Communications of the ACM **65**(1), 99–106 (2021)
20. Mitter, C., Prayer, D., Brugger, P.C., Weber, M., Kasprian, G.: In vivo tractography of fetal association fibers. PloS one **10**(3), e0119536 (2015)
21. Mitter, C., Prayer, D., Brugger, P.C., Weber, M., Kasprian, G.: In vivo tractography of fetal association fibers. PloS one **10**(3), e0119536 (2015)
22. Müller, T., Evans, A., Schied, C., Keller, A.: Instant neural graphics primitives with a multiresolution hash encoding. ACM transactions on graphics (TOG) **41**(4), 1–15 (2022)
23. Müller, T., Rousselle, F., Novák, J., Keller, A.: Real-time neural radiance caching for path tracing. arXiv preprint arXiv:2106.12372 (2021)

24. Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., Lerer, A.: Automatic differentiation in pytorch (2017)
25. Pietsch, M., Christiaens, D., Hutter, J., Cordero-Grande, L., Price, A.N., Hughes, E., Edwards, A.D., Hajnal, J.V., Counsell, S.J., Tournier, J.D.: A framework for multi-component analysis of diffusion mri data over the neonatal period. NeuroImage **186**, 321–337 (2019). https://doi.org/https://doi.org/10.1016/j.neuroimage.2018.10.060, https://www.sciencedirect.com/science/article/pii/S1053811918320329
26. Pontabry, J., Rousseau, F., Oubel, E., Studholme, C., Koob, M., Dietemann, J.L.: Probabilistic tractography using q-ball imaging and particle filtering: application to adult and in-utero fetal brain studies. Medical image analysis **17**(3), 297–310 (2013)
27. Rousseau, F., Glenn, O.A., Iordanova, B., Rodriguez-Carranza, C., Vigneron, D.B., Barkovich, J.A., Studholme, C.: Registration-based approach for reconstruction of high-resolution in utero fetal mr brain images. Academic radiology **13**(9), 1072–1081 (2006)
28. Sitzmann, V., Zollhöfer, M., Wetzstein, G.: Scene representation networks: Continuous 3d-structure-aware neural scene representations. Advances in neural information processing systems **32** (2019)
29. Tan, J., Zhang, X., Qing, C., Yang, C., Zhang, H., Li, G., Xu, X.: Fetal mri reconstruction by global diffusion and consistent implicit representation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 329–339. Springer (2024)
30. Tournier, J.D., Calamante, F., Connelly, A.: Robust determination of the fibre orientation distribution in diffusion mri: non-negativity constrained super-resolved spherical deconvolution. Neuroimage **35**(4), 1459–1472 (2007)
31. Tournier, J.D., Calamante, F., Gadian, D.G., Connelly, A.: Direct estimation of the fiber orientation density function from diffusion-weighted mri data using spherical deconvolution. Neuroimage **23**(3), 1176–1185 (2004)
32. Tournier, J.D., Christiaens, D., Hutter, J., Price, A.N., Cordero-Grande, L., Hughes, E., Bastiani, M., Sotiropoulos, S.N., Smith, S.M., Rueckert, D., et al.: A data-driven approach to optimising the encoding for multi-shell diffusion mri with application to neonatal imaging. NMR in Biomedicine **33**(9), e4348 (2020)
33. Turner, R., Le Bihan, D., Maier, J., Vavrek, R., Hedges, L.K., Pekar, J.: Echo-planar imaging of intravoxel incoherent motion. Radiology **177**(2), 407–414 (1990)
34. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. Proc. Adv. Neural Inf. Process. Syst. **30** (2017)
35. Wu, J., Chen, L., Li, Z., Li, X., Sun, T., Wang, L., Wang, R., Wei, H., Zhang, Y.: 3d isotropic high-resolution fetal brain mri reconstruction from motion corrupted thick data based on physical-informed unsupervised learning. IEEE Journal of Biomedical and Health Informatics (2025)
36. Wu, J., Sun, T., Wang, L., Zhang, Y., Wei, H.: Saffron: A physics-informed, label-free self-supervised deep learning framework for fast and accurate 3d fetal brain mri reconstruction. IEEE Journal of Selected Topics in Signal Processing (2025)
37. Wu, J., Wei, H., Zhang, Y.: Svrmamba: Slice-to-volume reconstruction from multiple mri stacks with slice sequence guided mamba. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 39, pp. 8404–8412 (2025)
38. Xu, J., Moyer, D., Gagoski, B., Iglesias, J.E., Grant, P.E., Golland, P., Adalsteinsson, E.: Nesvor: implicit neural representation for slice-to-volume reconstruction in mri. IEEE transactions on medical imaging **42**(6), 1707–1719 (2023)

39. Xu, J., Moyer, D., Grant, P.E., Golland, P., Iglesias, J.E., Adalsteinsson, E.: Svort: Iterative transformer for slice-to-volume registration in fetal brain mri. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 3–13. Springer (2022)
40. Young, S.I., Balbastre, Y., Fischl, B., Golland, P., Iglesias, J.E.: Fully convolutional slice-to-volume reconstruction for single-stack mri. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11535–11545 (2024)
41. Zanin, E., Ranjeva, J.P., Confort-Gouny, S., Guye, M., Denis, D., Cozzone, P.J., Girard, N.: White matter maturation of normal human fetal brain. an in vivo diffusion tensor tractography study. Brain and behavior **1**(2), 95–108 (2011)