From Promise to Practical Reality: Transforming Diffusion MRI Analysis with Deep Learning Enhancement

XINYI WANG

THE UNIVERSITY OF SYDNEY

Lead Supervisor: A/Prof. Weidong Cai
Associate Supervisor: A/Prof. Chenyu Wang
Associate Supervisor: Dr. Mariano Cabezas Grebol

A thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

School of Computer Science
Faculty of Engineering
The University of Sydney

The research reported was supported by the award of Research Training Program International Scholarship to the PhD candidate

28 February 2025

Statement of Originality

I hereby affirm that, to the best of my knowledge, the content of this thesis is entirely my own work. This thesis has not been previously submitted for any academic degree or other formal recognition. Every source, reference, and dataset utilized in this work has been properly cited and acknowledged in accordance with established academic integrity and ethical guidelines. I have taken all necessary precautions to ensure the authenticity and originality of this work.

Xinyi Wang

February 2025

Acknowledgements

Over the past four years as a PhD candidate, I have been fortunate to receive immense support and encouragement from numerous individuals. This journey has been marked by significant challenges, including the disruptions brought about by the COVID-19 pandemic and the nearly two-year wait for my visa. Despite these obstacles, I am grateful to have finally reached this milestone. I would like to extend my sincere gratitude to all those who have supported and guided me throughout my PhD studies.

First and foremost, I would like to express my deepest gratitude to my principal supervisor, A/Prof. Weidong (Tom) Cai, for his unwavering support from my Honours study through to the completion of my PhD. I am profoundly grateful to him for providing me with the opportunity to step into the world of research, an experience that has not only shaped my academic path but also broadened my horizons. Beyond being an exceptional research mentor, he has been a source of guidance and encouragement in my personal life. His continuous encouragement and steadfast support, particularly during the challenges of the pandemic and the prolonged wait for my visa, have been instrumental in helping me persevere and maintain confidence in my research. It has been an immense privilege to be supervised by him from the very beginning of my academic pursuits, and I will always be thankful for his mentorship and guidance.

I extend my heartfelt gratitude to A/Prof. Chenyu (Tim) Wang for providing me with the opportunity and resources to explore the significance of diffusion MRI (dMRI) in the rapidly evolving field of neuroscience. His support in connecting me with collaborators, securing essential datasets, and offering guidance has been invaluable. Beyond his role as a supervisor, he has always been willing to provide advice and assistance, whether in academic pursuits or personal matters.

I am also sincerely grateful to Dr. Mariano Cabezas Grebol for his invaluable mentorship and guidance across all my research projects. His insights and expertise have played a crucial role in shaping my academic growth, and I deeply admire his professionalism and rigorous approach to research. I am also deeply thankful to Dr. Zihao Tang for his collaboration and support in all our projects.

My appreciation extends to Prof. Fernando Calamante for his consistently insightful advice in the field of fibre orientation distribution (FOD). His meticulous attention to detail has never failed to impress and inspire me. I am also grateful to all my co-authors and colleagues for their support, including Prof. Michael Barnett, Dr. Arkiev D'Souza, Dr. Sicong Tu, Dr. Dongnan Liu, and Sheng Chen, as well as other collaborators across various institutions who may not have been mentioned by name but whose contributions have been invaluable.

Eight years ago, my mother, Jun Tan, and my father, Xiaoqiang Wang, made the wise decision to send me to The University of Sydney. Their unwavering financial and emotional support throughout my undergraduate years laid the foundation for my academic journey. They have been my strongest pillars, and I sincerely wish them good health and prosperity in the years ahead. I am also deeply grateful to my cousin, Ling Tan, who has always been like an elder sister to me, providing invaluable advice on both my life choices and academic pursuits. As her daughters, Chengyi (Claire) Zhang and Ruoyin (Emma) Zhang, embark on their own educational paths, I wholeheartedly wish them success and fulfillment in their studies.

I would like to extend my heartfelt thanks to Yixin Dai, my companion in this journey of studying abroad. She has been my travel partner, my source of comfort, and my biggest supporter, tolerating and caring for me all these years. Without her encouragement and companionship, I might have given up long ago.

Stepping into research has profoundly changed my life, and I owe much of this transformation to Tiange Xiang, an exceptionally talented and capable researcher. He introduced me to the world of research, patiently guided me from the very beginning, and helped shape my work ethic and approach. I sincerely wish him all the best in his PhD journey and hope he finds great success in his future endeavors.

I am also grateful to Kunze Wang, whose warmth and support have helped me in many aspects of my life. He is a reliable and responsible guide, as well as a compassionate and understanding supporter. My personal growth would not have been possible without his unwavering help. I hope he continues to find happiness and success in his journey ahead.

My heartfelt thanks also go to my roommate, Jiabao Gao, who took care of me and cooked for me when I was too occupied with my studies to do so myself. I wish her baby, Yunze Sun, every success and joy on his own learning journey.

To my high school and university deskmate, Heng Rao, I feel incredibly lucky to have shared this journey with you from high school to university. The bond we have built over the years is truly special.

I would also like to express my appreciation to Peizhen Guo, for always offering me comfort and emotional support, no matter the circumstances. I am also profoundly thankful to my grandparents, Shuiqing Man and Feng Guo, who relocated their entire family to Wuhan to accompany me during my time at The No. 1 Middle School Affiliated to Central China Normal University. Their presence meant a lot to me. And I'll never forget the delicious stir-fried cabbage my grandfather cooked—it remains one of my most cherished memories.

Before submitting my thesis, I visited Famen Temple in Xi'an and made three wishes. Famen Temple has brought me much good fortune along my academic journey, and I hope this time my wishes will also come true. After all, I believe that a girl who smiles often is bound to have good luck. I also wish my cousin in Xi'an, Yuzhu Liu, steady progress in her studies.

A special thank you to Shiyuan Huang, my strongest ally in the world of 'Honor of the King', for bringing me so much joy. I wish him a future filled with limitless possibilities and success, especially in this special year of the Snake.

There are many more family members, friends, and colleagues who have supported me along the way. I cherish the moments I shared with my dance partners from my undergraduate years, and I appreciate all my labmates for their camaraderie and encouragement. I hope that each of them finds a bright and fulfilling future.

Throughout my PhD, I have faced significant challenges, both academically and mentally. I am immensely grateful for the emotional support I received from those around me. This year, I also experienced the loss of my grandmother, Fumin Zhang, and my grandfather, Xiwan Tan, both of whom passed away in the Year of the Loong. I hope that, like the stars in the night sky, they continue to watch over me as I complete this chapter of my life. Just as in Coco, I will always remember them, ensuring that their legacy and love remain alive.

As I step into the Year of the Snake, I embrace a fresh start—ready to take control of my destiny, embark on new adventures, and live life to the fullest. Here's to new beginnings and the journey ahead!

Abstract

Magnetic Resonance Imaging (MRI) is an advanced, non-invasive imaging modality widely utilized in radiology for generating high-resolution anatomical and physiological images of the human body. Diffusion Magnetic Resonance Imaging (dMRI) is a non-invasive imaging technique that leverages the diffusion properties of water molecules to probe tissue microstructure, offering valuable insights into white matter architecture in both healthy and diseased brains. However, dMRI data often suffer from limitations such as low angular and spatial resolution, as well as artifacts arising from sub-optimal acquisition protocols, which can impede accurate Fiber Orientation Distribution (FOD) estimation and subsequent analyses.

Deep learning-based approaches have emerged as promising solutions to mitigate the limitations of dMRI data, including disruptions caused by acquisition artifacts and resolution constraints. Enhancing the angular resolution of dMRI data, while remaining within the practical limitations of clinical MRI scanners, remains a key challenge. One prominent method for FOD enhancement involves predicting single FOD estimates from patches of single-shell acquisitions, achieving performance comparable to multi-shell High Angular Resolution Diffusion Imaging (HARDI). However, dMRI also suffers from low spatial resolution, which introduces spatial blurring, further degrading FOD quality and affecting the reliability of downstream analyses. While recent advancements, such as FOD-Net, have demonstrated improvements in FOD estimation, they have primarily been evaluated on healthy subjects, leaving their broader clinical applicability underexplored. Although previous research has addressed these limitations individually, a unified approach capable of simultaneously tackling angular and spatial resolution constraints remains lacking. Developing such an integrated solution is essential to advancing dMRI enhancement methods and ensuring their applicability in clinical settings.

This thesis aims to bridge the gap between the potential of deep learning methods and their practical application in clinical neuroscience, particularly concerning dMRI data. It addresses several key challenges that hinder the widespread clinical adoption of deep learning methods in dMRI analysis, including computational inefficiencies, the lack of comprehensive clinical evaluation, underexplored acquisition protocol dependencies, limited explainability, and the fragmented handling of data disruptions. Firstly, Chapter 2 introduces FOD-Net 2.0, an improved fully convolutional model that refines FOD angular resolution and is validated on both healthy and clinical datasets. It also investigates optimal acquisition protocols by analyzing the trade-off between gradient direction count and enhancement quality. Chapter 3 extends this work by developing FastFOD-Net, a more practical and efficient solution, alongside a comprehensive clinical evaluation pipeline that assesses FOD enhancement at multiple levels, including fixel- and connection-wise analyses, and clinical applications. With the most comprehensive clinical evaluation to-date, our work demonstrates the potential of FastFOD-Net in accelerating clinical neuroscience research, empowering diffusion MRI analysis for disease differentiation, improved interpretability in connectome applications, reduced measurement errors to lower sample size requirements. Additionally, Chapter 4 explores uncertainty estimation in deep learning-based FOD enhancement, analyzing the reliability of FOD improvements and their impact on downstream tasks such as fiber bundle element estimation, thus enhancing model transparency. Chapter 5 then introduces UFREE (Unified FOD angular and spatial Resolution Enhancement framEwork), a novel two-stage learning strategy that unifies FOD enhancement across different kinds of resolutions, aligning low-resolution features with high-resolution priors to provide a generalized solution.

By introducing these robust, scalable, and clinically relevant deep learning solutions and establishing a comprehensive clinical evaluation platform, this work transforms the theoretical promise of these methods into practical reality, with significant implications for clinical neuroscience. The findings are analyzed from both technical and clinical perspectives, with an emphasis on clinical translation. Finally, future research directions are proposed to further integrate deep learning methodologies into real-world clinical applications.

List of Publications

$*$ means co-first author. The papers are arranged in chronological order, with the papers of which I am the first/co-first author, followed by the papers that are the result of collaborations.

Journal Papers

(P1) **X. Wang**, M. Barnett, F. Boonstra, Y. Barnett, M. Cabezas, A. D'Souza, M. Kiernan, K. Kyle, M. Law, L. Masters, Z. Tang, S. Tisch, S. Tu, A. Van Der Walt, D. Wang, F. Calamante, W. Cai, C. Wang. 'From Promise to Practical Reality: Transforming Diffusion MRI Analysis with Fast Deep Learning Enhancement', 2025. *Under Review*.

(P2) T. Xiang, C. Zhang, **X. Wang**, Y. Song, D. Liu, H. Huang, and W. Cai, 'Towards bi-directional skip connections in encoder-decoder architectures and beyond', *Medical Image Analysis*, vol. 78, p. 102420, 2022.
https://www.sciencedirect.com/science/article/pii/S1361841522000718

(P3) S. Aja-Fernández, C. Martín-Martín, Á. Planchuelo-Gómez, A. Faiyaz, M. N. Uddin, G. Schifitto, A. Tiwari, S. J. Shigwan, R. K. Singh, T. Zheng, Z. Cao, D. Wu, S. B. Blumberg, S. Sen, T. Goodwin-Allcock, P. J. Slator, M. Y. Avci, Z. Li, B. Bilgic, Q. Tian, **X. Wang**, Z. Tang, M. Cabezas, A. Rauland, D. Merhof, R. M. Maria, V. P. Campos, T. Santini, M. A. da Costa Vieira, S. HashemizadehKolowri, E. DiBella, C. Peng, Z. Shen, Z. Chen, I. Ullah, M. Mani, H. Abdolmotalleby, S. Eckstrom, S. H. Baete, P. Filipiak, T. Dong, Q. Fan, R. de Luis-García, A. Tristán-Vega, and T. Pieciak, 'Validation of deep learning techniques for quality augmentation in diffusion MRI for clinical studies', *NeuroImage: Clinical*, vol. 39, p. 103483, 2023.
https://www.sciencedirect.com/science/article/pii/S2213158223001742

Conference Papers

(P4) **X. Wang***, T. Xiang*, C. Zhang, Y. Song, D. Liu, H. Huang, and W. Cai, 'BiX-NAS: Searching Efficient Bi-Directional Architecture for Medical Image Segmentation', The 24th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2021.
https://link.springer.com/chapter/10.1007/978-3-030-87193-2_22

(P5) Z. Tang*, **X. Wang***, M. Cabezas, A. D'Souza, F. Calamante, D. Liu, M. Barnett, C. Wang, and W. Cai, ·Diffusion MRI Fibre Orientation Distribution Inpainting', The 25th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), The 13th International Workshop on Computational Diffusion MRI (CDMRI), pp65-76, 2022.
https://link.springer.com/chapter/10.1007/978-3-031-21206-2_6

(P6) **X. Wang**, Z. Tang, M. Cabezas, A. D'Souza, F. Calamante, D. Liu, M. Barnett, S. Tu, W. Cai, and C. Wang, 'FOD-Net 2.0: End-to-End FOD Enhancement for Low Angular Diffusion Acquisitions Using Deep Learning', The 32nd Annual Meeting of the International Society for Magnetic Resonance in Medicine (ISMRM Power Pitch), 2023.
https://archive.ismrm.org/2023/0952.html

(P7) **X. Wang**, Z. Tang, M. Cabezas, A. D'Souza, D. Liu, M. Barnett, F. Calamante, C. Wang, and W. Cai, 'Deep Learning-Driven Enhancement of Fibre Orientation Distribution: Effect of Choice of Gradient Direction Number', The 33rd Annual Meeting of the International Society for Magnetic Resonance in Medicine (ISMRM), 2024.
ISMRM 2024 Educational Stipend Award
https://archive.ismrm.org/2024/2165.html

(P8) **X. Wang**, M. Cabezas, Z. Tang, A. D'Souza, D. Liu, C. Wang, and W. Cai, 'Estimating Uncertainty on Deep Learning-Driven Fibre Orientation Distribution Enhancement', The 31st International Conference on Neural Information Processing (ICONIP Oral), 2024.
https://iconip2024.org/

(P9) **X. Wang**, Z. Tang, F. Calamante, W. Cai, M. Barnett, F. Boonstra, S. Kolbe, M. Law, H. Butzkueven, A. Van Der Walt, and C. Wang, 'Enhanced Diffusion Fibre Orientation Distribution for Patients with Multiple Sclerosis', The 40th Congress of European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS), 2024.
https://journals.sagepub.com/doi/10.1177/13524585241269219

(P10) **X. Wang**, K. Wang, Z. Tang, M. Cabezas, C. Wang, and W. Cai, 'UFREE: Unified Fiber Orientation Distribution Resolution Enhancement', 2025 IEEE International Symposium on Biomedical Imaging (ISBI), 2025.
https://biomedicalimaging.org/2025/

(P11) **X. Wang**, Z. Liu, S. Chen, H. Wang, C. Wang, M. Barnett, W. Cai, and Z. Tang, 'SdADC: Semi-supervised diffusion MRI artefact detection and classification', The 34th Annual Meeting of the International Society for Magnetic Resonance in Medicine (ISMRM), 2025.
https://www.ismrm.org/25m/

(P12) Z. Tang, **X. Wang**, L. Zhu, M. Cabezas, D. Liu, M. Barnett, W. Cai, and C. Wang, 'TW-BAG: Tensor-wise Brain-aware Gate Network for Inpainting Disrupted Diffusion Tensor Imaging', 2022 International Conference on Digital Image Computing: Techniques and Applications (DICTA), pp. 1-8, IEEE, November 2022.
https://ieeexplore.ieee.org/abstract/document/10034593

(P13) Z. Tang, **X. Wang**, M. Cabezas, L. Zhu, D. Liu, M. Barnett, W. Cai, and C. Wang, 'Reducing the impact of disrupted brain regions in Diffusion Tensor Imaging with

inpainting', The 32nd Annual Meeting of the International Society for Magnetic Resonance in Medicine (ISMRM), 2023.
ISMRM 2023 Educational Stipend Award
https://www.ismrm.org/23m/

(P14) S. Chen, Z. Tang, M. Cabezas, **X. Wang**, A. D'Souza, M. Barnett, F. Calamante, W. Cai, and C. Wang, 'Enhancing angular resolution via directionality encoding and geometric constraints in brain diffusion tensor imaging', The 31st International Conference on Neural Information Processing (ICONIP Oral), 2024.
ICONIP 2024 Best Student Paper Award
https://iconip2024.org/

(P15) S. Chen, Z. Tang, **X. Wang**, C. Wang, and W. Cai, 'Unsupervised dMRI artifact detection via angular resolution enhancement and cycle consistency learning', Australasian Joint Conference on Artificial Intelligence (AJCAI Oral), Springer Nature Singapore, pp. 417-431, 2024.
https://link.springer.com/chapter/10.1007/978-981-96-0351-0_31

(P16) Z. Tang, **X. Wang**, M. Cabezas, A. D'Souza, M. Barnett, F. Calamante, W. Cai, and C. Wang, 'Fiber Population-guided Pre-training for 3D Spatial Super-Resolution on Multimodal Brain Diffusion MR Imaging', Proceedings of the 6th ACM International Conference on Multimedia in Asia (MM Asia), 2024.
https://dl.acm.org/doi/abs/10.1145/3696409.3700288

(P17) Y. Chen, Z. Tang, T. Chen, **X. Wang**, Y. Sun, C. Tjoean, J. Lv, M. Barnett, F. Calamante, M. Irish, C. Wang, and W. Cai, 'Decoding task-based functional connectivity: Exploring Kolmogorov-Arnold networks for task-evoked fMRI classification', The 34th Annual Meeting of the International Society for Magnetic Resonance in Medicine (ISMRM), 2025.
ISMRM 2025 Educational Stipend Award
https://www.ismrm.org/25m/

Recognitions

Invitation to present ‘From Promise to Practice: Transforming Diffusion MRI Analysis with Deep Learning Enhancement’ at 2024 Brain and Intelligence Science Alliance (BISA) Young Scholar Forum, Shanghai, China.

Recipient of the Paulette Isabel Jones Career Development Award, University of Sydney, 2025.

Authorship Attribution Statement

Chapter 2 of this thesis is published as (P6) and (P7). I am the first author of these papers. I designed the method, conducted the experiments, analyzed the results, and wrote the manuscripts with the co-authors. The Preliminary study is published as (P3), which is about QUaD22 'Quality augmentation in diffusion MRI for clinical studies: Validation in migraine', a MICCAI CDMRI challenge. I am one of the authors of (P3). I attended this challenge, conducted and analyzed part of experiments and results with the co-authors.

In Chapter 3, part of the preliminary study on multiple sclerosis patients is published as (P9). About 80% of this part is from several sections in (P1), which is a method paper with the most comprehensive clinical evaluation to-date in dMRI enhancement and is under review. I am the first author of these publications. I designed the method and evaluation pipeline, conducted the experiments, analyzed the results, prepared visualizations, and wrote the manuscript with the co-authors.

Chapter 4 of this thesis is published as (P8). I am the first author of this paper. I designed the method, conducted the experiments, analyzed the results, and wrote the manuscripts with the co-authors.

Chapter 5 of this thesis is published as (P10). I am the first author of this paper. I designed the method, conducted the experiments, analyzed the results, and wrote the manuscripts with the co-authors.

In addition to the statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Student Name: Xinyi Wang

Date: 28 February 2025

As the supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Supervisor Name: Weidong Cai

Date: 28 February 2025

Glossary of Abbreviations

Abbreviation	Definition
5TT	Five-Tissue-Type
ACC / $\mathbf{r}_{\mathbf{Angular}}$	Angular Correlation Coefficient
ACT	Anatomically-Constrained Tractography
ALS	Amyotrophic Lateral Sclerosis
CC	Corpus Callosum
CGM	Cortical Gray Matter
CNN	Convolutional Neural Network
CSF	Cerebrospinal Fluid
CST	Corticospinal Tract
CSD	Constrained Spherical Deconvolution
DK	Desikan-Killiany
dMRI	Diffusion Magnetic Resonance Imaging
DT	Dystonic Tremor
DTI	Diffusion Tensor Imaging
DWI	Diffusion-Weighted Imaging
ET	Essential Tremor
FBA	Fixel-Based Analysis
FD / AFD	Fiber Density / Apparent Fiber Density
$\mathbf{E}_{\mathbf{FD}}$	FD Error / AFD Error
FOD	Fiber Orientation Distribution
FOV	Field of View
GM	Gray Matter
GT	Ground Truth

HARDI	High Angular Resolution Diffusion Imaging
HC	Healthy Controls
HCP	The Human Connectome Project
LARDI	Low Angular Resolution Diffusion Imaging
LDM	Latent Diffusion Model
MAE / $\mu\mathbf{E}_{\mathbf{Angular}}$	Mean Angular Error
MLP	Multilayer Perceptrons
MND	Motor Neurone Disease
MRI	Magnetic Resonance Imaging
MS	Multiple Sclerosis
MSE	Mean Squared Error
MSMT CSD	Multi-Shell Multi-Tissue Constrained Spherical Deconvolution
PD	Parkinson's Disease
PSNR	Peak Signal-to-Noise Ratio
$\mathbf{E}_{\mathbf{Peak}}$	Peak Error
ROI	Region of Interest
SGM	Subcortical Gray Matter
SIFT2	Spherical Deconvolution Informed Filtering of Tractograms 2
sMRI	Structural MRI
SS3T CSD	Single-Shell Three-Tissue Constrained Spherical Deconvolution
TE	Echo Time
TR	Repetition Time
VQ-VAE	Vector Quantized Variational Autoencoders
WM	White Matter

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

1.1 Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is an advanced, non-invasive imaging modality widely utilized in radiology for generating high-resolution anatomical and physiological images of the human body. By leveraging the principles of Nuclear Magnetic Resonance (NMR), MRI employs strong magnetic fields, magnetic field gradients, and radio frequency (RF) pulses to manipulate hydrogen protons within tissues, producing detailed images without the use of ionizing radiation. This distinguishes MRI from imaging techniques such as computed tomography (CT) and positron emission tomography (PET), which rely on X-rays or radioactive tracers.

MRI is extensively employed in hospitals and clinical settings for medical diagnosis, disease staging, and longitudinal patient monitoring. Compared to CT, MRI offers superior soft tissue contrast, making it particularly valuable for imaging structures such as the brain and abdominal organs. Examples are shown in Figure 1.1. However, MRI has limitations, including long scan durations, high acoustic noise levels, and the enclosed nature of conventional scanners, which can cause patient discomfort. Although open MRI systems have alleviated some of these issues, safety concerns persist for individuals with metallic implants or ferromagnetic materials, necessitating careful pre-screening protocols.

Originally known as Nuclear Magnetic Resonance Imaging (NMRI), MRI was renamed to avoid the negative connotations associated with the term 'nuclear' [2]. Hydrogen nuclei, abundant in biological tissues, are primarily used in MRI, allowing for effective mapping of

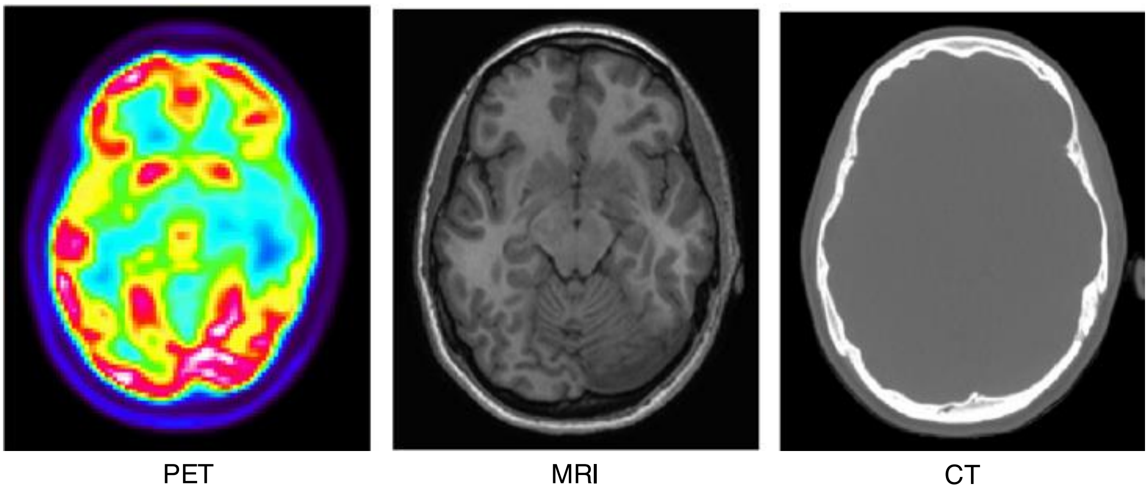


FIGURE 1.1. Reconstructed images from positron emission tomography (PET), Magnetic Resonance Imaging (MRI), and computed tomography (CT) scans of the same patient [1]. Images provided by University College London Hospital.

water and fat distribution. By applying RF pulses and magnetic field gradients, MRI spatially localizes signals and differentiates tissues based on their relaxation properties, enhancing contrast in diagnostic imaging.

Since its inception in the 1970s and 1980s, MRI has evolved into a versatile imaging modality. Beyond anatomical imaging, advancements such as diffusion MRI and functional MRI (fMRI) enable visualization of neuronal pathways and blood flow, broadening its applications in neuroscience and physiology. However, increasing reliance on MRI in healthcare raises concerns about cost-effectiveness and the potential for overdiagnosis [3, 4].

The key components of an MRI scanner include the main magnet, which polarizes the sample; shim coils, which correct inhomogeneities in the main magnetic field; a gradient system, which enables spatial localization of the scanned region; and an RF system, which excites the sample and detects the resultant NMR signal. Clinical MRI systems typically operate at 1.5 teslas (T), though field strengths range from 0.2T to 7T, with higher-field systems offering improved signal-to-noise ratios and resolution. Advanced research systems exceeding 14T are currently in conceptual or engineering design phases.

1.2 Brain Magnetic Resonance Imaging

MRI is a powerful diagnostic tool, offering superior soft tissue contrast compared to CT, enabling precise differentiation of brain structures. It minimizes artifacts, particularly in the brainstem and cerebellum, and provides high-resolution imaging of the pituitary gland [5, 6]. MRI is widely used to detect neurological conditions such as aneurysms, strokes, tumors, and traumatic brain injuries [7], as well as to differentiate neurodegenerative diseases like Parkinson's and Alzheimer's from other disorders [8, 9]. Recent advancements with deep learning solutions have demonstrated its remarkable success in automated tumor detection and classification [10].

1.2.1 Structural Magnetic Resonance Imaging

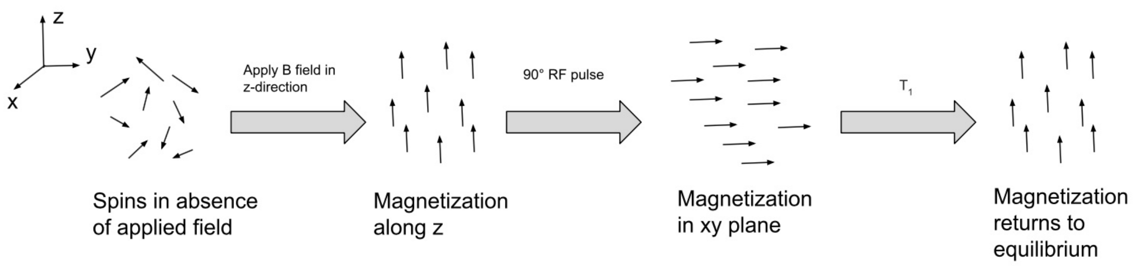


FIGURE 1.2. Diagram showing the variation in magnetization and spin orientations during the spin-lattice relaxation experiment [11].

Structural MRI (sMRI) focuses on visualizing and analyzing brain anatomy by detecting variations in the magnetic field of protons in soft tissues, leading to differences in image intensities across tissue types [13]. These images are influenced by water content, which can affect tissue differentiation, particularly in infants [14]. As a child matures, the reduction in water content and the increase in myelination enhance tissue contrast [15]. The primary brain tissue types—white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF), are critical for understanding brain development and neurological disorders [16].

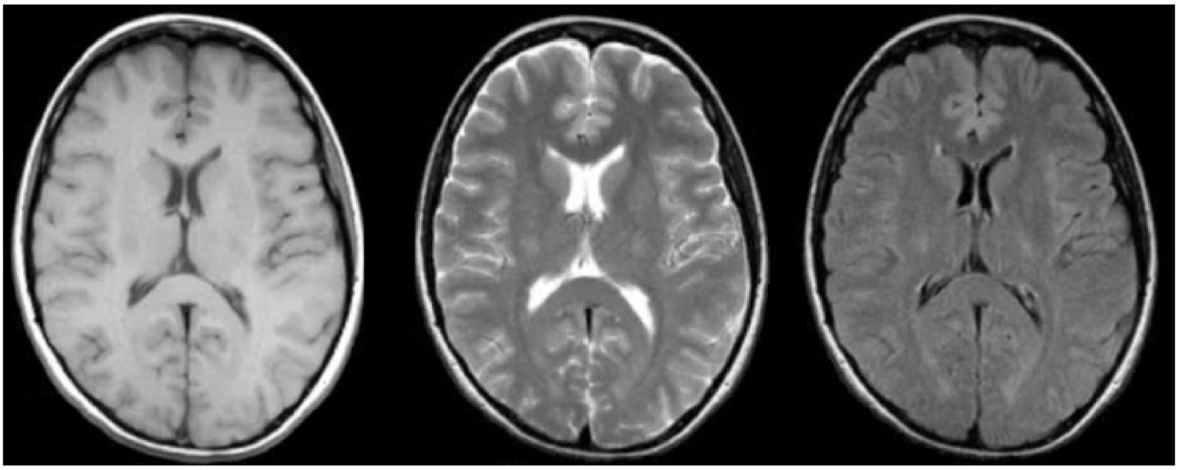


FIGURE 1.3. A conventional scan of a healthy subject, showing T1-, T2-, and FLAIR-weighted brain images from left to right [12].

sMRI images are created by varying the sequence of RF pulses. The key parameters influencing the contrast are repetition time (TR), the interval between RF pulses, and echo time (TE), the time between the RF pulse and the echo signal. MRI contrasts arise from variations in relaxation rates after perturbation by RF pulses. When a sample is placed in a magnetic field B_0, magnetization aligns along the z-axis. A 90° RF pulse flips the magnetization into the xy-plane, and after the pulse, the magnetization returns to equilibrium along the z-axis. This recovery time defines T1 (spin-lattice relaxation, Figure 1.2), with the relaxation rate $\frac{1}{T_1} = R_1$. The decay of transverse magnetization defines T2 (spin-spin relaxation), with the relaxation rate $\frac{1}{T_2} = R_2$.

In T1-weighted imaging, magnetization recovers before the MR signal is measured, making it ideal for visualizing the cerebral cortex and characterizing lesions, including post-contrast imaging. In T2-weighted imaging, magnetization decays before signal capture, which aids in detecting edema, inflammation, and white matter lesions. The T2 fluid-attenuated inversion recovery (FLAIR) sequence, with even longer TR and TE, suppresses CSF signal, improving the detection of abnormalities like White Matter Hyperintensities (WMH) by differentiating them from surrounding CSF [17]. The appearances of WM, CSF, and GM on T1-, T2-, and FLAIR-weighted images are shown in Figure 1.3.

1.2.2 Diffusion Magnetic Resonance Imaging

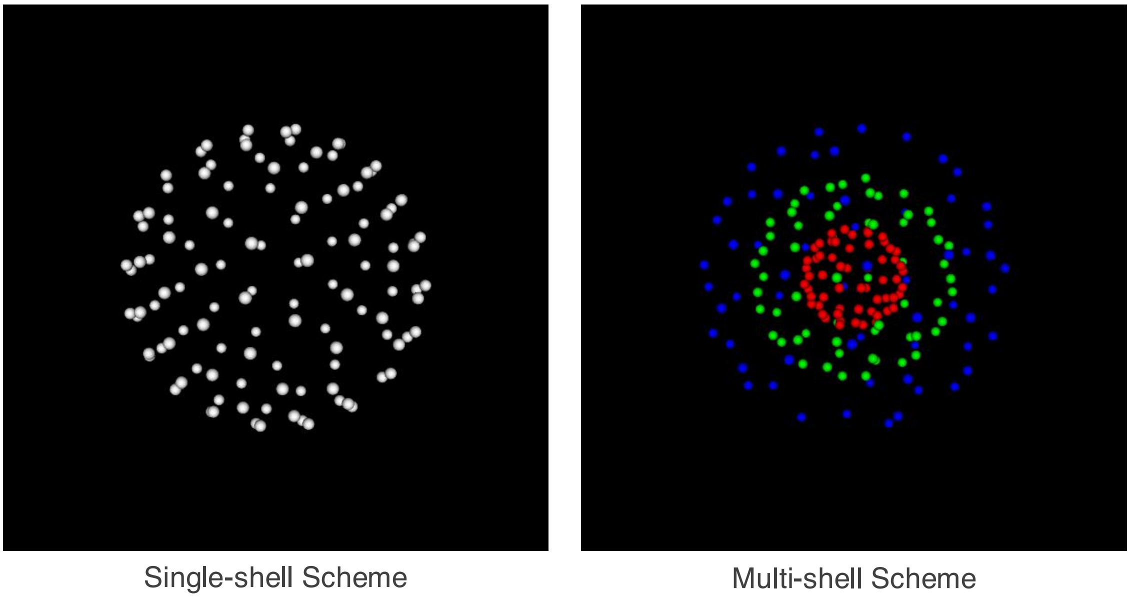


FIGURE 1.4. Single-shell v.s. multi-shell sampling schemes in dMRI [18]. The left image illustrates a single-shell sampling scheme, where diffusion measurements are uniformly distributed on a single spherical shell. The right image depicts a multi-shell scheme, consisting of multiple concentric shells (colored red, green, and blue), enabling richer microstructural characterization by capturing diffusion information at different b-values. Uniform distribution in both schemes ensures rotational invariance with respect to tissue structure and laboratory coordinates.

Diffusion Magnetic Resonance Imaging (dMRI) encompasses a suite of non-invasive imaging techniques that exploit the diffusion properties of water molecules to probe tissue microstructure. Among these, Diffusion-Weighted Imaging (DWI) is a common method that sensitizes MRI signals to the random motion of water molecules (Brownian motion) [19]. DWI is particularly useful for detecting abnormalities, such as in acute ischemic stroke, and for investigating the microstructure of white matter in both healthy and diseased brains. dMRI also includes advanced techniques like Diffusion Tensor Imaging (DTI), whole-brain tractography, and connectomics [20], which allow for the analysis of brain connectivity at global and regional levels. These methods have wide applications in neuroscience and clinical practice, where they enhance diagnosis, enable precise monitoring, and support surgical planning for various neurological conditions.

The mobility of molecules in tissue components could not be identical in all directions (anisotropic) due to the three-dimensional diffusion process. The occurrence of obstacles and mediums may regulate the displacement in some specific directions, resulting in anisotropy of diffusion. Since the utilization of magnetic field gradient pulses [21] can encode the diffusion process as MRI signals, the movements of molecules along the gradient directions are perceptible. Such variations of diffusion estimations with the gradient pulse directions changed can be observed to explore the anisotropy effect of diffusion. The advanced and popular application of DWI is to detect the anisotropy effects in brain white matter [22, 23]. Diffusion movements along the fiber direction are faster than that along the perpendicular direction. Thus, our assumption is that the orientation of fibers is demonstrated by that of the speediest diffusion [24].

There are two key parameters in DWI acquisition: the b-value and the b-vector. The b-value quantifies the diffusion-sensitizing factor, reflecting the strength and timing of the applied gradients during image acquisition. A higher b-value can be achieved by increasing the gradient amplitude and duration. However, the optimal b-value depends on various factors, including hardware constraints, the acquisition protocol, anatomical characteristics, and the type of pathology being investigated. The b-vector indicates the direction of diffusion sensitivity. In clinical practice, DWI is often obtained with fewer than 24 diffusion gradient directions and a single b-value (single-shell DWI) [25, 26]. Conversely, multi-shell DWI involves acquiring multiple b-values, each with a large number of gradient directions. This approach is typically necessary for advanced diffusion analysis, offering advantages over single-shell DWI [27]. The gradient directions are typically distributed uniformly, as shown in Figure 1.4. This distribution ensures that the sampling is orientation-invariant with respect to tissue structures or the laboratory coordinate frame.

While DWI sequences can be implemented in various ways, single-shot echo planar imaging (EPI) is the most commonly used approach due to its rapid acquisition, which enhances robustness against phase errors caused by patient motion [28]. However, EPI is susceptible to B0 field inhomogeneities and the induction of eddy currents when strong diffusion-weighted gradients are switched off [29]. Additionally, DWI is prone to various MRI artifacts, including subject motion and physiological noise [30, 31]. Pre-processing typically involves brain extraction [32, 33], denoising [34], and artifact corrections for Gibbs ringing [35], bias field inhomogeneities [36, 37], geometric distortions [38] , eddy currents [30], signal drift [39], and motion [40].

Diffusion is typically denoted by one scalar parameter D, which is termed as the diffusion coefficient. Attenuation A represents the diffusion effects on MRI spin-echo signals. A is based on D and b values in which b denotes the attributes of MRI gradient pulses. The relation can be formulated as equation 1.1:

$$\mathbf{A} = e^{(-b\mathbf{D})}. \tag{1.1}$$

Following acquisition, tissue microstructure can be modeled using various mathematical models, including DTI [41], and fiber orientation distribution (FOD) [42].

1.2.2.1 Diffusion Tensor Imaging

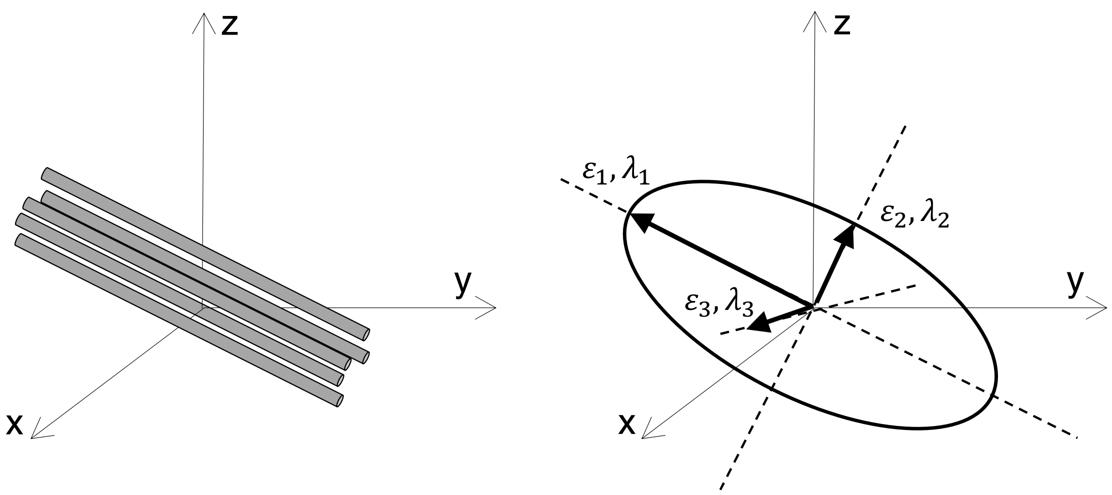


FIGURE 1.5. Left: Fiber tracts with arbitrary direction regarding to the coordinate system of scanner (x, y, z). Right: The ellipsoid models the three dimensional diffusivity whose shape and orientations are defined by eigenvalues and eigenvectors of the tensor model.

Diffusion Tensor Imaging (DTI) is widely used to extract microstructural information from DWI by leveraging tissue-specific diffusivity properties. The DTI model characterizes the primary directionality of water diffusion and its corresponding anisotropy [43, 44].

Since scalar parameters are insufficient to describe anisotropic diffusion, the tensor model $\mathbf{D}$ represents the mobility of water molecules and their interdependencies. This model is visualized as an ellipsoid, as shown in Figure 1.5, where the axes reflect diffusivity along different orientations, with the principal axis indicating the maximum diffusivity.

Mathematically, the tensor $\mathbf{D}$ is diagonalized as:

$$\mathbf{D} = \begin{bmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{bmatrix} = \mathbf{E} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \mathbf{E}^T \tag{1.2}$$

where $\mathbf{E}$ is a matrix of eigenvectors, and $\lambda_1, \lambda_2, \lambda_3$ contains the corresponding eigenvalues. The eigenvectors define the principal axes of the diffusion ellipsoid, while the eigenvalues quantify diffusivity along these axes. This transformation reorients the coordinate system from the MRI scanner's default x, y, and z axes to align with tissue diffusion properties.

The diffusion tensor $\mathbf{D}$ is a symmetric 3×3 matrix, with only six independent parameters. At least six diffusion-encoded measurements from different gradient directions are required for accurate estimation. The relationship between original signal intensity S_o and diffusion-weighted signal S_i in the presence of a diffusion-sensitizing gradient g_i is described by the Stejskal-Tanner equation [45]:

$$S_i = S_o e^{-\mathbf{b} g_i^T \mathbf{D} g_i}, \tag{1.3}$$

where **b** represents the diffusion weighting factor, incorporating gradient strength and pulse duration. The baseline image S_o (without diffusion weighting) corresponds to **b** $\approx 0\ s/mm^2$, producing a T2-weighted image, while diffusion-weighted images S_i typically have **b** $\approx 1,000\ s/mm^2$.

To solve for the six independent parameters of $\mathbf{D}$, at least seven images are required: six diffusion-weighted images (each from a different gradient direction) and one baseline image without diffusion weighting. Least squares fitting is commonly used for estimation. Increasing the number of encoding directions enhances model accuracy, though it requires balancing precision and acquisition time. In clinical practice, a higher number of gradient directions is generally preferred [21, 46].

To better understand the properties of the tensor, scalar metrics are derived from its three primary eigenvalues, ordered from largest to smallest. Commonly used metrics in clinical and neurological research include fractional anisotropy (FA), axial diffusivity (AD), mean diffusivity (MD), and radial diffusivity (RD), which characterize water diffusivity and facilitate group-wise comparisons [41, 47].

DWI summarizes tensor information into a single scalar or an RGB-based visualization, providing insight into tissue microstructure. FA measures relative anisotropy, while AD and RD quantify diffusion along and perpendicular to the principal fibre direction, respectively. MD, also referred to as the apparent diffusion coefficient (ADC), represents the average diffusion magnitude and is derived from the mean of the tensor's eigenvalues. Along with these metrics, MD plays a crucial role in studying neurological disorders such as multiple sclerosis (MS), amyotrophic lateral sclerosis (ALS), Alzheimer's disease (AD), Parkinson's disease (PD), and epilepsy [48, 49]. In pathological regions, FA typically decreases due to oedema, demyelination, gliosis, and inflammation [50, 51, 52], and MS patients often exhibit higher MD and lower FA in normal-appearing white matter [48, 53]. The sum of the eigenvalues, known as the trace, is another measure of diffusion, correlates with water movement at the voxel level, making it a valuable clinical marker for early stroke detection [48, 50].

1.2.2.2 Fiber Orientation Distribution

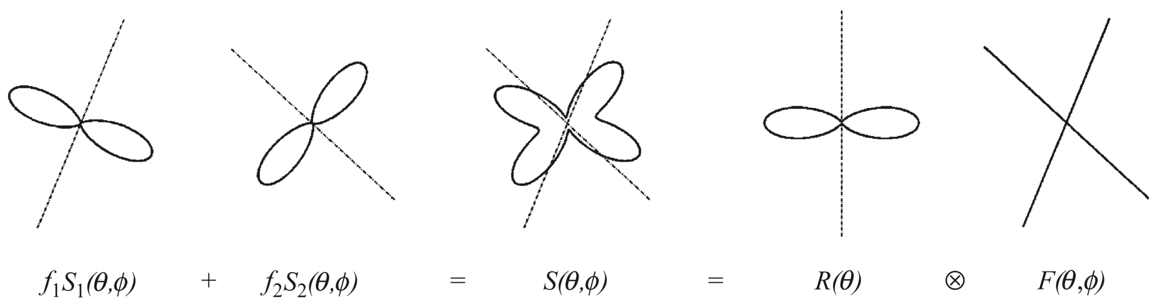


FIGURE 1.6. The voxel includes two fiber populations with different orientations (θ_1, ϕ_1) and (θ_2, ϕ_2). The dotted lines represent the fiber directions [54].

Most diffusion metrics discussed above rely on accurate estimation of the DTI model. However, this model is inherently limited in cases where multiple fiber populations exist within a single voxel [54]. The tensor model provides only a single principal orientation per voxel, making it insufficient for characterizing complex fiber structures. To overcome this limitation, Fiber Orientation Distribution (FOD), also known as orientation distribution function (ODF), [54, 55] is introduced, offering a means to estimate the the frequency distribution (or statistical function) of the fibre orientation angles within a voxel, regardless of the number of fiber populations.

This method assumes that all fiber populations share identical diffusion properties but differ in orientation. Figure 1.6 illustrates a case where two orientationally independent fiber populations ($N = 2$) coexist within a voxel. The diffusion-weighted signal attenuation, $S(\theta, \phi)$, is sampled at specific directions (θ_1, ϕ_1) and (θ_2, ϕ_2), where θ and ϕ represent the elevation and azimuthal angles in spherical coordinates. The fiber orientation distribution function $F(\theta, \phi)$ is then obtained through spherical deconvolution with a symmetric response function $R(\theta)$, which can be measured directly from regions of diffusion signals likely to include one single coherently directional fiber population:

$$F(\theta, \phi) = R^{-1} S(\theta, \phi). \tag{1.4}$$

1.2.2.3 Spherical Deconvolution

The spherical convolution operation is defined as the convolution of the kernel moving on the surface of the sphere and surface function [56]. In this context, the response function $R(\theta)$ is the kernel and the signal $S(\theta, \phi)$ is the function being convolved.

The spherical harmonic representation S^n of $S(\theta, \phi)$ at nth order can be formulated as $S^n = R^n F^n$, where F^n is a vector of length $(2n + 1)$ representing the n-th order spherical harmonic decomposition of $F(\theta, \phi)$. R^n is a $(2n + 1) \times (2n + 1)$ matrix corresponding to the rotational harmonic decomposition of $R(\theta)$.

The spherical convolution can thus be served as a sequence of matrix multiplications, and the spherical deconvolution can be obtained by performing an inversion of R^n to compute F^n. In terms of diffusion, the computation can be simplified further since components at the odd orders can all be set to zeros due to the symmetric diffusion procedure and only components at even orders need to be estimated.

Normally, the maximum harmonic order n_{max} is set to 8 to limit the number of independent samples. In this case, the number of parameters is 45. The harmonic coefficients at higher orders are more sensitive to noise in practice, which might be ignored to limit the negative effects.

1.2.2.4 Constrained Spherical Deconvolution

Constrained Spherical Deconvolution (CSD) builds upon the principles of spherical deconvolution, and the primary challenge in CSD is to solve for the FOD coefficients x that best match the observed diffusion signal, while enforcing a non-negativity constraint that ensures the FOD coefficients represent physically meaningful fiber orientations.

The CSD method formulates the problem as an optimization task. The goal is to minimize the error between the observed signal d and the predicted signal Cx, where C is a matrix linking the FOD coefficients to the measured diffusion signal:

$$\hat{x} = \arg\min_{x} \frac{1}{2} \|Cx - d\|^2 \quad \text{subject to} \quad Ax \geq 0, \tag{1.5}$$

where A is a matrix that enforces the non-negativity constraint on the FOD coefficients. When multi-shell diffusion data is available, CSD can be extended to incorporate measurements from multiple b-values:

$$\hat{x} = \arg\min_{x} \frac{1}{2} \left\| \begin{bmatrix} C_1 \\ \vdots \\ C_n \end{bmatrix} x - \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} \right\|^2 \quad \text{subject to} \quad Ax \geq 0$$

Through spherical decomposition and CSD, the final FOD image (i.e., FOD coefficients) is obtained, which can then be used for downstream analyses like fiber tractography and measuring fiber density in regions of complex fiber structure.

1.2.2.5 Fiber Bundle Element 'Fixel'

FODs are continuous functions on the sphere representing the complex fiber configurations, especially in regions where fibers cross or fan out in multiple directions. Following spherical deconvolution, the FOD in each voxel is represented as a set of coefficients in the spherical harmonic basis, but the orientations and integrals of these FOD lobes is of high interest to investigate the potential of generating high quality tractography of human brains. For clarity, the peak of a FOD is considered as the previse direction on the unit hemisphere for which the FOD amplitude is a maximum and the lobe is served as the angular spread of each peak.

Within each voxel of an FOD image, the smallest structural unit is a 'fixel', representing a discrete fiber bundle component. To generate a fixel-based representation, FOD images are sampled across 1281 basis directions on the unit hemisphere to capture distinct FOD lobes. A fast-marching level set algorithm assigns FOD samples to potential lobes, followed by peak refinement using Newton optimization. The final fixel representation encodes both orientation and spread, which can be used to estimate apparent fiber density [57, 58]. Each voxel can be divided into one or multiple fixels, where the fixel number represents the number of underlying white matter fibres within the voxel [59].

The resulting fixels f include not only the peak information $f^{\ \mathbf{Peak}}$ but also their orientation spread $f^{\ \tilde{\mathbf{d}}}$ and integral which can be then computed to reveal the apparent fibre density [57]. This fixel-wise measure $f^{\ \mathbf{FD}}$, more widely referred to as 'apparent fiber density (AFD)' or 'fiber density' (FD), serves as an indicator of WM microstructure with its value being approximately proportional to the total intra-axonal volume under specific conditions [57, 60].

1.2.2.6 Tractography and Connectome

Tractography is a technique used to estimate the trajectories of WM tracts from dMRI data, enabling the mapping of the brain's structural connectivity, particularly for specific pathways of interest [62]. Figure 1.7 shows the example tractography of the whole brain. The process involves tracking fibers from a seed point in both forward and backward directions. To

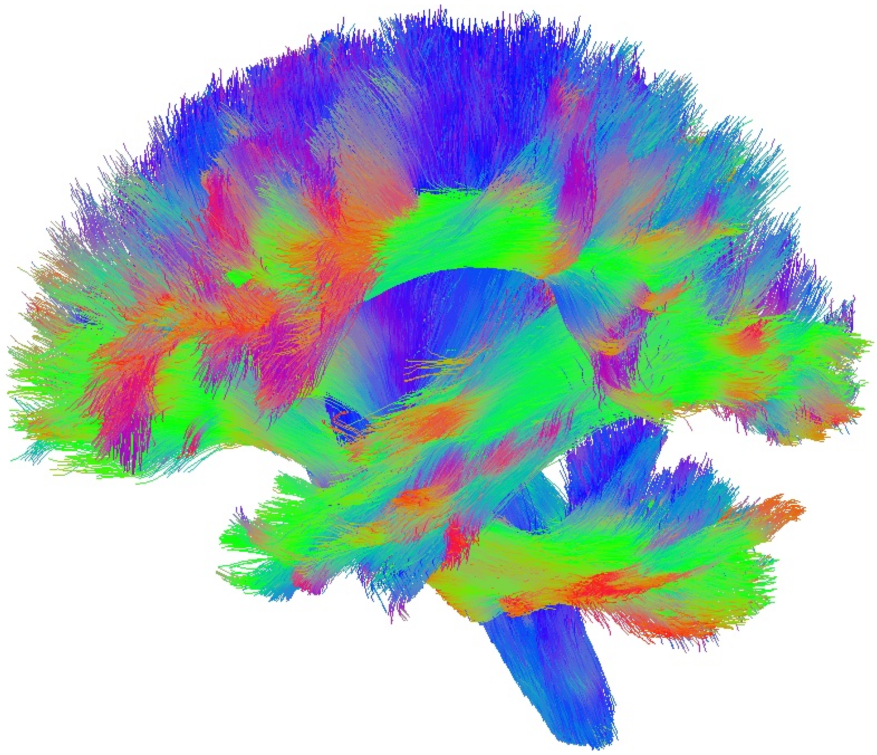


FIGURE 1.7. A visualization example of whole brain tractography [61].

enhance this, structural MRI (sMRI) is integrated to segment tissue into distinct types such as WM, GM, subcortical GM (e.g., amygdala and basal ganglia), CSF, and pathological tissue, resulting in a five-tissue-type (5TT) image [63], which is essential for subsequent tractography analysis [64].

Streamline-based tractography methods [65, 66, 50] are the most widely used, where the direction of the fastest diffusion (indicated by the major eigenvector) guides the tracking, and the resulting trajectories follow the eigenvectors. These methods typically consist of three main steps: correction of the tractography, identification of WM pathways, and quantification. The tractography process produces a tractogram, a collection of streamlines that represent the 3D reconstruction of WM fiber trajectories. Streamlines (i.e., each is formerd by a set of ordered points) are generated by following the diffusion directions and can be mapped using different mathematical models to capture the full range of WM connections.

DTI and FOD models are commonly used to generate tractography [65, 67]. However, FOD has a distinct advantage over DTI, as it can accurately represent crossing and intersecting fibers—an issue that DTI struggles with. Since approximately 90% of the brain's WM contains crossing fibers, FOD-based tractography provides more reliable results, whereas DTI tends to produce inaccurate fiber orientation estimates [68]. FOD is typically estimated using CSD, which yields more robust tractography through a probabilistic streamline algorithm [55, 58, 69].

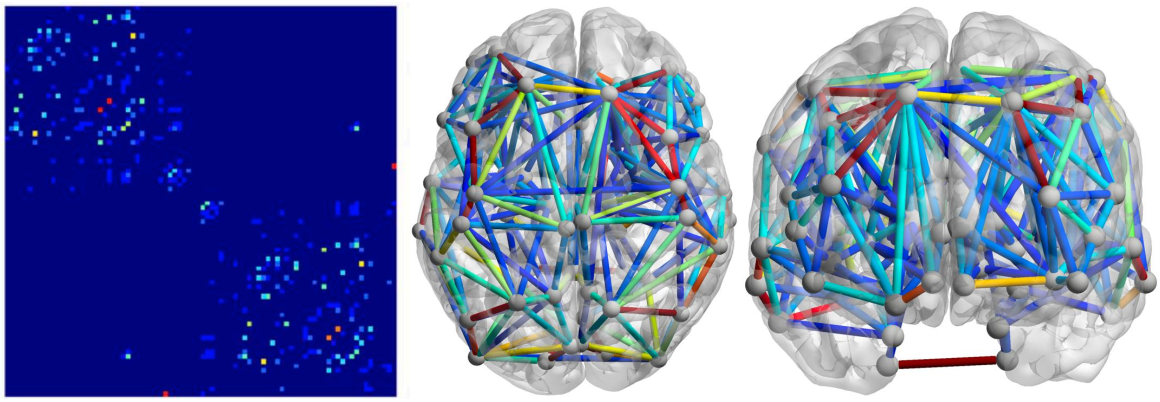


FIGURE 1.8. Whole-brain connectome visualization.

Connectome refers to a 2D brain connectivity matrix that is derived from tractography as illustrated in Figure 1.8, where both the rows and columns correspond to predefined regions of interest (ROIs) within the brain, with the selection of these ROIs determined by the specific atlas used. Desikan-Killiany (DK) atlas [70], which delineates 84 brain regions is a commonly employed atlas for constructing connectomes. Each element in the resulting matrix quantifies the strength of connectivity between the paired ROIs.

Probabilistic tractography [55, 64] uses dynamic seeding to generate millions of streamlines. Each streamline can be weighted using Spherical Deconvolution Informed Filtering of Tractograms 2 (SIFT2) [69], which integrates biological data into connectivity estimates [71]. A parcellation map can then be created based on DK atlas, with each integer representing a specific node, and voxels without assigned nodes are marked as 0. By assigning streamlines to these parcellation nodes, a structural connectome matrix is generated, revealing the brain's structural connectivity [63].

1.3 Deep Learning-based Image Enhancement

Image enhancement focuses on generating images by mapping a source image to a desired target domain, and is crucial for applications like inpainting [72, 73] and super resolution [74]. Recent advancements in deep learning have significantly improved the effectiveness of different tasks in image-to-image enhancement.

1.3.1 Deep Learning

Deep learning, a subset of machine learning, has gained immense popularity for its ability to learn directly from complex data without heavy reliance on manual feature engineering [75]. It is inspired by the signal-processing mechanisms of human brain, utilizing deep neural networks that process data through layers of inputs, weights, and biases.

The development of artificial neural networks dates back to 1979 with Fukushima's Neocognitron [72], and the term 'convolutional neural networks' (CNNs) was introduced in 1988. CNNs feature multiple hierarchical layers, including convolutional and pooling layers, and in 1989, LeCun demonstrated the practical application of back-propagation to learn convolution kernel coefficients from images of hand-written numbers [76], which significantly advanced image recognition and classification. However, early models faced limitations due to computational constraints, particularly in graphics processing unit (GPU) processing. The rise of powerful GPUs, notably from NVIDIA and supported by software like compute unified device architecture (CUDA) [77], allowed for more efficient training of deep learning models. This progress culminated in 2012 with AlexNet [78], which introduced key concepts like rectified linear units (ReLU) activation [79] and dropout [80], and further advancements came with VGG's deeper 16-19 layer networks in 2014 [81]. Later, also saw the introduction of U-Net [82], which revolutionized natural and medical imaging [83, 84, 85, 86, 87] analysis tasks with its encoder-decoder structure and skip connections.

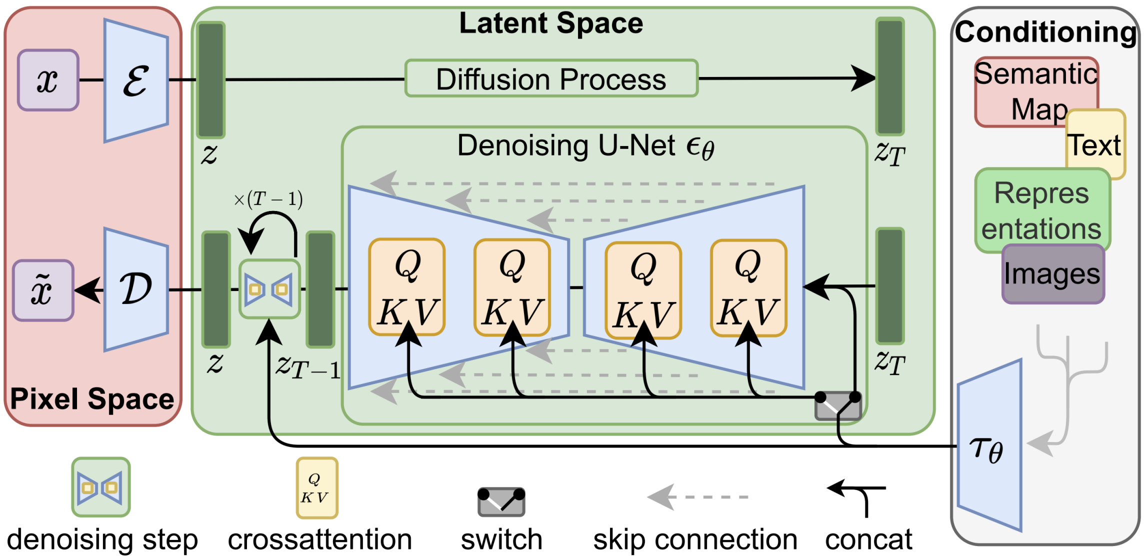


FIGURE 1.9. The overview of latent diffusion model (LDM) [88].

1.3.2 Natural Image Enhancement

Image enhancement has made remarkable strides with the advent of deep learning, particularly in tasks like inpainting and super resolution. Inpainting aims to restore missing or corrupted regions of an image by predicting pixel values based on surrounding context, maintaining texture and color consistency. super resolution, on the other hand, enhances the resolution of low-quality images, improving fine details and clarity.

Central to the progress in image enhancement are Generative Adversarial Networks (GANs), which have proven effective in image synthesis tasks. GANs consist of two networks: a generator that creates synthetic images and a discriminator that differentiates between real and generated images. The adversarial training process between these networks enables the generation of highly realistic images. Variants such as Deep Convolutional GANs (DCGAN) [89] and Pix2Pix, a conditional GAN [90], further refine this approach by integrating convolutional layers for better feature extraction.

Despite their success, challenges persist. For example, traditional convolutional methods for inpainting treat all pixels uniformly, which can be inefficient when handling corrupted areas. Partial Convolution (PConv) [73] improves upon this by using a binary mask to re-weight feature maps, focusing attention on corrupted regions. Gated convolution [91] improves this further by replacing the fixed binary mask with a learnable one, allowing the model to dynamically focus on corrupted regions and enhance the accuracy of inpainting.

Recent advancements in natural image synthesis have seen a notable shift towards diffusion models, which have demonstrated superior performance over GANs in generating high-quality, high-resolution images. Unlike GANs, which operate directly in the image space, Latent Diffusion Models (LDMs) encoding high-resolution features into discrete codes via Vector Quantized Variational Autoencoders (VQ-VAE), leading to more stable and controlled image generation.

The success of LDMs builds upon the use of VQ-VAE as the first stage of the process (Figure 1.9). VQ-VAE plays a crucial role by encoding high-resolution features into a discrete codebook, which is then used by the diffusion model to generate high-resolution images. FeMaSR (Feature Matching super resolution) [74] integrates VQ-GAN and implicit high-resolution priors to improve image restoration, particularly in the challenging task of blind super resolution. By using a pretrained codebook and decoder, it matches low-resolution inputs to high-resolution counterparts without needing explicit high-resolution references. This approach reduces artifacts and instability. Together, these advances in VQ-VAE and LDMs represent a new paradigm in image enhancement, delivering better results than previous approaches.

1.3.3 Diffusion Magnetic Resonance Imaging Enhancement

In recent years, DWI has gained significant attention in brain MRI research due to its ability to provide insights into white matter microstructure. However, DWI images are often compromised by pathologies [92], image distortions and artefacts [93], or sub-optimal acquisition protocols [94]. To address these challenges, inpainting techniques have been developed to reconstruct disrupted images. Additionally, low angular and spatial resolution in DWI acquisitions can hinder the precise capture of complex fiber bundle details. Consequently, improving DWI image quality through advanced computational methods has become a critical area of research.

Several deep learning-based methods have been proposed to address disruptions in dMRI data caused by artifacts or sub-optimal acquisition protocols. For example, the Tensor-Wise Brain-Aware Gate (TW-BAG) Network [94] inpaints disrupted regions in DTI images by estimating diffusion tensor coefficients based on valid information from undisrupted areas, enhancing DTI analysis by filling in missing data due to sub-optimal acquisition protocols and artefacts. Order-wise Coefficient Estimation Network (OCE-Net) [95] addresses disruptions in the FODs, which is crucial for accurate tractography. Using gating convolutions, OCE-Net reconstructs missing or corrupted FODs with unknown distribution, improving the continuity and accuracy of tractography even in regions affected by data loss or artifacts.

Enhancing the angular resolution of dMRI data within the constraints of clinical MRI scanners remains a significant challenge. Studies on DWI enhancement [96, 97, 98, 99, 100] have often encountered high computational complexity due to the diversity of acquisition protocols and model training requirements. A recent advancement, the DirGeo-DTI method, improves DTI metric accuracy by incorporating directionality encoding and geometric constraints [101]. The UdAD-AC framework automatically detects and corrects artifacts in dMRI data, using the DTI angular resolution enhancement task as a proxy [102, 103].

The limited angular resolution in low angular resolution images leads to the loss of essential information about complex fiber bundles, which reduces the accuracy of downstream analyses relying on FOD derivatives. One notable approach to enhance FOD data involves predicting single FOD estimates from patches of single-shell LARDI acquisitions [104]. This technique has shown promise in achieving enhancements comparable to multi-shell HARDI acquisitions. Despite its promising results, FOD-Net requires training for several weeks to converge to satisfactory results and inference can take hours for a single subject with a single GPU due to the voxel-wise nature of the predictions. Several subsequent studies [105, 106, 107] have built upon this method, including FOD-Net, which has focused primarily on healthy subjects and FOD metrics. However, the broader impact of these techniques in clinical research remains underexplored.

Low spatial resolution, meanwhile, introduces spatial blurring that could also degrade FOD quality. Another recent work [108] presented a fiber-population guided framework to address the low spatial resolution issue of DWI by directly applying super resolution on FODs to mitigate the erroneous differences in FODs and the associated derivatives.

1.4 Challenges

Despite the promising potential of deep learning in dMRI enhancement, several critical challenges need to be overcome to unlock its full clinical potential in practical reality: (1) the lack of a well-defined clinical rationale and inherent model design limitations, (2) insufficient large-scale validation across diverse neurological conditions, (3) unexplored acquisition protocol requirements, (4) limited interpretability and explainability, and (5) the absence of a unified framework for different kinds of resolution enhancement.

1.4.1 Challenge 1: Clinical Rationale and Method Design

A key challenge in the enhancement of dMRI is establishing a clear clinical rationale for the use of deep learning methods.

Results from OCE-Net [95] demonstrated substantial differences in the percentage of significant edges in the generated connectomes in empty ROIs. However, noisy ROIs only exhibited 2%. This suggests that FOD inpainting may not be critical in noisy regions. In contrast, enhancing angular resolution, as demonstrated by FOD-Net [109], significantly improves both tractography and connectome generation. Compared to DTI, enhancing FODs is more crucial, as it generates higher quality tractography and connectomes that better capture complex fiber structures. These findings underscore the focus of subsequent efforts on improving the angular resolution of FODs.

However, FOD-Net also has limitations, particularly its high computational cost. Despite promising results, FOD-Net requires weeks of training to achieve satisfactory convergence, and inference can take hours per subject, even with a single GPU, due to the voxel-wise nature of the training paradigm. Therefore, there is a pressing need for more efficient deep learning solutions that maintain result quality while addressing these computational challenges.

1.4.2 Challenge 2: Clinical Validation

While deep learning models have demonstrated promising results on healthy subjects, a significant barrier to their clinical adoption is the lack of an established validation platform. Currently, there is no standardized pipeline to evaluate generated FODs, tractography, and connectomes, limiting their use in clinical research. To be effectively integrated into clinical workflows, these models need to undergo rigorous clinical validation, ensuring they deliver reliable and actionable insights for clinicians.

1.4.3 Challenge 3: Acquisition Protocol Requirement

Another major challenge is determining the optimal acquisition protocols requirement for using deep learning techniques effectively. The specific acquisition parameters that best support angular resolution enhancement remain underexplored. Research focused on refining these protocols is needed to ensure that the data fed into deep learning models achieves the optimal trade-off, thereby maximizing model performance.

1.4.4 Challenge 4: Explainability of Deep Learning

The 'black box' nature of deep learning methods is another challenge, particularly in medical imaging where transparency is essential. To ensure clinicians can trust deep learning-generated results, it is important to explore model explainability and understand the patterns that the models are learning. Improved transparency will enable continuous method optimization.

1.4.5 Challenge 5: Unified Solutions for Multiple Resolutions

In addition to enhancing angular resolution, other challenges such as improving spatial resolution also play a significant role in the quality of dMRI enhancement. While previous research has addressed these issues individually, there is still a need for a unified approach that can address all of these resolution-related challenges simultaneously. Developing such a comprehensive solution would allow for more robust and effective dMRI enhancement methods, facilitating their application in clinical environments.

1.5 Thesis Contributions and Organization

The contribution and structure of this thesis can be summarized as below.

Chapter 2 introduces FOD-Net 2.0, an improved, fully convolutional extension of FOD-Net designed to enhance the angular resolution of FODs. This preliminary solution is applied not only to healthy subjects but also validated on clinical data, addressing the first challenge of clinical rationale and method design. This chapter also investigates the optimal acquisition protocols for utilizing these deep learning-based solutions, specifically focusing on the number of gradient directions. It analyzes the relationship between gradient direction count and FOD enhancement quality, addressing the third challenge by determining an optimal trade-off between acquisition protocol and model performance.

Chapter 3 advances FOD-Net 2.0 by introducing FastFOD-Net, a more practical solution to the first challenge outlined in the previous section. It also establishes a comprehensive clinical evaluation platform for FOD enhancement, encompassing FOD, fixel, and connection-wise assessments, addressing the second challenge. Additionally, this chapter explores clinical applications, such as fixel-based analysis between clinical groups, validating the real-world impact of deep learning methods and promoting the adoption of FOD enhancement in clinical neuroscience research using diffusion MRI. With the most extensive clinical evaluation to date, our work highlights FastFOD-Net's potential to accelerate neuroscience research, enhance disease differentiation, improve interpretability in connectome studies, and reduce measurement errors to lower sample size requirements.

In Chapter 4, enhancing FODs from clinical scans of low angular resolution is a complex task that is formulated as a multivariate regression problem. This is the first work to apply uncertainty techniques to uncover how improvements in FOD coefficients affect downstream tasks, such as fibre bundle element estimation. This approach enhances the reliability and trustworthiness of deep learning-driven enhancements in dMRI data and make the solution more white box, addressing Challenge 4.

Chapter 5 proposes the Unified FOD Angular and Spatial Resolution Enhancement Framework (UFREE), which estimates high-resolution (HR) FODs from any low-resolution (LR) FOD. UFREE employs a two-stage learning strategy: in the first stage, HR feature learning is shared across tasks, and in the second stage, LR FOD features are aligned with pre-trained HR FOD priors through latent feature matching. This solution directly addresses the fifth challenge by providing a unified method for resolution enhancement.

This thesis addresses the gap between the potential of deep learning methods and their actual impact in clinical neuroscience, particularly in the context of diffusion MRI data. Despite the promising capabilities of deep learning, several challenges hinder its widespread adoption, including inefficiencies, the lack of comprehensive clinical evaluation platforms, underexplored acquisition protocol requirements, limited explainability, and a fragmented approach to handling diverse data disruptions. Our work directly addresses these challenges by introducing methods that bring deep learning solutions from theoretical promise to practical reality. The significance of our contributions is examined both in technical and clinical contexts, with a particular focus on their potential for clinical translation. Finally, we propose directions for future research to build on these findings and further advance the integration of deep learning into clinical neuroscience.

CHAPTER 2

FOD-Net 2.0: Learning-driven FOD Enhancement for Low Angular Diffusion Acquisitions

---

Modern structural brain connectome pipelines and tractography techniques heavily rely on the quality of the diffusion weighted image acquisition (and angular resolution in particular) and the subsequent estimation of the FODs for each voxel. Generating reliable connectomes from low angular single shell acquisitions in clinical scenarios remains a challenging task. This chapter presents an end-to-end deep learning framework, termed FOD-Net 2.0, to enhance FOD estimates according to multi shell acquisitions from low angular single shell acquisitions to guarantee high quality tractography and connectomes within acceptable time and resources. This approach serves as a preliminary solution to **Challenge 1: Clinical Rationale and Method Design**, introduced in Section 1.4.1 from the previous chapter.

However, the enhancement capacity with a different number of gradient directions is not fully characterized. This chapter aims to explore the impact of initial gradient directions on FOD enhancement capacity of clinically accessible single-shell diffusion data (Figure 2.1). We employ a FOD enhancement framework on single-shell (b = 1000) data with different numbers of gradient directions. The enhanced FODs and derivatives are evaluated through FOD-based, fixel-based and connectomic metrics to explore the optimal trade-off between the learning-based FOD enhancement outcome and the choice of the number of gradient directions

In addition, this chapter provides guidelines for the optimal design of dMRI acquisition protocols meeting the expectations of clinical research on FOD enhancement according to the capability of learning-based frameworks, ensuring high quality tractography and

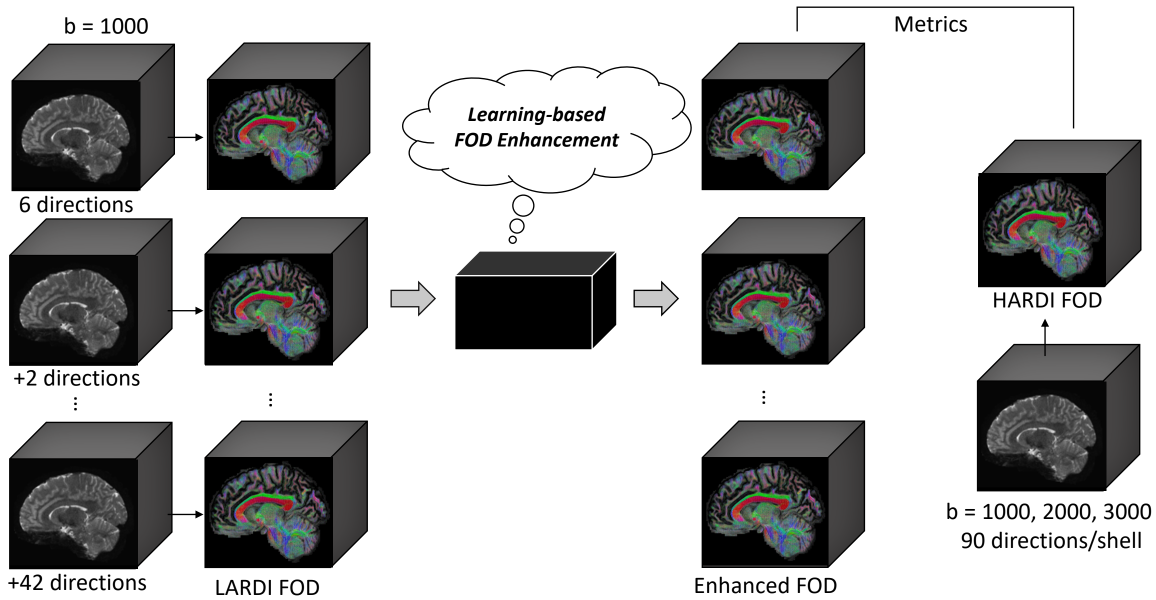


FIGURE 2.1. Preview of learning-driven FOD enhancement. Low angular FODs (left) with increasing gradient directions undergoes enhancement via a deep learning model (center), producing high quality FOD reconstructions (right). The results are compared against the reference high angular HARDI FOD (top right) for evaluation.

connectomes without the need for multi-shell high angular resolution diffusion imaging (HARDI) protocols. Our experiments and results directly address **Challenge 3: Acquisition Protocol Requirement**, as discussed in Section 1.4.3 from the previous chapter.

2.1 Introduction

dMRI techniques can map structural brain connectivity non-invasively with the construction of structural connectome matrices [110, 111], whose reliability is heavily dependent on the quality of the acquisition and the subsequent estimation of the FODs. Through CSD [27, 42, 54], high quality FOD estimates for complex fiber bundle configurations can be reliably computed from multi-shell acquisitions of high angular resolution diffusion imaging (HARDI), yielding state-of-the-art tractography [112]. However, the availability of multi-shell HARDI data is constrained by protocol limitations and the acquisition time required in clinical studies. Therefore, single-shell low angular resolution diffusion imaging (LARDI) is commonly acquired. However, such sequences cannot capture necessary information on complex fiber structures yielding to potentially misleading structural connectomes [113]. A previous work [104] on FOD enhancement for LARDI acquisitions focused on feeding patches cropped at each voxel to predict a single FOD estimate. Despite its promising results, FOD-Net requires training for several weeks to converge to satisfactory results. Furthermore, due to the voxel-wise nature of the predictions, inference can take hours for a single subject. Our work presents an end-to-end solution to mitigate time-consuming voxel-wise predictions by predicting the whole patch directly, thus making the super-resolution solution more practical applicable. We evaluated our method on a Human Connectome Project (HCP) dataset [114] and an in-house clinical dataset with FOD, fixel-based, and connectomic metrics to thoroughly analyze the impact of our enhancement approach.

Learning-based methods [104, 115] have shown promising results in FOD enhancement comparable to multi-shell HARDI acquisitions from limited single-shell. However, the effect of the number of directions on such enhancement frameworks remain unexplored, usually focusing on a fixed protocol. This study aims to bridge the gap between deep learning engineering and clinical practice by analysing the impact of gradient directions when enhancing FODs from clinical accessible single-shell LARDI acquisitions. We employ an end-to-end FOD enhancement framework [115] on the public HCP dataset [114], sampling different evenly distributed gradient directions, and thoroughly evaluate the capability of this framework in terms of FOD-based, fixel-based, and connectomic metrics.

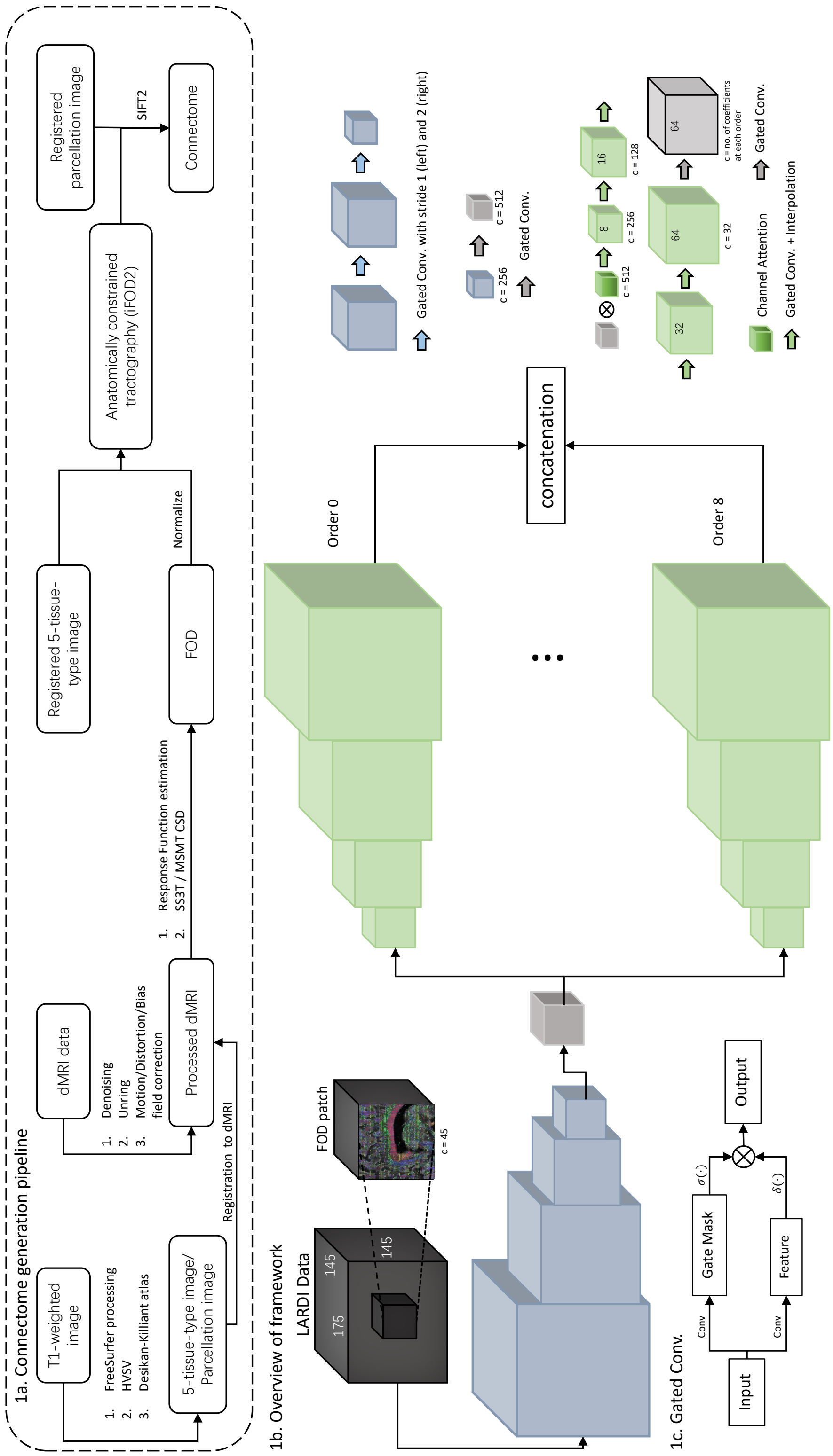


FIGURE 2.2. Illustration of an end-to-end FOD enhancement solution termed FOD-Net 2.0.

2.2 Methods

2.2.1 Dataset and Preprocessing

100 samples (50 training, 50 testing) from the HCP dataset were selected as ground truth from the HCP dataset, which were then subsampled to emulate LARDI data [104]. The in-house dataset contained 97 samples of 3 different clinical groups: 39 amyotrophic lateral sclerosis (ALS) subjects, 20 healthy controls (HCs), and 39 subjects with other motor neurone disorders. Figure 2.2a summarizes the preprocessing and connectome generation pipeline (using MRtrix) [116, 63]. The FODs were computed using the single-shell three-tissue (SS3T) [117] CSD and multi-shell multi-tissue (MSMT) [27] CSD routines, respectively, to obtain 45 spherical harmonic coefficients regardless of the gradient directions. We cropped and z-score normalized input patches of $64 \times 64 \times 64 \times 45$ with a sliding window of stride 32.

2.2.2 Architecture

Our network (Figure 2.2b) consists of a shared feature encoder and five order-wise decoders. The encoder adopts gated convolutions as the building blocks [91] (Figure 2.2c) to learn and highlight encoded features at different resolution levels. To analyze the dependencies between lower order coefficients and higher order ones [118], each decoder incorporates a channel attention mechanism [119]. The model was trained with mean squared error (MSE) loss and the Adam optimizer for 50 epochs on the HCP data only. We also tested the HCP trained model on the clinical dataset.

2.2.3 Requirement for Acquisition Protocol

As demonstrated in Figure 2.3, different numbers of gradient directions from 6 to 48 were sampled from the acquisitions using the Kennard-Stone algorithm [104, 115].

A total of 100 subjects were selected from the HCP dataset, split into 80 cases for training and 20 for testing. The mean b0 volume was then added to the sampled volumes to simulate a

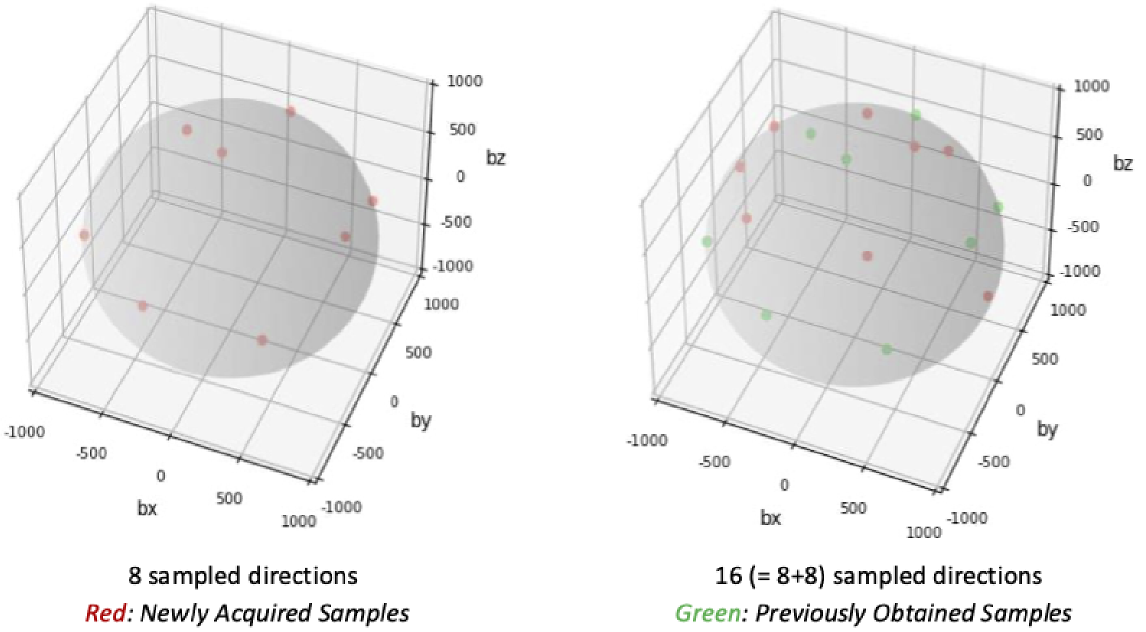


FIGURE 2.3. Data sampling. To simulate single-shell LARDI, we applied Kennard-Stone sampling to select varying numbers of gradient directions from the b1000 acquisitions, incorporating the mean b0 image.

single-shell LARDI acquisition. Figure 2.4a and 2.4c provide an overview of the preprocessing and connectome generation pipeline using Mrtrix [116, 63].FODs were computed using both SS3T [117] and MSMT [27] CSD routines, resulting in 45 spherical harmonic coefficients, regardless of the number of gradient directions. Input patches were cropped using a sliding window and were normalized by coefficient-wise z-score of the whole subject. We then applied an end-to-end learning-based framework (FOD-Net 2.0) (Figure 2.4b) to enhance the entire FOD patch [115].

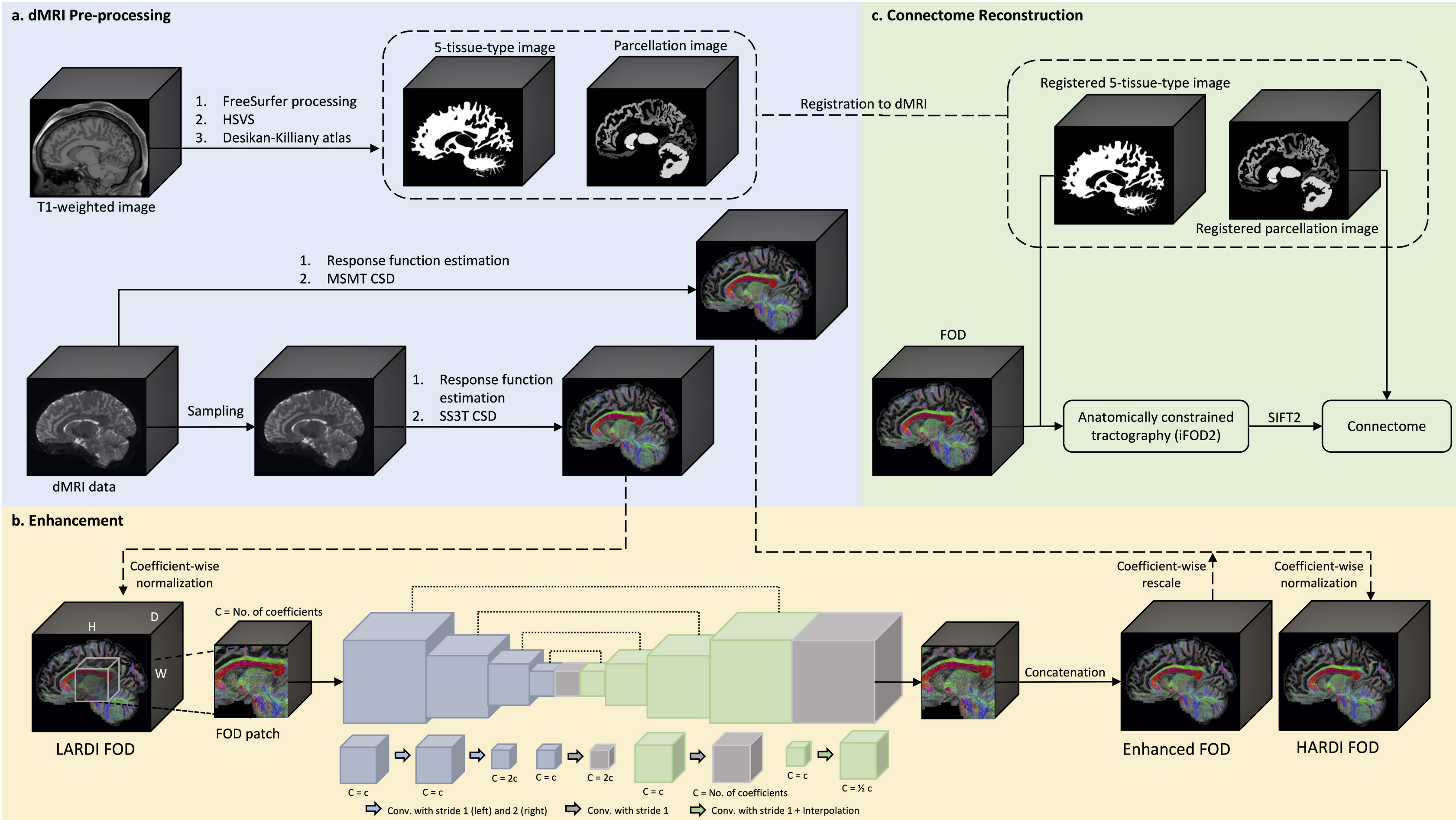


FIGURE 2.4. Overview of the FOD enhancement pipeline and its derivatives for acquisition protocol analysis using deep learning. a, dMRI pre-processing pipeline. b, FOD-Net 2.0. c, Connectome generation pipeline.

ROIs	SS3T	FOD-Net	FOD-Net 2.0
WM	0.7649	0.8948	**0.9293**
WM+CGM	0.5347	0.7832	**0.8611**
WM+SGM	0.4769	0.6417	**0.7513**

FIGURE 2.5. Mean angular correlation coefficient (ACC) on HCP datasets on pure or partial-volumed white matter voxels.

2.3 Results

In this chapter, we established an initial pipeline for evaluating the generated data from the perspectives of FOD, tractography, and connectomics. Section 3.2.5 in Chapter 3 further refines the metrics and addresses previously overlooked aspects.

2.3.1 Performance

Peak Signal-to-Noise Ratio (PSNR) and the angular correlation coefficient (ACC) were used to assess the quality and accuracy of the estimated FOD maps in regions with pure WM and those with GM partial volumes [104]. Figure 2.5 shows that our method obtains the best mean ACC 0.9293 approximated on the HCP dataset. Furthermore, our analysis on all the groups including ALS, HC, other, on the clinical dataset (Figure 2.6) demonstrates that our method is robust on both healthy subjects and patients from an unseen protocol during training.

To study fixel based metrics, we defined 3 regions of interest (ROIs) with different numbers of crossing fibers (1, 2 and 3) to analyze the mean angular error (MAE) of the fibers (important for tractography) [104]. Figure 2.7 and 2.8 show our method attains the minimum MAE in all ROIs on HCP, and ROI 1 and 2 on the clinical dataset.

Finally, to measure the accuracy of the reconstructed connectomes when compared to the reference multi-shell data, we computed the percentage of different edges (%) according to the t-test between methods using a false discovery rate (FDR) correction for multiple

Group	ROIs	SS3T	FOD-Net 2.0
Global	WM	0.7916	**0.8370**
	WM+CGM	0.6091	**0.7360**
	WM+SGM	0.5755	**0.6443**
ALS	WM	0.7895	**0.8360**
	WM+CGM	0.6045	**0.7339**
	WM+SGM	0.5714	**0.6388**
HC	WM	0.7975	**0.8371**
	WM+CGM	0.6264	**0.7408**
	WM+SGM	0.5951	**0.6561**
Other	WM	0.7909	**0.8381**
	WM+CGM	0.6048	**0.7357**
	WM+SGM	0.5695	**0.6439**

FIGURE 2.6. Mean angular correlation coefficient (ACC) on in-house clinical dataset between different groups.

ROIs	SS3T	FOD-Net	FOD-Net 2.0
ROI1	0.4217	0.3420	**0.2549**
ROI2	0.4387	0.3043	**0.2129**
ROI3	0.3955	0.2531	**0.2073**
WM	0.4324	0.3421	**0.2565**

FIGURE 2.7. Mean angular error (MAE) of regions of interest (ROIs) of different crossing fiber structures and white matter tissue on HCP dataset.

Group	ROIs	SS3T	FOD-Net 2.0
Global	ROI1	0.3764	**0.3265**
	ROI2	0.3055	**0.2781**
	ROI3	**0.2998**	0.3560
	WM	0.3757	**0.3256**
ALS	ROI1	0.3772	**0.3265**
	ROI2	0.3133	**0.2811**
	ROI3	**0.3026**	0.3606
	WM	0.3774	**0.3263**
HC	ROI1	0.3743	**0.3257**
	ROI2	0.2899	**0.2714**
	ROI3	**0.2913**	0.3395
	WM	0.3738	**0.3248**
Other	ROI1	0.3767	**0.3268**
	ROI2	0.3063	**0.2786**
	ROI3	**0.3016**	0.3606
	WM	0.3749	**0.3254**

FIGURE 2.8. Mean angular error (MAE) of regions of interest (ROIs) on clinical data between different clinical groups.

comparisons [104]. Our model decreases the number of significantly different edges by 40% on the HCP dataset (Figure 2.9). However, that improvement drastically decreases, although still significant, to less than 10% for all groups on the clinical dataset due to the differences in acquisition protocol. To further study the connectome differences, connectome disparity matrices (Figure 2.10, 2.12 and 2.11) were computed between the predicted connectomes and

HCP			
Metric	SS3T	FOD-Net	**FOD-Net 2.0**
Edge	75.50%	33.39%	**32.30%**

Clinical Data			
Metric	Group	SS3T	FOD-Net 2.0
Edge	Global	66.72%	**63.17%**
	ALS	56.17%	**50.14%**
	HC	39.67%	**34.02%**
	Other	55.65%	**50.46%**

FIGURE 2.9. Percentages of different (relative to reference multi-shell multi-tissue result) edges with FDR correction on HCP and clinical data. FOD-Net 2.0 can improve the connectome quality of single-shell with less different edges. The FOD-Net result on clinical data was omitted here due to the high computational cost of inference on relatively large cohorts. In later sections, we revisited this analysis by introducing FastFOD-Net, a significantly accelerated and enhanced variant of FOD-Net, designed for clinical scalability. Comparative results between methods are presented in Section 3.3.3.

the multi-shell ones [104]. It can be observed qualitatively and quantitatively that our model generated the connectome with the lowest disparity, outperforming the single shell estimates on both the HCP and private dataset (almost by half in terms of mean disparity).

Figure 2.9: It is unclear why the performance of FOD-Net was not reported for the clinical dataset. Please clarify whether it was omitted intentionally (e.g., due to limitations or unsatisfactory results), and consider including the results for completeness or explaining the rationale for exclusion.

HCP			
Metric	SS3T	FOD-Net	**FOD-Net 2.0**
Mean disparity	4.96E+02	1.85E+02	**1.84E+02**

Clinical Data			
Metric	Group	SS3T	**FOD-Net 2.0**
Mean disparity	Global	4.80E-01	**2.63E-01**
	ALS	4.89E-01	**2.61E-01**
	HC	4.36E-01	**2.68E-01**
	Other	4.92E-01	**2.63E-01**

FIGURE 2.10. Quantitative results of connectome disparity matrices. The SuperResolved (FOD-Net 2.0 estimates) generate connectomes with lower mean disparity on HCP and in-house clinical data.

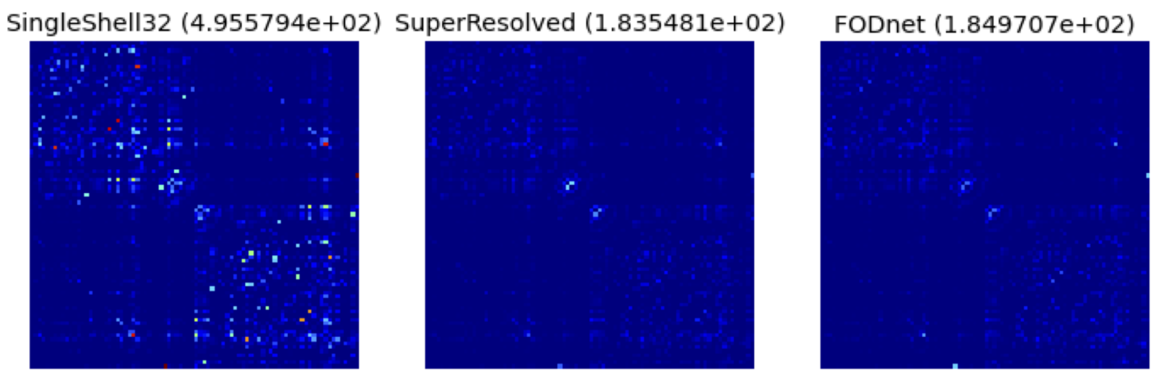


FIGURE 2.11. Visualization of connectome disparity matrices on HCP.

2.3.2 Effect of Choice of Gradient Direction Number

Figure 2.13 illustrates how the enhancement framework significantly enhances the quality and accuracy of FODs for all the sampled directions. With 20 or more gradient directions, it achieves a satisfactory mean angular correlation coefficient exceeding 0.9 in pure WM volumes, 0.8 in WM with cortical GM (CGM), and 0.7 in WM with subcortical GM (SGM), accompanied by a notable quality boost, although the improvement becomes more gradual at this stage.

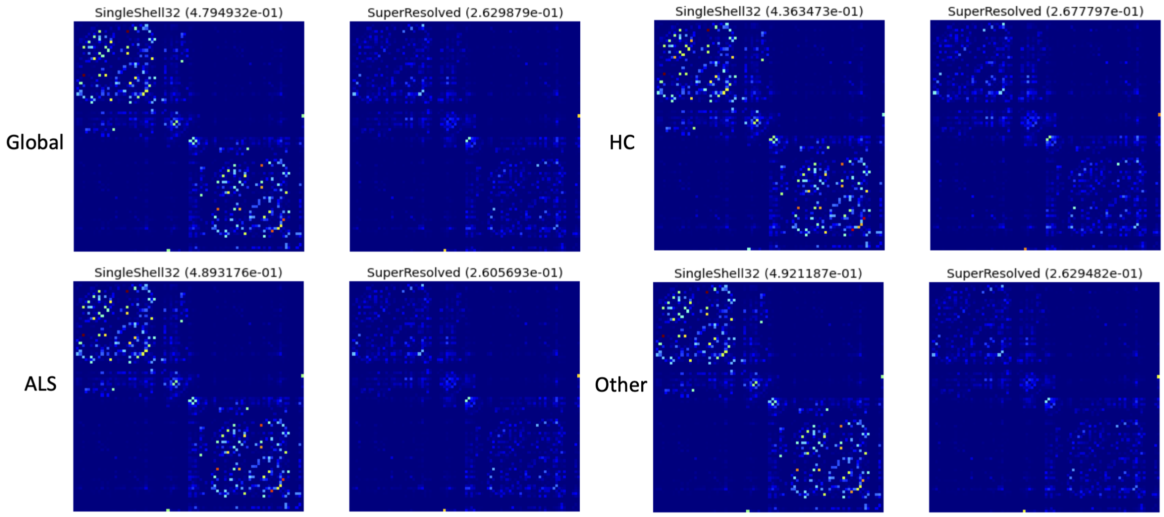


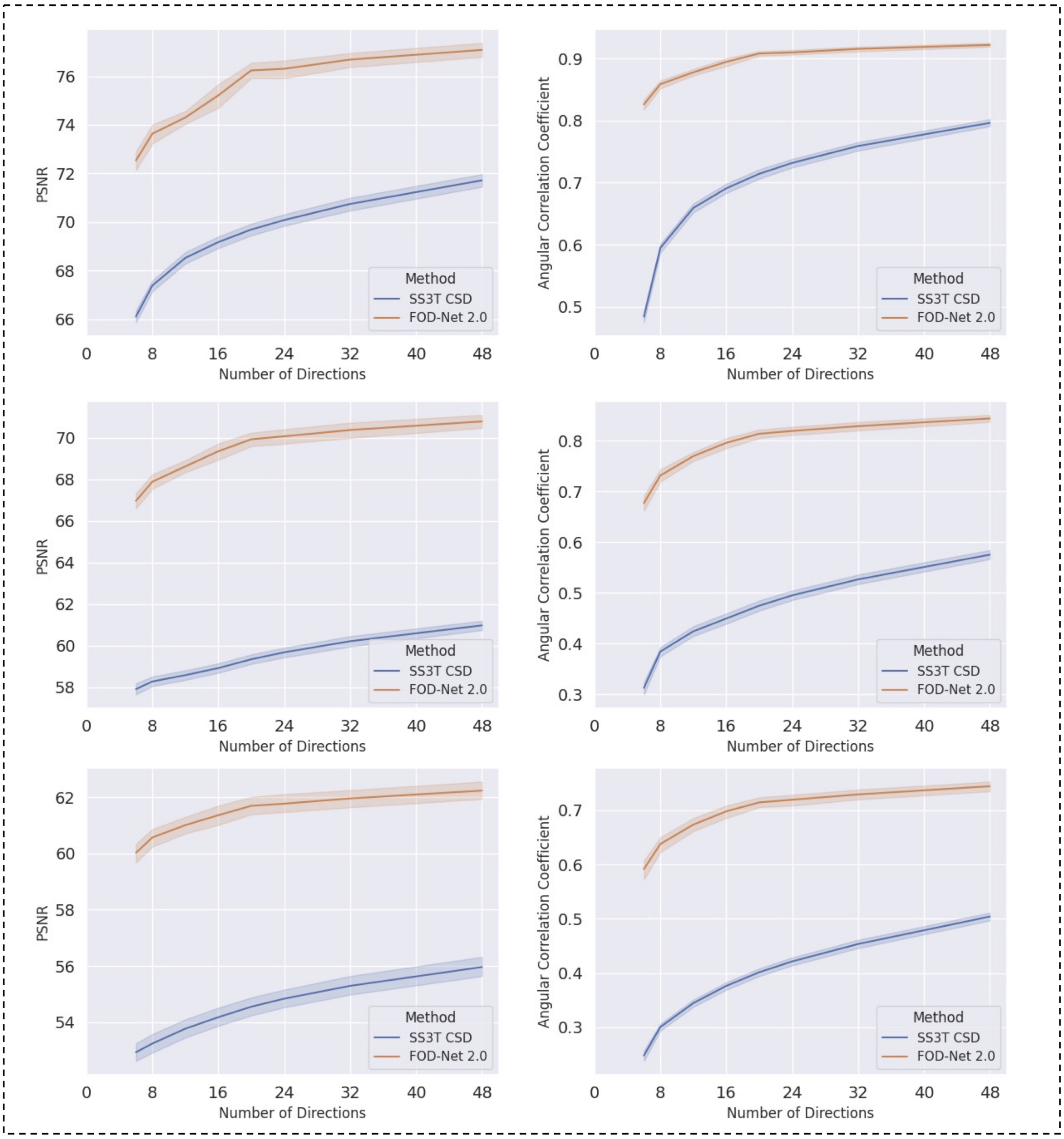
FIGURE 2.12. Visualization of connectome disparity matrices on clinical data.

Tractography reconstruction accuracy was also evaluated using fixel-based metrics in two ROIs with an increasing number of crossing fibers (2, and 3) [104, 115]. Notably, as the number of gradient directions reaches 20 in both ROIs, the mean angular error of enhanced FODs exhibits a substantial decrease (crucial for tractography [104]) (Figure 2.14a). Nonetheless, the AFD and peaks error show slower reductions, reaching a plateau after surpassing 24 gradient directions (Figure 2.14b,c).

Mean disparity matrices [104] were also computed to compare connectomes from all the sampled sets of directions. The Kendall's Tau rank correlation coefficient was used to measure ranking similarity of the connection strengths. When increasing the number of sampled directions, mean disparity (Figure 2.15a, c-e) reaches substantial quantitative and qualitative reductions while the ranking correlation coefficient (Figure 2.15b) shows significant improvements which both plateaued when surpassing a threshold of 24.

2.4 Discussions

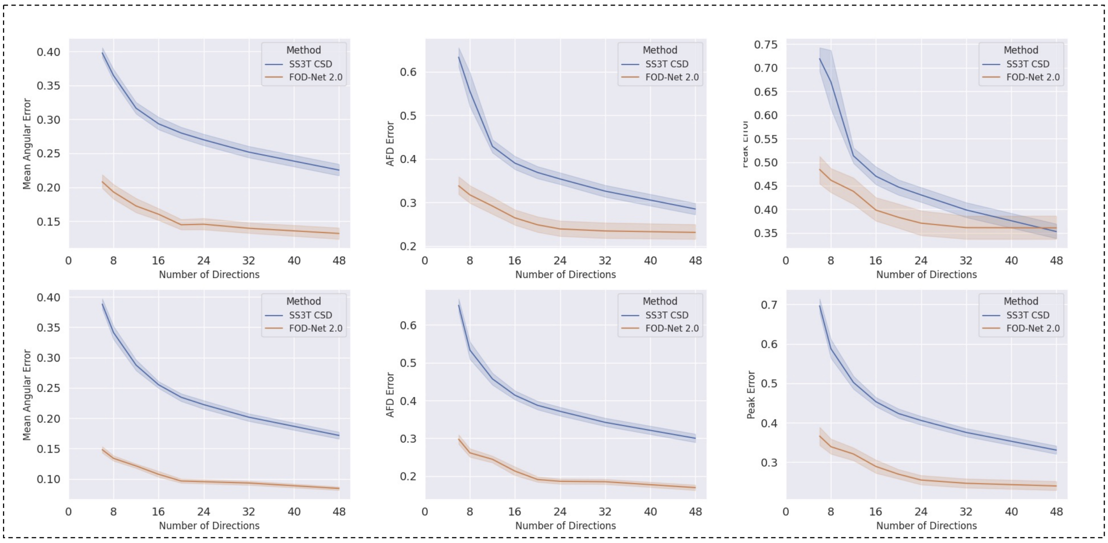
Our method can accurately enhance FOD from clinical single-shell acquisitions with a limited number of directions (directions = 32). Furthermore, the method can generalize to unseen



a. Peak Signal-to-Noise Ratio v.s. Number of Directions *b. Angular Correlation Coefficient v.s. Number of Directions*

FIGURE 2.13. All the FOD-Net 2.0 results (i.e., even the 6 directions case) give superior results to ALL the SS3T CSD results. a, The quantitative analysis of Peak Signal-to-Noise Ratio (PSNR) estimating FOD quality across different numbers of gradient directions. b, The quantitative analysis of angular correlation coefficient estimating FOD accuracy across different numbers of gradient directions. Mean values with a confidence interval of 95% (error bands) are reported.

protocols. Moreover, we significantly reduced training and testing time (2 days and a few minutes, respectively) when compared to FOD-Net (2 weeks and 2-3 hours, respectively). Interestingly, some metrics of HC were slightly better than that of patient data especially for



a. Mean Angular Error v.s. Number of Directions *b. AFD Error v.s. Number of Directions* *c. Peak Error v.s. Number of Directions*

FIGURE 2.14. Analysis on attributes of fibers. a, The quantitative analysis of mean angular error [104, 115] of fibers across different numbers of directions. b, The quantitative analysis of AFD error [104] of fibers across different numbers of gradient directions. c, The quantitative analysis of Peak Error [104] of fibers across different numbers of gradient directions. Mean values with a confidence interval of 95% (error bands) are reported.

SS3T methods, which suggests that the variability on patient brains is not properly captured by LARDI data and should be further studied.

Across FOD, fixel-based, and connectomic measures, the learning-based enhancement framework consistently improves single-shell LARDI FODs across different gradient directions, demonstrating its effectiveness and generalizability. Notably, the extent of improvement plateaus beyond approximately 24 gradient directions for FOD and its subsequent tractography and connectome reconstructions, with 6-24 directions making a notable improvement gap. This finding aligns with tensor modeling requisites (at least 28 for low b-values) for stablizing dMRI signal characterization [120], implying a potential shortage of dMRI information in single-shell acquisitions and a more substantial enhancement achievable through deep learning by incorporating diffusion information from multi-shell acquisitions before reaching 24 directions. Here, we demonstrate the capability of FOD enhancement frameworks regarding gradient directions, which has at least two clinical implications: (1) in the planning of future dMRI studies, and (2) to determine whether the application of enhancement on retrospectively collected datasets would be appropriate.

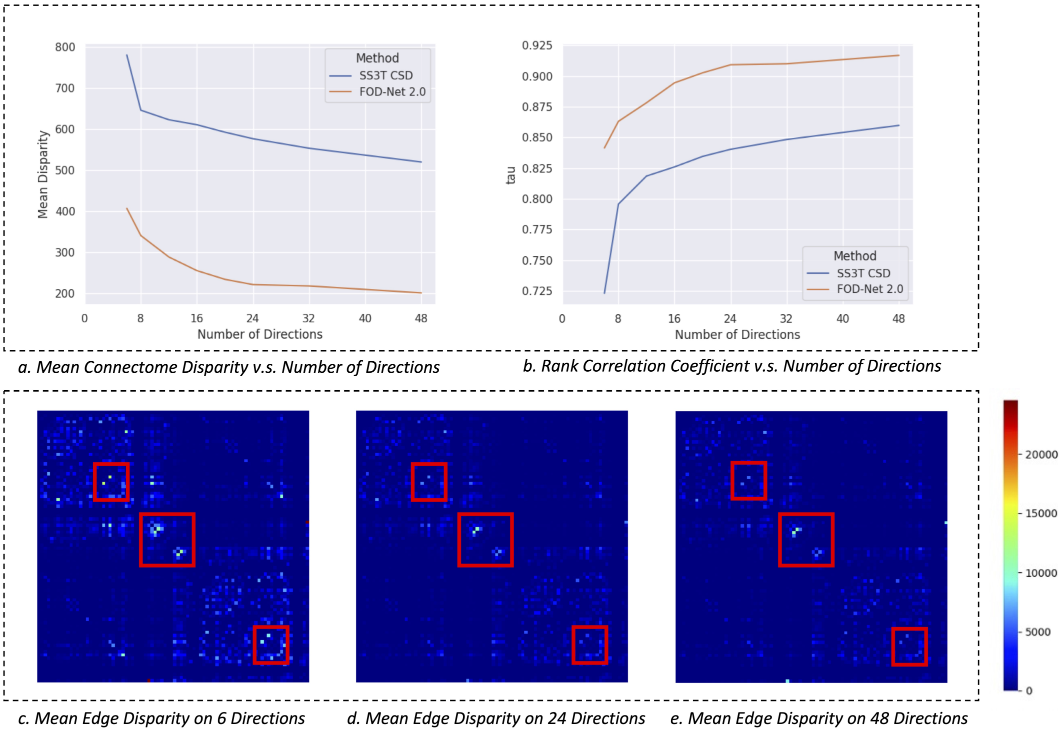


FIGURE 2.15. Analysis on brain connectivity. a, The quantitative analysis of mean connectome disparity (i.e., subject-wise average of all mean edge disparities) across different numbers of directions. b, The quantitative analysis of rank correlation coefficients of connectomes across different numbers of directions. c-e, The qualitative analysis of mean edge disparity matrices [104] across different numbers of directions. Red boxes denote edges with significant improvement plateauing after 24 directions.

2.5 Summary

We present an end-to-end deep learning framework to enhance FOD from LARDI acquisitions that accurately sifts spurious fibers and recovers missing fiber tracks. The method is robust in reconstructing high quality tractography and connectome and it generalizes to unseen protocols with data from different clinical groups.

Our result prove the effectiveness of learning-based FOD enhancement across a wide range of gradient directions with a low b-value of 1000 and acquiring more than 24 directions yields only marginal improvements for the enhanced FODs. This suggests that the optimal trade-off between the enhancement outcome and the number of gradient directions is achieved around 24 directions.

CHAPTER 3

FastFOD-Net: Advancing Diffusion MRI Analysis with Fast Deep Learning Enhancement

---

FOD is an advanced diffusion MRI modeling technique to represent complex white matter fiber configurations within a voxel, and a key step for subsequent brain tractography and connectome analysis. Its reliability and accuracy however heavily rely on the quality of the MRI acquisition and the subsequent estimation of the FODs at each voxel. Generating reliable FODs from widely available clinical protocols with single shell and low angular resolution acquisitions remains a challenging task, but potentially could be addressed with recent advances in deep learning-based enhancement techniques, such as the recently proposed FOD-Net algorithm. Despite the development of several related methods to enhance the angular resolution of the FODs in recent years, majority have been assessed on healthy subjects, which have proved to be a major hurdle for their clinical adoption.

In this chapter, we validate a newly optimized enhancement framework, FastFOD-Net, across healthy controls and six neurological disorders, which is a more practical solution to **Challenge 1: Clinical Rationale and Method Design** (Section 1.4.1). This accelerated end-to-end deep learning framework enhancing FODs with superior performance and delivering training/inference efficiency for clinical use (60$\times$ faster comparing to its predecessor). With the most comprehensive clinical evaluation to-date, our work demonstrates the potential of FastFOD-Net in accelerating clinical neuroscience research, empowering diffusion MRI analysis for disease differentiation, improved interpretability in connectome applications, reduced measurement errors to lower sample size requirements. By addressing **Challenge 2: Clinical Validation** (Section 1.4.2), this chapter provides the missing piece needed to increase the widespread adoption and build clinical trust in deep learning-based methods for diffusion MRI enhancement, enabling reliable analysis of real world clinical diffusion MRI data comparable to that achievable with high-quality research-quality protocols.

3.1 Introduction

Analysis of DWI can map structural brain connectivity non-invasively. Through CSD [27, 42, 54], high-quality FOD estimates for complex WM fiber bundle configurations can be reliably computed from multi-shell (i.e., DWI data acquired with multiple b-values) acquisitions of HARDI, yielding accurate WM tractography [121, 112]. In most clinical research settings, however, LARDI is more commonly acquired due to the constraints of non-research-focused clinical sites and their limited acquisition time for clinical patients [104]. This results in a large discrepancy between high-quality research-grade protocols and those used in clinical practice. Even historical retrospective datasets from research-quality protocols often lack multi-shell acquisitions necessary for advanced analysis [122, 123]. Consequently, clinical diffusion MRI data often failed to capture sufficient information to reliably reconstruct complex fiber structures, leading to misleading results with potentially dangerous consequences [113]. These challenges underscore the need for enhancement methods that can produce reliable results from data acquired with non-state-of-the-art research protocols.

Deep learning has been introduced recently to bridge the gap between the need for advanced research protocols and the constrained settings of clinical diffusion MRI protocols. One of these recent methods directly enhanced single FOD estimates obtained from patches of minimum pre-processed single-shell, clinical-quality data, achieving enhancements comparable to those from multi-shell HARDI acquisitions [104]. Despite promising results, few other related attempts [105, 106, 124] have only been validated on small scale datasets of healthy subjects and FOD-level metrics (i.e., an intermediate step in the analysis pipeline). This has contributed to their very limited practical adoption for clinical studies and is another instance where deep learning has not yet delivered the real-world practical impact and lived to its promise due to the lack of clinical-trust evaluation.

In this chapter, we establish an evaluation guideline for learning-based FOD enhancement (Figure 3.1), to compare and clinically validate a newly optimized variant FastFOD-Net. By introducing patch-wise training and inference instead of time-consuming voxel-wise predictions, the new FastFOD-Net algorithm obtains superior performance while greatly improving the computational efficiency by at least 20 times for training and 60 times for inference compared to its predecessor, making its clinical deployment practical. We conducted the most comprehensive clinical validation to date, comparing state-of-the-art diffusion analysis methods, including Fixel-Based Analysis (FBA) [125] and connectome analysis [126], to fully investigate the potential of the proposed enhancement framework. The clinical validation presented in this work encompasses applications not only in healthy controls but also in dataset from diverse neurological conditions, including Parkinson's Disease (PD), Essential Tremor (ET), Dystonic Tremor (DT), Amyotrophic Lateral Sclerosis (ALS), other motor neurone disease, and multiple sclerosis. Our results highlight the reliability of using FastFOD-Net to bridge the gap between advanced research protocols and constrained clinical settings, and transform its promise into real-world clinical impact. This will contribute to providing the missing piece for allowing the wider adoption of deep learning enhancement solutions to advance clinical neuroscience research using diffusion MRI data. The project page of FastFOD-Net with evaluation guideline is made available at: https://fastfodnet.github.io/.

3.2 Methods

3.2.1 Datasets

The comparison between datasets used in FOD-Net [104] and our work is summarized in Table 3.1.

We employed 80 training and 20 validation samples from the Human Connectome Project (HCP) [114], with an additional 15 validation samples drawn from HCP anomaly cases ($\text{HCP}_{Anomaly}$), which are datasets with unexpected brain abnormalities. Furthermore, we

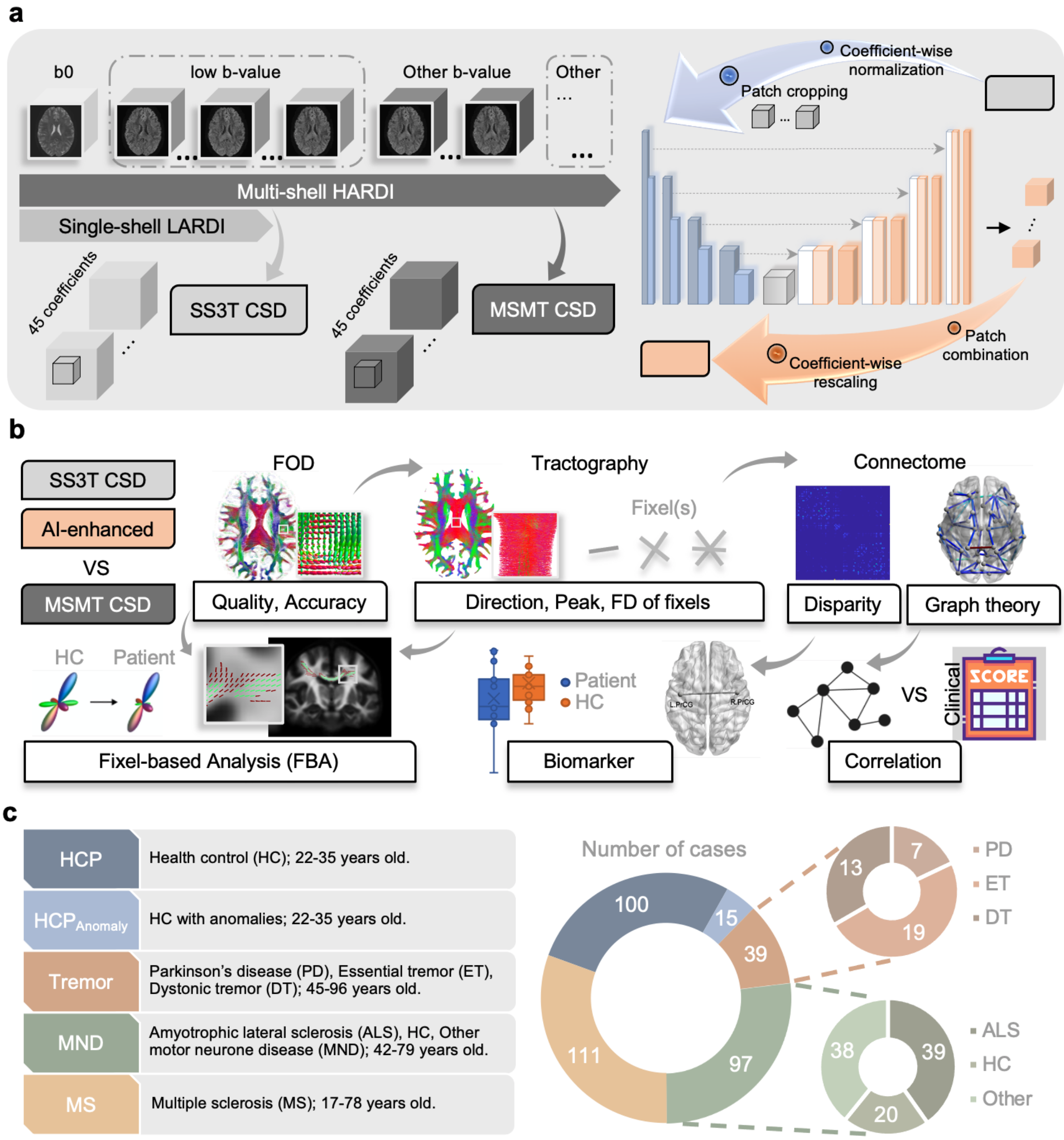


FIGURE 3.1. Overview of analysis workflow and data composition. a, Left: Illustration of input and ground truth (GT) data preparation. Multi-shell HARDI DWI data were acquired at multiple b-values with a large number of diffusion gradient directions (typically over 60 directions) [104], while LARDI data typically combine a single b0 image with a smaller set of directions, often from a single low b-value acquisition. FOD coefficients were generated using SS3T CSD [117] for LARDI (input in light gray) and MSMT CSD [27] for HARDI (GT in drak gray). Right: We developed a fast deep learning method termed FastFOD-Net to estimate high-quality FOD coefficients (in orange) from SS3T CSD data. b, The evaluation guideline contains diverse end-point diffusion analysis metrics to fully investigate the potential of deep learning solutions in clinical neuroscience research. c, Composition of the data used in this study: the Human Connectome Project (HCP), HCP subjects with MRI anomalies ($\text{HCP}_{Anomaly}$), Tremor, MND, and MS dataset.

conducted 5-fold cross-validation using in-house datasets (denoted as Tremor, MND, MS) from patients with tremors [127], motor neuron diseases, and multiple sclerosis. All these datasets were acquired with different HARDI protocols that we selected as the ground truth (GT) data - see details of acquisition parameters and data processing steps below.

3.2.1.1 The Human Connectome Project

The HCP [114] launched in 2010 stands out as one of the largest and most commonly used datasets within the neuroimaging community. For each subject, DWI sequences were acquired using advanced multi-shell HARDI protocols [128]. The HCP study aimed to study human brain connectivity from an initial cohort of 1200 young adults aged between 22 and 35 years.

100 unrelated samples were selected, and this dataset comprises minimally processed DWI data captured at an isotropic resolution of 1.25 mm with a 3T scanner. The repetition time (TR), echo time (TE) and field of view (FOV) were 5520 ms, 89.5 ms and 256 mm, respectively. Three gradient shells with b-values of 1000, 2000, and 3000 s/mm^2 were utilized, encompassing 90 diffusion-encoding directions per shell. The acquisition matrix was set to 145×145 to acquire 174 coronal slices for each subject. Furthermore, the protocol included 18 b0 volumes. Each subject also included a high-resolution T1-weighted dataset, which was acquired with an isotropic voxel size of 0.7 mm, TR/TE = 2400/2.14 ms, and flip angle = $8°$.

Additionally, 15 HCP subjects with focal structural anomalies (denoted as HCP$_{Anomaly}$), but considered benign during structure quality control (QC) [114, 129], were assessed separately to evaluate the enhancement performance. The regions of interest (ROIs) were delineated by trained neuroimaging analyst and guided by the associated report[1].

[1] https://www.humanconnectome.org/study/hcp-young-adult/document/1200-subjects-data-release

3.2.1.2 Clinical Datasets

We also included three neurological study cohorts: (i) a Tremor dataset encompassing patients who exhibited various tremor disorders, (ii) and a MND dataset including healthy controls and patients, and (iii) a MS dataset with clinical outcomes.

The Tremor dataset described previously [127], contains patients with Parkinson's disease (PD), Essential Tremor (ET) and Dystonic Tremor (DT). Imaging data were acquired from patients diagnosed with one of the neurological movement disorders at a single centre using a 3T MRI scanner (SIGNA Architect, General Electric, Milwaukee). A total of 39 subjects with consistent diffusion protocol were included, which consisted of 8 b0 volumes and 3 gradient shells of b-values of 700, 1,000, 2,800 s/mm^2 with 25, 40, 75 directions respectively, and the following acquisition parameters: TR = 6250 ms, TE = 106 ms, flip angle = $90°$, FOV = 230 mm, acquisition matrix = 128×128, slice thickness = 2 mm, and a total of 72 axial slices. The structural imaging protocol consisted of a sagittal 3D IR-FSPGR T1-weighted image with the following parameters: TR: 8 ms, TE: 3.24 ms, flip angle: $10°$, FOV: 256 mm, acquisition matrix: 256×256, slice thickness: 1.2 mm, and a total of 146 sagittal slices. The age range of this patient cohort spanned from 45 to 96 years, including a higher proportion of older individuals.

The second clinical dataset denoted as MND dataset, was collected from 97 samples of 3 different clinical groups using a 3T MRI scanner (GE MR750, DV29, 32-channel Nova head coil): 39 ALS subjects, 20 healthy controls, and 38 subjects with other motor neuron disorders. The age of this cohort is from 42 to 79 years old. The diffusion acquisition protocol contains 4 b-values of 0, 700, 1000, 2800 s/mm^2 with 8, 25, 40 and 75 directions, respectively. The parameters were set as follows: TE/TR = 100/3245 ms, flip angle = $90°$, FOV = 240 mm, acquisition matrix = 128×128, an isotropic voxel size of 2 mm, and a total of 70 axial slices. The T1-weighted images were acquired with the matrix size of 256×256 and an isotropic voxel size of 1 mm along the coronal orientation. Other parameters were set to an isotropic voxel size of 1 mm, TE/TR = 2.3/6.2 ms, Flip Angle = $12°$, FOV = 240 mm, and the number of slices = 204.

The last dataset consists of 111 MS subjects with well characterized clinical outcomes such as lesion volume, brain atrophy, disease duration and EDSS within 12 months. This population presented a large age range from 17 to 78 years old. The dataset was acquired using a 3T SIEMENS Skyra scanner. The acquisition parameters of DWI: TE/TR = 111/3300 ms, flip angle = 90, FOV = 256 mm, acquisition matrix = 128 $\times$ 128, slice thickness = 2 mm, number of slices = 76. Due to various in the different diffusion acquisition protocol used, we focused on cases that had 30 directions for the 1000 and 2500 b-values. The T1-weighted images were acquired with the matrix size of 256 $\times$ 256 and an isotropic voxel size of 1 mm along the sagittal orientation. Other parameters were set to an isotropic voxel size of 1 mm, TE/TR = 2.98/2300 ms, Flip Angle = $9°$, FOV = 256 mm, and the number of slices = 176.

TABLE 3.1. Details of the datasets used in FOD-Net [104] and our work, including the acquisition parameters all pertaining to the diffusion imaging protocols, are summarized below. HC denotes healthy control while $\text{HC}_{Anomaly}$ represents healthy cases with abnormalities on focal brain. For the Tremor datast, DT, ET, and PD denote patient cases with Parkinson's Disease (PD), Essential Tremor (ET) and Dystonic Tremor (DT). For the MND dataset, ALS denotes amyotrophic lateral sclerosis cases while other denotes other motor neuron disorder patients. The MS dataset consists of patients diagnosed with multiple sclerosis.

	FOD-Net				**FastFOD-Net**				
Name	HCP	$\text{HCP}_{Anomaly}$	Local	Multicenter	HCP	$\text{HCP}_{Anomaly}$	Tremor	MND	MS
Type	HC	$\text{HC}_{Anomaly}$	HC	HC	HC	$\text{HC}_{Anomaly}$	DT, ET, PD	ALS, other, HC	MS
#Data[1]	110	3	4	3	100	15	39	97	111
Age	22−35		25−35	23, 26, 23	22−35		45−96	42−79	17−78
Spacing[2]	1.25 mm		2 mm	-	1.25 mm		1.8 mm	2 mm	2 mm
Matrix[3]	145×145		128×128	-	145×145		128×128	128×128	128×128
#Slices[4]	174		70	-	174		72	70	76
Orientation	Coronal		Axial	Axial	Coronal		Axial	Axial	Axial
Shell bvalues[5]	0, 1000, 2000, 3000			-	0, 1000, 2000, 3000		0, 700, 1000, 2800		0, 200, 500, 1000, 2500
Shell sizes[6]	18, 90, 90, 90		8, 32, 32, 72	-	18, 90, 90, 90		8, 25, 40, 75		5/11, 3, 6 30, 30
Clinical outcome	-	-	-	-	-	-	-	-	✓

[1] #Data represents the number of samples in this dataset. [2] Spacing is the voxel size.
[3] Size and [4] #Slices denote the size of acquisition matrix and the number of slices acquired along the orientation.
[5] Shell b-values list the average b-value for each shell, and [6] shell sizes list the number of volumes in each shell (from lowest to highest).

3.2.2 Data Processing

Upon acquiring and pre-processing single-shell clinical quality DWI data and its corresponding multi-shell HARDI counterparts, the subsequent step focused on the generation of FOD maps with MRtrix [116, 63] by adopting SS3T CSD [117] on single-shell sampled data and MSMT [27] CSD on the multi-shell HARDI data, respectively. The data processing workflow is outlined in Figure 3.2.

3.2.2.1 DWI Pre-processing

Since the DWI data are subject to various imaging artifacts, particularly the inherent subject motion and physiological noise, a range of pre-processing steps are recommended to reduce the confounding effects inherent in DWI acquisitions [29]. Except minimally processed data from HCP and HCP$_{Anomaly}$, other clinical datasets were pre-processed following recommendations by MRtrix [116, 63] and FSL [32]. Denoising was the first step to be applied since noise amplifications lower the sensitivity of DWI data. Then, unringing was performed to remove Gibbs ringing artifacts using the local subvoxel-shifts method [130]. Motion and distortion correction were conducted to address the misalignment of slices and microstructural parameters [29, 30]. Finally, bias field correction was performed as a pre-processing step to eliminate low frequency intensity inhomogeneities across the image [37].

3.2.2.2 Pre-processing of T1-weighted Image

Additionally, the T1-weighted images were segmented using Freesurfer [131] according to the DK atlas [70], to generate a parcellation image consisting of 84 discrete nodes. The pure and partial WM tissues were defined through the segmentation of WM and GM using hybrid surface and volume segmentation (HSVS) [132]. This segmentation can be used to construct a 5TT mask for each subject, which encompasses cortical GM, subcortical GM, WM, CSF, and pathological tissue (which remains empty on healthy subjects). Subsequently, the 5TT and the parcellation image were registered to the processed DWI for structural connectome reconstruction.

3.2.2.3 FOD Generation

To emulate single-shell LARDI data, we used the Kennard-Stone subsampling strategy [133] on the original multi-shell HARDI diffusion data. This strategy enabled the selection of 32 uniformly distributed directions from the lowest b-value shell (b = 1000 or 700 s/mm^2; 30 directions for MS data). Subsequently, these subsets were combined with the b0 data. The subsequent step involves the generation of FOD maps with MRtrix [116, 63]. First, the 'dhollander' algorithm [134] was used to estimate WM, GM and CSF response functions for each subject and compute the group average across the entire subjects of each dataset [125]. Note that the response functions for the subsampled DWI data were recomputed since the downsampling process can introduce slight changes in voxel values on the DWI data and ensuring accurate response functions was crucial for subsequent modeling process. Then, two distinct operations were performed to generate the corresponding FOD maps: SS3T CSD [117] on the single-shell sampled data, and MSMT [27] CSD on the multi-shell HARDI data. A maximum harmonics order (l_{max}) equal to 8 [54, 104] was used, yielding a total of 45 spherical harmonic coefficients regardless of number of gradient directions used during acquisition.

FODs are continuous functions that can be discretized into a finite number of fiber directions, referred to as 'fixels', within each voxel [58]. White matter bundles were segmented using fixel data derived from MSMT FODs through TractSeg [135].

3.2.2.4 Connectome Generation

Connectome [20] is a complete map of brain connections, which can be used to investigate global changes in brain connectivity and probe the connectivity of specific ROIs with the rest of the brain. The procedure for constructing the human connectome [20] of each subject consists of four primary stages: (1) whole-brain tractography; (2) quantifying the weight of each streamline to ensure quantitative connectivity results [69]; (3) selecting a GM atlas; (4) constructing the connectome by assessing the degree of structural connectivity between pairs of nodes.

For the first step, we employed anatomically-constrained whole-brain tractography on all subjects using the HCP tractography pipeline in MRtrix [116, 63]. Driven by second-order integration over fiber orientation distributions (iFOD2), the tractogram under the framework of ACT [64] was computed by taking a given FOD image as the input and providing the corresponding 5TT mask. Notable parameters included a step size of $0.5 \times voxel\ size$, a maximum curvature of 45 degrees per step, a length range from 5 to 250 mm, and a FOD amplitude threshold of 0.06. Seeding was randomized from the interface between GM and WM, generating 10 million streamlines. Following tractography, the SIFT2 algorithm [69] was applied to assign weights to each streamline, enhancing the quantitative and biologically relevant properties of the tractogram [71]. Subsequently, in the third step, the T1-weighted image of each subject was segmented into meaningful brain nodes using the DK atlas [70]. In the fourth phase, we constructed the connectome matrix for each FOD-based tractogram by summing the SIFT2 weighted of all streamlines connecting each pair of nodes.

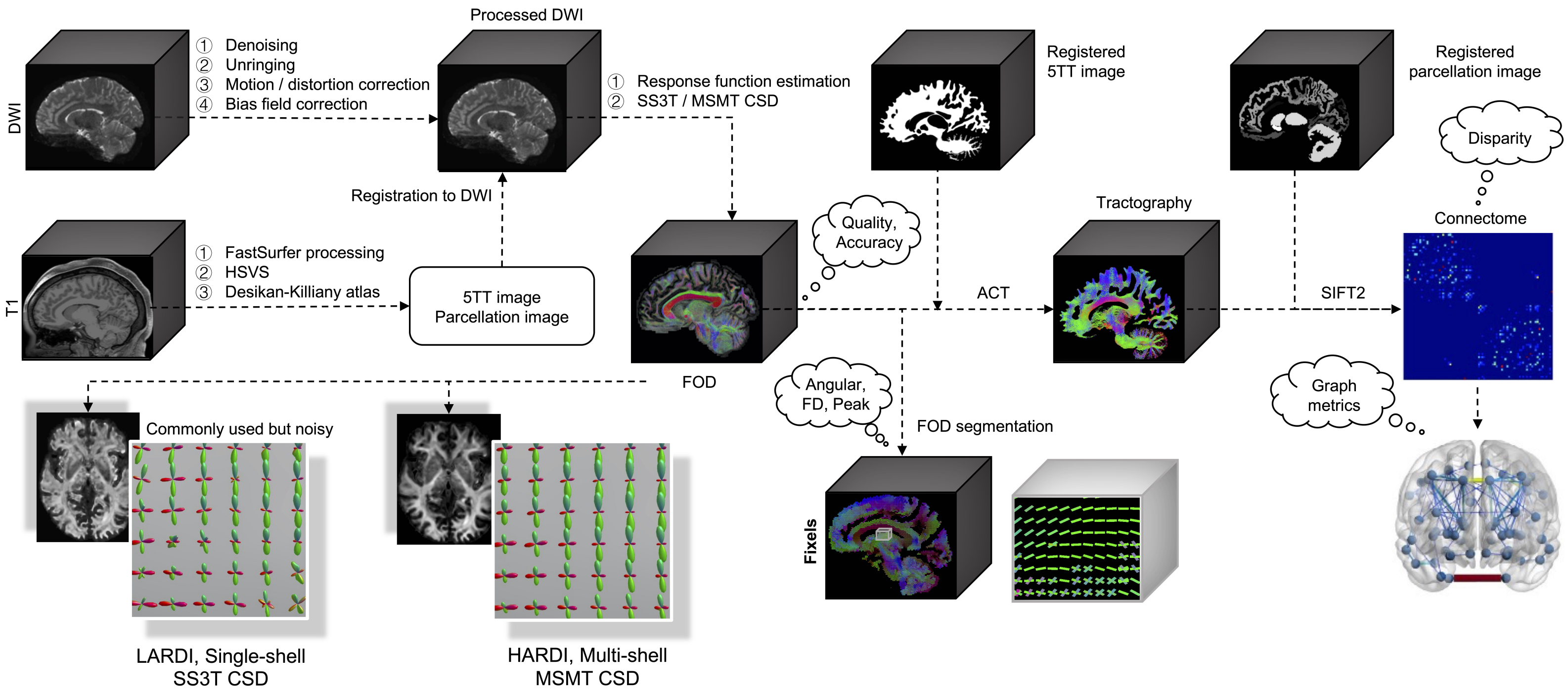


FIGURE 3.2. Workflow of data processing. The pipeline includes DWI pre-processing, T1-weighted image pre-processing. FODs are estimated using SS3T or MSMT CSD. FOD segmentation enables fiber bundle element ‘fixel’ analysis from direction, fiber density (FD), and peak properties. Anatomically-constrained tractography (ACT) is performed utilizing registered five-tissue-type (5TT) images to generate tractograms. Fixel analysis can be served as a proxy evaluation of tractography. Finally, structural connectomes are constructed using registered parcellations and Spherical deconvolution Informed Filtering of Tractograms 2 (SIFT2) re-weighted streamlines, followed by analyses focusing on connectivity disparities and graph metrics.

3.2.3 Architecture

We extended the FOD-Net design [104] to directly predict entire patches by adopting a fully convolutional backbone [136]. This not only learns better features of FODs by overcoming the limitations of FOD-Net but also conserves computational resources allowing for greater practical applicability.

FOD-Net [104] applies a combined architecture of CNNs and MLPs to enhance coefficients of a single FOD estimate voxel-by-voxel and for each spherical harmonics order separately based on single-shell LARDI data; this approach is time consuming due to the high complexity of MLPs, and overlooks the inherent 3D attributes of imaging data and inter-order coefficient relationships. Here, we extend the FOD-Net design to enable the direct prediction of entire patches by adopting a fully convolutional backbone [136]. This not only learns better features of FODs by overcoming the above limitations but also conserves computational resources allowing for greater practical applicability. Due to the capacity of U-Net to effectively handle diverse yet small datasets and yield acceptable predictions even with limited annotated data, researchers have proposed an appropriate training framework without architectural changes that automatically adapts to a given dataset for semantic segmentation [137].

Here, we optimized the training paradigm for learning-based FOD enhancement. Firstly, to accelerate the convergence of our model and reduce its sensitivity to outliers, coefficient-wise standardization following multi-tissue informed intensity normalization [138] was undertaken on FODs and all coefficients were brought to a consistent scale. Moreover, to allow for the consideration of information from neighboring FODs within the same patch and interactive learning between different orders, our model abandoned separate order learning and FOD estimates were trained on a patch-wise basis regardless of the original data size; this also makes the solution more practically applicable. Coefficient-wise rescaling was eventually applied to the combined prediction to adjust each coefficient back to its original scale.

3.2.4 Implementation Details

Our model was trained with the mean squared error (MSE) loss and Adam optimizer with the number of epochs set to 50. The learning rate was set to 0.01 with a batch size of 4. U-Net [136] is a state-of-the-art architecture that can be generalized to any end-to-end medical imaging application. The 'U' shape architecture includes a contracting path that captures features of FODs through 3D convolutional and downsampling layers, as well as an expansive path that upsamples feature maps to generate high-resolution FOD patches. The inclusion of skip connections facilitates the transfer of fine-grained information between the encoder and decoder. It consists of 4 encoding blocks, each corresponding to a decoding block at the same scale. The number of channels doubles with each increasing level during encoding, and it decreases by half with each subsequent level during decoding. Convolutional blocks with stride 2 were applied for downsampling and interpolation were used for upsampling. We cropped input patches of $64 \times 64 \times 64 \times 45$ with a sliding window of stride 32 and then fed them into the model for training.

3.2.5 Quantitative Diffusion MRI Assessment

In the last chapter, we developed an initial approach to evaluating the generated data from FOD, tractography, and connectome perspective. In this chapter, we refined the evaluation metrics and addressed aspects that were previously overlooked. We presented a clinical-trust evaluation pipeline following the analysis workflows of current state-of-the-art diffusion analysis techniques.

3.2.5.1 Overall FOD Assessment

The quantitative FOD assessment was designed to determine whether the GM partial volume influenced the performance of FOD enhancement methods. Peak Signal-to-Noise Ratio ($\mathbf{PSNR}$) quantified the impact of enhancement on image quality and evaluating the similarity of FODs between GT (MSMT CSD) and single-shell based methods (SS3T CSD, FOD-Net and FastFOD-Net). Additionally, the angular correlation coefficient ($\mathbf{r_{Angular}}$) [139] provides a measurement of how correlated the spherical harmonic coefficients are between GT data and other estimates as formulated in Equation 3.1.

Angular correlation coefficient $\mathbf{r_{Angular}}$ [139] can be calculated as follows:

$$\mathbf{r_{Angular}}(u, v) = \frac{\sum_{c=1}^{C} u_c {v_c}^{\mathbf{T}}}{\sqrt{\sum_{c=1}^{C} {u_c}^2}\sqrt{\sum_{c=1}^{C} {v_c}^2}} \tag{3.1}$$

where u and v refer to the GT FODs and the FODs estimated from LARDI (with or without enhancement), respectively; C denotes the total number of spherical harmonic coefficients and c refers to the c^{th} spherical harmonic coefficient. The coefficient provides a quantitative measure of the agreement between two FOD estimates.

3.2.5.2 Fiber Bundle Element Assessment

Three key metrics were utilized to assess specific fiber bundle elements within a voxel (referred to as 'fixels') between GT and LARDI methods: mean angular error ($\mu\mathbf{E_{Angular}}$), peak error ($\mathbf{E_{Peak}}$), and fiber density (FD) error ($\mathbf{E_{FD}}$) [104]. Although the original FOD-Net paper quantified the accuracy of fiber bundle element reconstruction, the effect of unmatched fiber elements between GT fixels and estimated fixels was not considered when computing these measures. Specifically, each fiber bundle element of a LARDI estimations was always matched to a GT fixel regardless of their angular error. It implies that matched fiber bundle elements could have perpendicular orientations in the worst-case scenario and lead to misleading results on misaligned pairs of fixels. False positive fixels would be ignored even if they were erroneous. In that sense, $\mu\mathbf{E_{Angular}}$, stands as one of the most valuable metrics because the accurate reconstruction of fiber orientations is key for understanding brain connectivity. To fairly compare fiber bundle elements between GT and other methods, estimated fixels with angular errors higher than a selected threshold are considered misaligned for the rest of the paper. Following our definitions of matched and misaligned fixels, we optimized the $\mathbf{E_{Peak}}$ and $\mathbf{E_{FD}}$ measures by also considering misaligned fixels.

To further assess the performance of enhancement frameworks on regions with varying fiber complexity, we generated masks based on TractSeg masks. Particularly, three main anatomical ROIs were defined [104]: a single fiber region (ROI 1) within a selection of the corpus callosum (CC) primarily composed of large and coherent fiber tracts; the region where the middle cerebellar peduncle (MCP) and corticospinal tract (CST) bundles intersect, both of which exhibit coherent orientations, which results in the voxels containing only two fixels in this intersection (ROI 2); the intersected region of superior longitudinal fascicle (SLF), CST, and CC with more complex fiber configurations (i.e., three distinct fixels, ROI 3). Since the regions were defined based on the number of crossing fiber bundles, we removed voxels that contained more fixels than expected for a fairer analysis regarding fixel complexity. For example, ROI 3 should only have 3-fixel voxels.

Mean angular error, denoted as $\mu\mathbf{E_{Angular}}$ [104], is defined as follows.

$$\mu\mathbf{E_{Angular}} = \frac{1}{n}\sum_{i=0}^{n} \frac{\mathbf{f^{\tilde{d}}} \cdot \mathbf{\hat{f}^{\tilde{d}}}}{\|\mathbf{f^{\tilde{d}}}\|_2 \, \|\mathbf{\hat{f}^{\tilde{d}}}\|_2} \tag{3.2}$$

Here, $\mathbf{f^{\tilde{d}}}$ is a unit vector that represents the direction of an estimated fixel in three-dimensional space (i.e., x, y, z axes), $\mathbf{\hat{f}^{\tilde{d}}}$ is the unit vector of its GT, $\cdot$ is the dot product, and $\| \; \|_2$ is the L2 norm. To compare fixels fairly, we selected a maximum threshold of 45 degrees when matching fixels. Fixels with angular errors higher than that are considered misaligned for the rest of the paper. For matched fixels, peak error is computed as the absolute difference between amplitude at the maximum peak of the estimated fixels $\mathbf{f^{Peak}}$ and their corresponding GT $\mathbf{\hat{f}^{Peak}}$. Moreover, FD error for matched fixels is computed in a similar manner as the absolute difference between the integral of a given fixel estimate $\mathbf{f^{FD}}$ and its corresponding GT $\mathbf{\hat{f}^{FD}}$ to estimate the density of fibers in a specific region. Misaligned fixels are considered as either extra $\mathbf{f_+}$ or missing $\mathbf{f_-}$ fixels on estimations. The whole value is then summed to compute $\mathbf{E_{Peak}}$ and $\mathbf{E_{FD}}$ as formulated in Equation 3.3.

$$\mathbf{E} = \sum_{i=0}^{n} \left|\mathbf{f} - \mathbf{\hat{f}}\right| + \sum_{i=0}^{n} \mathbf{f_+} + \sum_{i=0}^{n} \mathbf{f_-} \tag{3.3}$$

Then we computed these improved metrics in three defined ROIs (single-fiber population, 2-way crossing, and 3-way crossing) to analyze the performance under distinct fiber configurations. Anomaly regions denoting ROIs surrounding the anomalous regions on HCP$_{Anomaly}$

cases, and lesions on MS cases, were also analyzed. Moreover, normal appearing WM (NAWM) outside lesions was evaluated on MS cases to determine whether deep learning methods performed consistently regardless of the presence of MS lesions.

3.2.5.3 Connectome Assessment

The structural connectome (**M**) [20], a comprehensive representation of brain connections, was generated using anatomically-constrained tractography [64] with the input FOD image, as detailed in Section 3.2.2. It serves as a graph matrix with nodes denoting gray matter regions and edges denoting brain connections between them. Full name of nodes and what lobe they belong to for DK atlas is presented in Table 3.2.

In this work, we first defined the disparity matrix as an edge-wise disparity measurement and employed Kendall ranking coefficient (τ) [140] to asses the overall difference between a LARDI connectome and its corresponding GT matrix. We defined the disparity matrix as an edge-wise disparity measurement to asses the overall difference between a LARDI estimate ($\mathrm{M}_{\mathbf{Est}}$) and its corresponding GT matrix ($\mathrm{M}_{\mathbf{GT}}$) measured as the mean absolute error. Subsequently, we computed the subject-wise average of mean edge disparities to summarize the overall difference, denoted as the mean connectome disparity ($\mu\mathbf{Disparity}$). We also counted the percentage of significantly different edges by extending the disparity analysis with edge-wise paried t-tests ($p < 0.05$) and false discovery rate (FDR) correction. While edge weights indicate connectivity magnitude, the absolute values may be less informative than the relative rankings, making rank-order correlation a more meaningful metric than direct numerical comparison.

In addition, to analyze the connectome enhancement quality from a network perspective, several graph theory metrics [141, 142] were computed via Brain Connectivity Toolbox (BCT) [143] and compared between methods. The difference ratio (DR) of the graph metrics of the reference multi-shell connectome and other estimates are expressed in percentages, was calculated as:

$$\mathbf{DR} = \frac{\mathbf{M_{Est}} - \mathbf{M_{GT}}}{\mathbf{M_{GT}}} \times 100\%. \tag{3.4}$$

TABLE 3.2. Full name of brain nodes (regions) and what lobe they belong to for Desikan-Killiany (DK) atlas [70].

Key	Name	Short Name	Key	Name	Short Name
0	Left-BanksSTS-Cortex	L.BSTS	42	Right-Thalamus	R.TH
1	Left-CaudalAnteriorCingulate-Cortex	L.CACG	43	Right-Caudate	R.CA
2	Left-CaudalMiddleFrontal-Cortex	L.CMFG	44	Right-Putamen	R.PU
3	Left-Cuneus-Cortex	L.CU	45	Right-Pallidum	R.PA
4	Left-Entorhinal-Cortex	L.EC	46	Right-Hippocampus	R.HI
5	Left-Fusiform-Cortex	L.FG	47	Right-Amygdala	R.AM
6	Left-InferiorParietal-Cortex	L.IPG	48	Right-Accumbens	R.AC
7	Left-InferiorTemporal-Cortex	L.ITG	49	Right-BanksSTS-Cortex	R.BSTS
8	Left-IsthmusCingulate-Cortex	L.ICG	50	Right-CaudalAnteriorCingulate-Cortex	R.CACG
9	Left-LateralOccipital-Cortex	L.LOG	51	Right-CaudalMiddleFrontal-Cortex	R.CMFG
10	Left-LateralOrbitoFrontal-Cortex	L.LOFG	52	Right-Cuneus-Cortex	R.CU
11	Left-Lingual-Cortex	L.LG	53	Right-Entorhinal-Cortex	R.EC
12	Left-MedialOrbitoFrontal-Cortex	L.MOFG	54	Right-Fusiform-Cortex	R.FG
13	Left-MiddleTemporal-Cortex	L.MTG	55	Right-InferiorParietal-Cortex	R.IPG
14	Left-ParaHippocampal-Cortex	L.PHIG	56	Right-InferiorTemporal-Cortex	R.ITG
15	Left-ParaCentral-Cortex	L.PaCG	57	Right-IsthmusCingulate-Cortex	R.ICG
16	Left-ParsOpercularis-Cortex	L.POP	58	Right-LateralOccipital-Cortex	R.LOG
17	Left-ParsOrbitalis-Cortex	L.POR	59	Right-LateralOrbitoFrontal-Cortex	R.LOFG
18	Left-ParsTriangularis-Cortex	L.PTR	60	Right-Lingual-Cortex	R.LG
19	Left-Pericalcarine-Cortex	L.PCAL	61	Right-MedialOrbitoFrontal-Cortex	R.MOFG
20	Left-PostCentral-Cortex	L.PoCG	62	Right-MiddleTemporal-Cortex	R.MTG
21	Left-PosteriorCingulate-Cortex	L.PCG	63	Right-ParaHippocampal-Cortex	R.PHIG
22	Left-PreCentral-Cortex	L.PrCG	64	Right-ParaCentral-Cortex	R.PaCG
23	Left-PreCuneus-Cortex	L.PCU	65	Right-ParsOpercularis-Cortex	R.POP
24	Left-RostralAnteriorCingulate-Cortex	L.RACG	66	Right-ParsOrbitalis-Cortex	R.POR
25	Left-RostralMiddleFrontal-Cortex	L.RMFG	67	Right-ParsTriangularis-Cortex	R.PTR
26	Left-SuperiorFrontal-Cortex	L.SFG	68	Right-Pericalcarine-Cortex	R.PCAL
27	Left-SuperiorParietal-Cortex	L.SPG	69	Right-PostCentral-Cortex	R.PoCG
28	Left-SuperiorTemporal-Cortex	L.STG	70	Right-PosteriorCingulate-Cortex	R.PCG
29	Left-SupraMarginal-Cortex	L.SMG	71	Right-PreCentral-Cortex	R.PrCG
30	Left-FrontalPole-Cortex	L.FP	72	Right-PreCuneus-Cortex	R.PCU
31	Left-TemporalPole-Cortex	L.TP	73	Right-RostralAnteriorCingulate-Cortex	R.RACG
32	Left-TransverseTemporal-Cortex	L.TTG	74	Right-RostralMiddleFrontal-Cortex	R.RMFG
33	Left-Insula-Cortex	L.IN	75	Right-SuperiorFrontal-Cortex	R.SFG
34	Left-Cerebellum-Cortex	L.CER	76	Right-SuperiorParietal-Cortex	R.SPG
35	Left-Thalamus	L.TH	77	Right-SuperiorTemporal-Cortex	R.STG
36	Left-Caudate	L.CA	78	Right-SupraMarginal-Cortex	R.SMG
37	Left-Putamen	L.PU	79	Right-FrontalPole-Cortex	R.FP
38	Left-Pallidum	L.PA	80	Right-TemporalPole-Cortex	R.TP
39	Left-Hippocampus	L.HI	81	Right-TransverseTemporal-Cortex	R.TTG
40	Left-Amygdala	L.AM	82	Right-Insula-Cortex	R.IN
41	Left-Accumbens	L.AC	83	Right-Cerebellum-Cortex	R.CER

The graph metrics analyzed include [143, 141, 142]:

- Global Efficiency: The global efficiency quantifies how efficiently information or signals can be exchanged between all pairs of nodes in a brain network, providing insights into the overall interconnectedness and communication effectiveness of a brain connectome.

- Transitivity: The transitivity is the ratio of triangles to triplets in the network and is an alternative to the clustering coefficient.
- Density: Density is the fraction of present connections to possible connections. Connection weights are ignored in calculations.
- Assortativity: The assortativity coefficient is a correlation coefficient between the degrees of all nodes on two opposite ends of a link. A positive assortativity coefficient indicates that nodes tend to link to other nodes with the same or similar degree.
- Mean Betweenness Centrality: Node betweenness centrality is the fraction of all shortest paths in the network that contain a given node. Nodes with high values of betweenness centrality participate in a large number of shortest paths.
- Mean Modularity: The optimal community structure is a subdivision of the network into nonoverlapping groups of nodes in a way that maximizes the number of within-group edges, and minimizes the number of between-group edges. The modularity is a statistic that quantifies the degree to which the network may be subdivided into such clearly delineated groups.
- Mean Strength: Node strength is a measure of the importance or influence of a node within a network and is computed as the sum of the weights of all the edges connected to a brain node. The strength of the entire brain connectome was then determined by averaging the strengths of all brain nodes.
- Mean Network Degree: Node degree is the number of links connected to the node. In directed networks, the in-degree is the number of inward links and the out-degree is the number of outward links. Connection weights are ignored in calculations.
- Mean Clustering Coefficient: The clustering coefficient is the fraction of triangles around a node and is equivalent to the fraction of node's neighbors that are neighbors of each other.
- Mean Local Efficiency: The local efficiency is the global efficiency (see above) computed on node neighborhoods, and is related to the clustering coefficient.

3.2.5.4 Clinical Research Application 1: Fixel-Based Analysis

Fixel-Based Analysis (FBA) [144, 59] is a technique to examine alterations in microstructure and macrostructure properties of WM tracts at the fixel level, most commonly used when comparing patients to control groups. Macroscopic differences of FD [59] can be assessed on a fixel-wise level by leveraging the registration of individual subject data to a common template space. Ultimately, the FBA pipeline produces fixel-wise statistical outcomes between clinical groups, assigning a specific p-value to each individual fixel, even when multiple distinct fixels coexist within the same voxel. In this study, our hypothesis is that more significant differences can be observed based on enhanced FODs when compared with no enhancement.

Several studies have conducted direct investigations into the association between structural brain abnormalities and the cognitive and behavioral profiles of relatively small samples of ALS patients [145, 146, 147, 148]. Based on findings from these studies, our study performs FBA analysis on FOD estimates in the CC and CST which are sensitive to impairments in ALS populations when compared to healthy controls (Figure 3.3). We evaluate whether the improvements brought by learning-based frameworks can yield similar results to those obtained from the analysis of multi-shell HARDI FODs (GT), and obtained $\mathbf{TP}$, $\mathbf{FN}$, $\mathbf{FP}$ and $\mathbf{TN}$. To summarize these measures, we computed the $\mathbf{Sensitivity}$, $\mathbf{Specificity}$ and $\mathbf{Precision}$ for each enhancement method. $\mathbf{F1}$ is also included and particularly valuable in scenarios where there is an imbalance between the two classes being classified, such as significantly different fixels and non-significantly different fixels, as it combines both $\mathbf{Sensitivity}$ and $\mathbf{Specificity}$.

To perform FBA, we take advantage of the publicly available MRtrix software package [63]. To compare the FBA results obtained from two population templates, we initially registered the population templates of each competing FOD estimation method into the multi-shell space. Subsequently, we warped the subject FOD images into this new template space. This warping allowed us to perform the statistical analysis in a common space for the reoriented fixels, assessing statistical differences between the patient group and the control group in terms of FD. However, an additional fixel matching step is necessary before making comparisons.

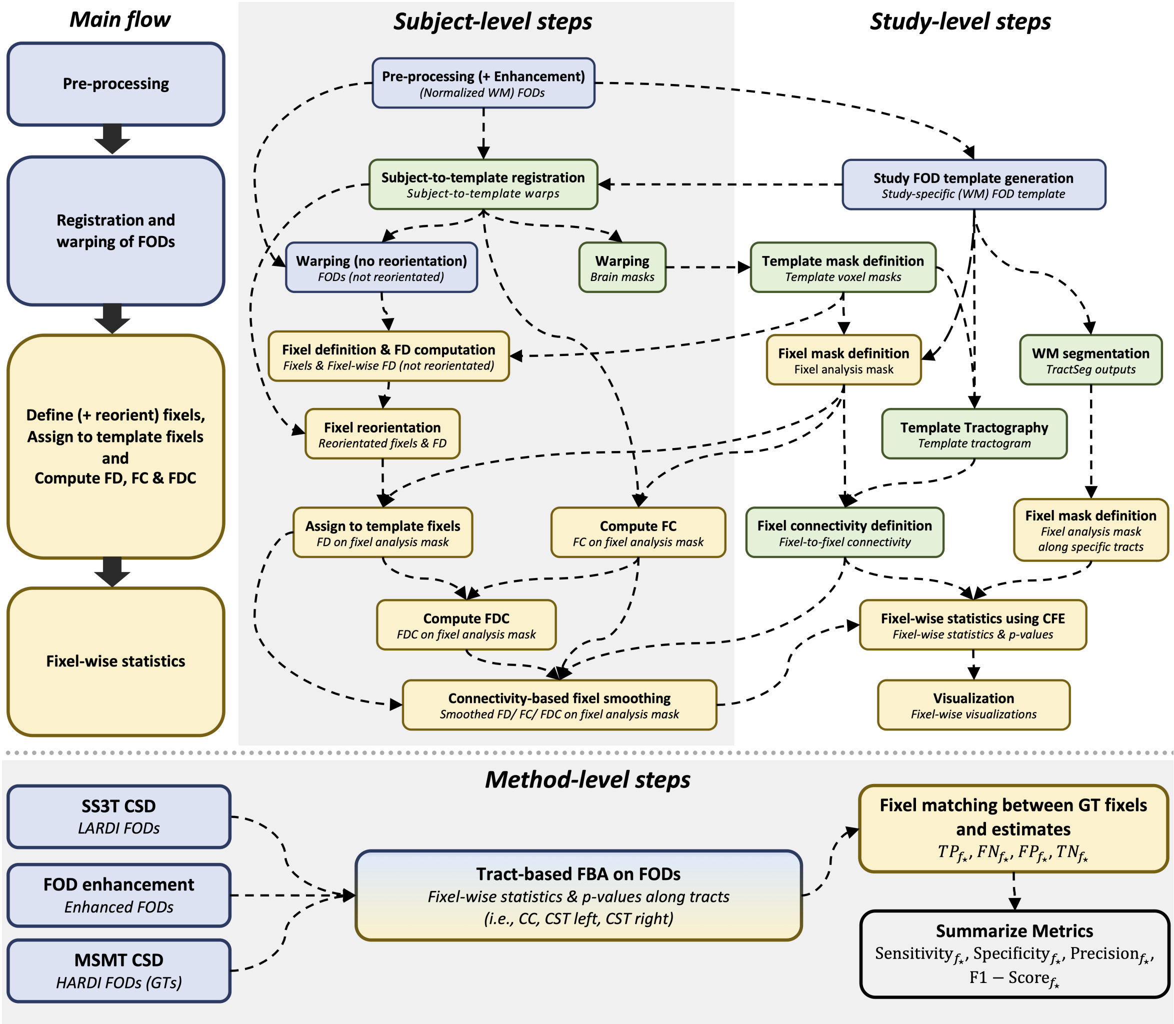


FIGURE 3.3. Our Fixel-Based Analysis (FBA) pipeline. Blue boxes denoting FOD data and yellow boxes denoting fixel-wise data. Each box names a processing step and its resulting output (in *italic*). Subject-level steps are performed once per individual subject, whereas study-level steps are operated only once for the study. For each FOD estimation method, i.e., SS3T CSD, FOD-Net, FastFOD-Net, and MSMT CSD, we performed a tract-based FBA study along specific tracts to generate fixel-wise statistics ($p < 0.05$) between clinical groups, and other subsequently method-level steps to compute quantitative metrics for comparisons between methods.

For that step we followed the methodology describe in the Section 3.2.5.2. Eventually, we determined whether the fixels showing significant differences between patients and controls in the multi-shell FODs (GT) could be also detected in our enhanced FOD estimates. To that end, some common detections are described below.

- $\mathbf{TP}$ represents the number of significantly different fixels correctly identified;
- $\mathbf{FN}$ represents the number of significantly different fixels that were not identified as such;
- $\mathbf{FP}$ represents the number of fixels incorrectly identified as significantly different when they are not;
- $\mathbf{TN}$ represents the number of fixels correctly identified as not significantly different.

From these, the $\mathbf{Sensitivity}$ metric was computed to measure the ability of a FOD enhancement technique to correctly identify fixels that were indeed significantly different, as formulated in Equation 3.5. $\mathbf{Specificity}$ (Equation 3.6 was also computed to measure the ability of each enhancement method to correctly identify fixels that are not significantly different. Finally, the $\mathbf{Precision}$ metric (Equation 3.7) measures the reliability when detecting significantly different fixels using enhanced FODs (i.e., the ratio between true positives and all significantly different fixels) and evaluates how well the method avoids false positives. To account for class imbalance between significantly different and non-significantly different fixels, we also included the F1-score (Equation 3.8).

$$\mathbf{Sensitivity} = \frac{\mathbf{TP}}{\mathbf{TP} + \mathbf{FN}} \tag{3.5}$$

$$\mathbf{Specificity} = \frac{\mathbf{TN}}{\mathbf{TN} + \mathbf{FP}} \tag{3.6}$$

$$\mathbf{Precision} = \frac{\mathbf{TP}}{\mathbf{TP} + \mathbf{FP}} \tag{3.7}$$

$$\mathbf{F1} = \frac{2\mathbf{TP}}{2\mathbf{TP} + \mathbf{FN} + \mathbf{FP}} \tag{3.8}$$

3.2.5.5 Clinical Research Application 2: Pathological Connection Analysis

For pathology-related connections, alterations of their connectivity value between clinical groups could be analyzed to further explore the practical impact of learning-based FOD enhancement in a real-world clinical study. Therefore, we hypothesized that the abnormal connection could serve as a potential biomarker, which should demonstrate a significant difference ($p < 0.05$) between patients and healthy controls in enhanced connectomes similar to that seen with multi-shell derived GTs. Furthermore, we also focused on a connection not expected to be affected, as this allow us to assess whether the enhancement process introduces potential false positive (abnormal) connections that, thereby ensuring the specificity of the biomarker identification process.

Previous research studies on MND have consistently highlighted deficits in the primary motor cortex among ALS patients [149]. Rose et al [150] proposed that measures of altered intra- and inter-hemispheric structural connectivity of the primary motor and somatosensory cortex would provide an improved assessment of corticomotor involvement in ALS. The findings demonstrate that the precentral gyrus interhemispheric connectivity would change between ALS patients and controls, while the postcentral gyrus interhemispheric connectivity would not change. To validate our hypothesis, we performed independent t-tests to measure whether specific structural connections from the primary motor cortex (i.e., Left PreCentral Cortex [L.PrCG] to the Right PreCentral Cortex [R.PrCG]) and from somatosensory cortex (i.e., Left Postcentral Cortex [L.PoCG] to the Right Postcentral Cortex [R.PoCG]) were significantly different between the ALS patient group and controls for each FOD method. We also performed analysis of sample size determination with respect to different levels of statistical power to assess if FastFOD-Net can effectively facilitate clinical studies. Effect size was also computed to quantitatively measure the magnitude difference between groups. Similarly, for the Tremor dataset, one-way ANOVA analysis was performed to determine whether there are significant differences of pathology-related connection, i.e., thalamus connectivity [127], between the means of three independent groups.

3.2.5.6 Clinical Research Application 3: Correlation Analysis

In addition to group analysis, we employed Pearson correlation analysis between the graph metrics derived from the estimated connectomes and clinical outcomes. This analysis aims to evaluate the potential of our method to reflect clinical outcomes, thereby demonstrating its applicability in clinical neuroscience.

3.3 Results

The experiments were conducted on datasets from the Human Connectome Project (HCP and HCP$_{Anomaly}$) [114], Tremor [127], MND, and MS - see demographics and data details in Figure 3.1c, Methods 3.2.2, and Table 3.1. Four different methods to compute the FOD were compared: SS3T CSD [117] and MSMT CSD [27], which are the state-of-the-art for analyzing single- and multi-shell DWI data, respectively; FOD-Net [104], a recent enhancement baseline method; and the new optimized FastFOD-Net method. The enhancement performance were evaluated by comparing pre-enhancement (SS3T CSD), post-enhancement (FOD-Net / FastFOD-Net), with the ground truth (GT - MSMT CSD).

3.3.1 FOD Assessment

The quantitative FOD assessments evaluated the signal-to-noise and angular accuracy in following brain regions [104]: (1) deep WM regions only; and (2) WM/GM border area including a mixture of WM and GM at juxtacortical (JCWM) or subcortical regions (WM-SGM). Figure 3.4a,b show that the proposed FastFOD-Net is significantly superior across all conditions for both FOD-level metrics - see details on metrics and results in Methods 3.2.5.1 and Table 3.3.

3.3.2 Fiber Bundle Element Assessment

The FOD data can therefore be discretized in the angular domain based on its fixels, where ‘fixel’ refers to a specific fiber bundle population within a voxel [125]. The impact of the

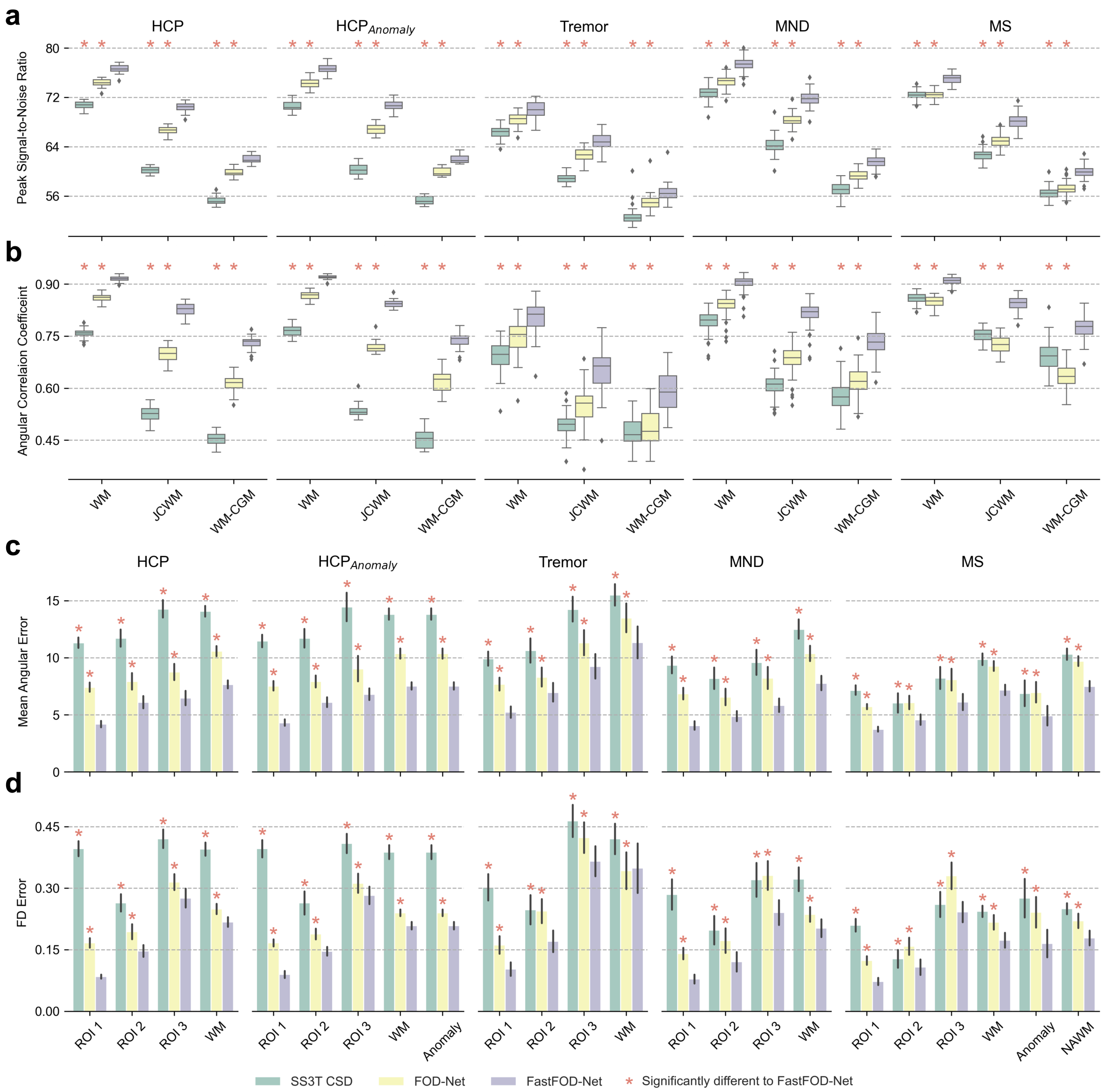


FIGURE 3.4. Quantitative analysis of FOD assessment and fixel assessment. Error bars represent standard derivation, and * denotes the results are significantly different from FastFOD-Net ($p < 0.05$). a, Box plots of Peak Signal-to-Noise Ratio (**PSNR**) between the ground truth (GT, corresponding to MSMT CSD) and other methods (SS3T CSD, FOD-Net, and FastFOD-Net) across patients and controls. WM denotes white matter; JCWM and WM-SGM represent the boundaries between WM and juxtacortical or subcortical gray matter. b, Box plots of angular correlation coefficient ($\mathbf{r_{Angular}}$) between GT and other methods. c, Bar plots of mean angular error ($\mu\mathbf{E_{Angular}}$) between GT fixels and other estimations across patients and controls. ROI 1, 2, and 3 represent areas with distinct fiber configurations: 1 single fiber population, 2-way fiber crossing, and 3-way fiber crossing, respectively. 'Anomaly' represents ROIs surrounding the anomalous regions on HCP$_{Anomaly}$ cases, and lesions on MS cases. NAWM represents normal appearing WM outside lesions. d, Bar plots of fiber density (FD) error ($\mathbf{E_{FD}}$).

TABLE 3.3. Quantitative analysis of mean peak signal-to-noise ratio[1] and angular correlation coefficient[2] of estimated FODs when compared to MSMT CSD-derived data[3].

		PSNR[1] (↑)				
Methods	Tissue Type	HCP	$\text{HCP}_{Anomaly}$[4]	Tremor	MND	MS
SS3T CSD [117]	WM[5]	70.7430* ± 0.6071	70.6472* ± 0.9050	66.3592* ± 1.0522	72.8055* ± 1.0879	72.3808* ± 0.7498
	JCWM[6]	60.2084* ± 0.5640	60.3150* ± 0.9830	58.8588* ± 0.7082	64.2163* ± 1.2248	62.6626* ± 0.8832
	WM-CGM[7]	55.2939* ± 0.7760	55.2862* ± 0.7453	52.7151* ± 1.4621	57.0784* ± 1.0483	56.4695* ± 0.8889
	Anomaly	NA	70.6472* ± 0.9050	NA	NA	50.6064* ± 5.4511
	NAWM	NA	NA	NA	NA	72.6303* ± 0.6927
FOD-Net [104]	WM	74.3551* ± 0.6627	74.3509* ± 0.9020	68.4481* ± 1.0767	74.6290* ± 0.9037	72.4098* ± 0.6572
	JCWM	66.6154* ± 0.6941	66.8379* ± 0.8597	62.7267* ± 1.0582	68.3410* ± 0.9037	64.9100* ± 0.8907
	WM-CGM	59.8562* ± 0.7748	59.9640* ± 0.7303	55.0599* ± 1.3900	59.3251* ± 0.9110	57.2308* ± 0.8559
	Anomaly	NA	74.3509* ± 0.9020	NA	NA	50.7694* ± 5.7220
	NAWM	NA	NA	NA	NA	72.7749* ± 0.6231
FastFOD-Net	WM	**76.6887** ± 0.7102	**76.6642** ± 0.9091	**70.0695** ± 1.2042	**77.3696** ± 1.0330	**75.0700** ± 0.7714
	JCWM	**70.3790** ± 0.8130	**70.6622** ± 0.9859	**64.8251** ± 1.3154	**71.8760** ± 1.1395	**68.0943** ± 1.0352
	WM-CGM	**61.9567** ± 0.7450	**62.1020** ± 0.7746	**56.5891** ± 1.3888	**61.5614** ± 0.9755	**59.8748** ± 0.8822
	Anomaly	NA	**76.6642** ± 0.9091	NA	NA	**54.0927** ± 6.1493
	NAWM	NA	NA	NA	NA	**75.4318** ± 0.7438
		$\mathbf{r_{Angular}}$[2] (↑)				
Methods	Tissue Type	HCP	$\text{HCP}_{Anomaly}$	Tremor	MND	MS
SS3T CSD	WM	0.7590* ± 0.0157	0.7652* ± 0.0167	0.6936* ± 0.0458	0.7917* ± 0.0280	0.8594* ± 0.0133
	JCWM	0.5265* ± 0.0227	0.5367* ± 0.0234	0.4960* ± 0.0346	0.6091* ± 0.0298	0.7535* ± 0.0185
	WM-CGM	0.4534* ± 0.0186	0.4554* ± 0.0307	0.4744* ± 0.0401	0.5755* ± 0.0394	0.6930* ± 0.0396
	Anomaly	NA	0.7652* ± 0.0167	NA	NA	0.8978* ± 0.0325
	NAWM	NA	NA	NA	NA	0.8442* ± 0.0138
FOD-Net	WM	0.8614* ± 0.0124	0.8675* ± 0.0117	0.7467* ± 0.0519	0.8404* ± 0.0247	0.8491* ± 0.0141
	JCWM	0.7013* ± 0.0244	0.7205* ± 0.0204	0.5534* ± 0.0590	0.6854* ± 0.0358	0.7258* ± 0.0247
	WM-CGM	0.6115* ± 0.0252	0.6196* ± 0.0362	0.4904* ± 0.0554	0.6219* ± 0.0399	0.6368* ± 0.0305
	Anomaly	NA	0.8675* ± 0.0117	NA	NA	0.8979* ± 0.0296
	NAWM	NA	NA	NA	NA	0.8293* ± 0.0145
FastFOD-Net	WM	**0.9155** ± 0.0086	**0.9201** ± 0.0072	**0.8045** ± 0.0490	**0.9042** ± 0.0193	**0.9097** ± 0.0108
	JCWM	**0.8286** ± 0.0200	**0.8439** ± 0.0130	**0.6549** ± 0.0623	**0.8147** ± 0.0321	**0.8447** ± 0.0191
	WM-CGM	**0.7294** ± 0.0224	**0.7372** ± 0.0288	**0.5933** ± 0.0577	**0.7335** ± 0.0370	**0.7738** ± 0.0276
	Anomaly	NA	**0.9201** ± 0.0072	NA	NA	**0.9263** ± 0.0303
	NAWM	NA	NA	NA	NA	**0.8979** ± 0.0113

[3]MSMT CSD-derived data is ground truth (GT). [4]$\text{HCP}_{Anomaly}$ denotes HCP healthy subjects with structural abnormalities. [5]WM denotes pure white matter. [6]JCWM and [7]WM-SGM represent the boundaries between WM and juxtacortical or subcortical gray matter. The best result is highlighted in **bold**, and * denotes the results are significantly different from FastFOD-Net ($p < 0.05$).

FOD enhancement can be then assessed using fixel-wise metrics, reflecting the angular and amplitude accuracy - see Methods 3.2.5.2 for details. Figure 3.4c,d demonstrate that FastFOD-Net significantly ($p < 0.05$) outperformed SS3T CSD and FOD-Net in terms of mean angular error $\mu\mathbf{E_{Angular}}$ and FD error $\mathbf{E_{FD}}$ of fiber bundle elements in all ROIs considered (see Methods 3.2.5.2 for ROI definitions) and all datasets. Figure 3.5 shows FastFOD-Net obtained comparable peak errors $\mathbf{E_{Peak}}$ to its predecessor FOD-Net. Although the improvements brought by FastFOD-Net sometimes were reduced in ROI with two fibers

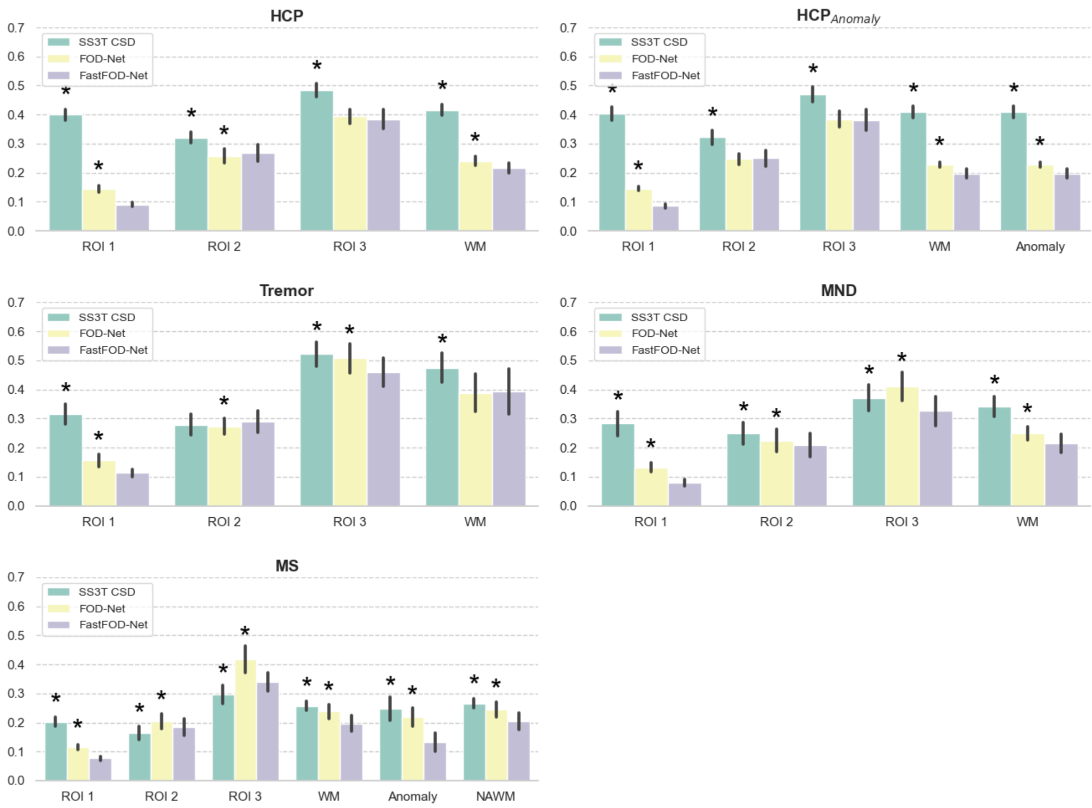


FIGURE 3.5. Quantitative analysis of fixel assessment for the peak property. Bar plots of peak error ($\mathbf{E_{Peak}}$) between ground truth (GT, corresponding to MSMT CSD) and other methods (SS3T CSD, FOD-Net, and FastFOD-Net) across patients and controls. ROI 1, 2, and 3 represent areas with distinct fiber configurations: 1 fiber population, 2-way fiber crossing, and 3-way fiber crossing, respectively. 'Anomaly' represents ROIs surrounding the anomalous regions on HCP$_{Anomaly}$ cases, and lesions on MS cases. WM denotes white matter and NAWM represents normal appearing white matter outside lesions.

crossing (ROI 2, where the middle cerebellar peduncle crosses the corticospinal tract), they became more significant on more complex the fiber configurations (ROI 3, where the superior longitudinal fasciculus, the corticospinal tract and corpus callosum cross). More detailed results are provided in Table 3.4, 3.5 and 3.6.

TABLE 3.4. Quantitative analysis of mean angular error[1] between estimated fixels and MSMT CSD-derived data[2].

		$\mu\mathbf{E}_{\mathbf{Angular}}$ [1] (↓)				
Methods	ROIs[3]	HCP	$\mathrm{HCP}_{Anomaly}$	Tremor	MND	MS
SS3T CSD	ROI 1	11.3195* ± 0.4641	11.4658* ± 0.5575	9.9299* ± 0.6195	9.3661* ± 0.7447	7.1650* ± 0.4200
	ROI 2	11.7188* ± 0.7512	11.7127* ± 0.8148	10.6309* ± 1.0655	8.1924* ± 0.9425	6.0463* ± 0.8429
	ROI 3	14.2819* ± 0.7576	14.4520* ± 1.2459	14.2531* ± 1.0848	9.5959* ± 1.1047	8.2231* ± 0.9621
	WM	14.0733* ± 0.4640	13.8267* ± 0.4996	15.5042* ± 0.9480	12.5016* ± 0.8436	9.8694* ± 0.5014
	Anomaly	NA	13.8267* ± 0.4996	NA	NA	6.8722* ± 1.1291
	NAWM	NA	NA	NA	NA	10.3116* ± 0.4946
FOD-Net	ROI 1	7.4300* ± 0.4019	7.5356* ± 0.4189	7.6910* ± 0.5656	6.8654* ± 0.5113	5.7171* ± 0.2357
	ROI 2	7.9312* ± 0.7158	7.9235* ± 0.5172	8.3075* ± 0.8332	6.5644* ± 0.7305	6.0729* ± 0.5926
	ROI 3	8.7554* ± 0.7193	9.0422* ± 1.1177	11.3200* ± 1.1015	8.2244* ± 0.9652	8.0804* ± 0.9460
	WM	10.5850* ± 0.4406	10.3668* ± 0.4374	13.4913* ± 1.2522	10.3761* ± 0.6715	9.2633* ± 0.4295
	Anomaly	NA	10.3668* ± 0.4374	NA	NA	6.9770* ± 0.8948
	NAWM	NA	NA	NA	NA	9.7037* ± 0.4174
FastFOD-Net	ROI 1	**4.1873** ± 0.2889	**4.3173** ± 0.2846	**5.2427** ± 0.4955	**4.0590** ± 0.3721	**3.7379** ± 0.2065
	ROI 2	**6.0944** ± 0.5396	**6.0960** ± 0.4258	**6.9762** ± 0.8086	**4.8787** ± 0.4501	**4.5656** ± 0.4558
	ROI 3	**6.4689** ± 0.6401	**6.8098** ± 0.5127	**9.2458** ± 1.0695	**5.8457** ± 0.5932	**6.1278** ± 0.7135
	WM	**7.6581** ± 0.3529	**7.5290** ± 0.3278	**11.3303** ± 1.3980	**7.7757** ± 0.6271	**7.1719** ± 0.4620
	Anomaly	NA	**7.5290** ± 0.3278	NA	NA	**4.9169** ± 0.8623
	NAWM	NA	NA	NA	NA	**7.5016** ± 0.4554

[2] MSMT CSD-derived data is ground truth (GT). [3] The regions labeled as ROI 1, 2, and 3 represent areas with distinct fiber configurations, each occurring with a varying number of fixels (1, 2, and 3, respectively); WM is short for white matter; Anomaly represents ROIs surrounding the anomalous regions on the $\mathrm{HCP}_{Anomaly}$ cases, and lesions on MS cases, respectively; and NAWM represents normal appearing WM outside lesions. The best result is highlighted in **bold**, and * denotes the results are significantly different from FastFOD-Net ($p < 0.05$).

TABLE 3.5. Quantitative analysis of FD error[1] of fixels between estimates fixels and MSMT CSD-derived data[2].

Methods	ROIs[3]	$\mathbf{E_{FD}}^1$ (↓) HCP	$\mathrm{HCP}_{Anomaly}$[4]	Tremor	MND	MS
SS3T CSD	ROI 1	0.3964* ± 0.0183	0.3964* ± 0.0212	0.3023* ± 0.0321	0.2849* ± 0.0369	0.2099* ± 0.0154
	ROI 2	0.2644* ± 0.0211	0.2642* ± 0.0282	0.2476* ± 0.0356	0.1978* ± 0.0348	0.1280* ± 0.0216
	ROI 3	0.4206* ± 0.0225	0.4091* ± 0.0236	0.4641* ± 0.0392	0.3205* ± 0.0413	0.2605* ± 0.0305
	WM	0.3955* ± 0.0161	0.3881* ± 0.0169	0.4203* ± 0.0371	0.3221* ± 0.0291	0.2437* ± 0.0138
	Anomaly	NA	0.3881* ± 0.0169	NA	NA	0.2759* ± 0.0469
	NAWM	NA	NA	NA	NA	0.2500* ± 0.0135
FOD-Net	ROI 1	0.1668* ± 0.0117	0.1668* ± 0.0086	0.1616* ± 0.0215	0.1408* ± 0.0140	0.1240* ± 0.0104
	ROI 2	0.1939* ± 0.0183	0.1884* ± 0.0131	0.2444* ± 0.0292	0.1724* ± 0.0298	0.1585* ± 0.0210
	ROI 3	0.3151* ± 0.0194	0.3123* ± 0.0233	0.4235* ± 0.0375	0.3312* ± 0.0351	0.3303* ± 0.0326
	WM	0.2495* ± 0.0126	0.2401* ± 0.0080	**0.3428*** ± 0.0448	0.2363* ± 0.0180	0.2171* ± 0.0178
	Anomaly	NA	0.2401* ± 0.0080	NA	NA	0.2414* ± 0.0373
	NAWM	NA	NA	NA	NA	0.2207* ± 0.0177
FastFOD-Net	ROI 1	**0.0845** ± 0.0049	**0.0898** ± 0.0087	**0.1029** ± 0.0163	**0.0786** ± 0.0107	**0.0728** ± 0.0087
	ROI 2	**0.1471** ± 0.0144	**0.1464** ± 0.0107	**0.1707** ± 0.0263	**0.1208** ± 0.0237	**0.1076** ± 0.0187
	ROI 3	**0.2761** ± 0.0228	**0.2828** ± 0.0213	**0.3660** ± 0.0365	**0.2409** ± 0.0303	**0.2421** ± 0.0250
	WM	**0.2177** ± 0.0119	**0.2088** ± 0.0092	0.3491 ± 0.0603	**0.2027** ± 0.0217	**0.1733** ± 0.0181
	Anomaly	NA	**0.2088** ± 0.0092	NA	NA	**0.1654** ± 0.0336
	NAWM	NA	NA	NA	NA	**0.1787** ± 0.0178

[2] MSMT CSD-derived data is ground truth (GT). [3] The regions labeled as ROI 1, 2, and 3 represent areas with distinct fiber configurations, each occurring with a varying number of fixels (1, 2, and 3, respectively); WM is short for white matter; Anomaly represents ROIs surrounding the anomalous regions on the $\mathrm{HCP}_{Anomaly}$ cases, and lesions on MS cases, respectively; and NAWM represents normal appearing WM outside lesions. The best result is highlighted in **bold**, and * denotes the results are significantly different from FastFOD-Net ($p < 0.05$).

TABLE 3.6. Quantitative analysis of peak error[1] of fixels between estimated fixels and multi-shell HARDI data[2].

		$\mathbf{E_{Peak}}^1$ (↓)				
Methods	ROIs[3]	HCP	$\text{HCP}_{Anomaly}{}^4$	Tremor	MND	MS
SS3T CSD	ROI 1	0.3996* ± 0.0184	0.4036* ± 0.0218	0.3159* ± 0.0348	0.2827* ± 0.0416	0.2027* ± 0.0149
	ROI 2	0.3211* ± 0.0189	0.3219* ± 0.0254	0.2787 ± 0.0362	0.2489* ± 0.0369	**0.1644*** ± 0.0231
	ROI 3	0.4832* ± 0.0226	0.4701* ± 0.0254	0.5225* ± 0.0416	0.3710* ± 0.0449	**0.2964*** ± 0.0311
	WM	0.4165* ± 0.0179	0.4095* ± 0.0193	0.4742* ± 0.0501	0.3411* ± 0.0346	0.2566* ± 0.0158
	Anomaly	NA	0.4095* ± 0.0193	NA	NA	0.2484* ± 0.0399
	NAWM	NA	NA	NA	NA	0.2655* ± 0.0156
FOD-Net	ROI 1	0.1455* ± 0.0111	0.1453* ± 0.0073	0.1560* ± 0.0208	0.1322* ± 0.0151	0.1152* ± 0.0086
	ROI 2	**0.2581*** ± 0.0246	**0.2475** ± 0.0186	**0.2734*** ± 0.0262	0.2241* ± 0.0384	0.2042* ± 0.0268
	ROI 3	0.3940 ± 0.0254	0.3839 ± 0.0276	0.5074* ± 0.0498	0.4112* ± 0.0484	0.4166* ± 0.0461
	WM	0.2408* ± 0.0155	0.2279* ± 0.0091	**0.3888** ± 0.0657	0.2480* ± 0.0231	0.2380* ± 0.0253
	Anomaly	NA	0.2279* ± 0.0091	NA	NA	0.2196* ± 0.0321
	NAWM	NA	NA	NA	NA	0.2450* ± 0.0250
FastFOD-Net	ROI 1	**0.0906** ± 0.0075	**0.0863** ± 0.0065	**0.1126** ± 0.0121	**0.0805** ± 0.0115	**0.0768** ± 0.0074
	ROI 2	0.2688 ± 0.0287	0.2497 ± 0.0275	0.2887 ± 0.0373	**0.2089** ± 0.0413	0.1841 ± 0.0282
	ROI 3	**0.3847** ± 0.0342	**0.3814** ± 0.0363	**0.4603** ± 0.0494	**0.3263** ± 0.0511	0.3394 ± 0.0314
	WM	**0.2161** ± 0.0181	**0.1974** ± 0.0163	0.3929 ± 0.0784	**0.2140** ± 0.0319	**0.1970** ± 0.0283
	Anomaly	NA	**0.1974** ± 0.0163	NA	NA	**0.1315** ± 0.0318
	NAWM	NA	NA	NA	NA	**0.2046** ± 0.0277

[3]Multi-shell HARDI data is ground truth (GT). [4]$\text{HCP}_{Anomaly}$ denotes HCP healthy subjects with structural abnormalities. [5]WM denotes pure white matter. [6]JCWM and [7]WM-SGM represent the boundaries between WM and juxtacortical or subcortical gray matter. The best result is highlighted in **bold**, and * denotes the results are significantly different from FastFOD-Net ($p < 0.05$).

3.3.3 Connectome Assessment

The structural connectome matrix computed from diffusion MRI fiber-tracking data serves as a summary depiction of the whole brain connectivity, offering insights into circuit-based alterations associated with neurological and psychiatric disorders [151, 152]. Details of the metrics used to reflect the accuracy of the estimated connectomes when compared to the GT connectome computed with MSMT CSD of multi-shell HARDI data are provided in Methods 3.2.5.3.

FastFOD-Net produced high-quality connectomes, with the least number of significant different edges relative to GT (Table 3.7 and Figure 3.8), the highest Kendall correlation (Table 3.7 and Figure 3.6), and the lowest disparity (Figure 3.7 and Table 3.7), surpassing both SS3T CSD and FOD-Net by a substantial margin.

To zoom in on the differences between the learning-based methods, FastFOD-Net reduced the disparity of inter-hemispheric connections (Figure 3.9 and 3.10), while FOD-Net did not reduce and even exacerbated the disparity on subjects with tremors and MND.

Finally, the network topology of the connectome is often summarized with graph theoretical metrics [153]. As illustrated by Figure 3.11, the DR to the GT for FastFOD-Net is the closest to 0% for most graph measures, confirming its superiority compared to the other LARDI based methods. Details of results on DR of graph metrics are provided in Table 3.8 and 3.9.

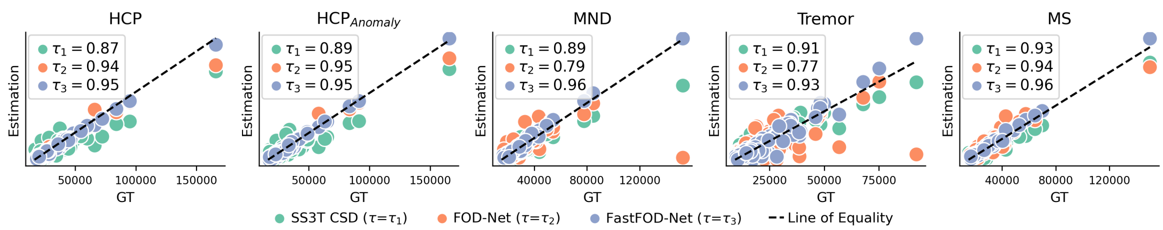


FIGURE 3.6. Correlation analysis of mean connectome matrices between GT via MSMT CSD and other estimates via SS3T CSD, FOD-Net and FastFOD-Net on HCP, HCP$_{Anomaly}$, Tremor, MND, and MS cases. τ denotes Kendall ranking correlation coefficient.

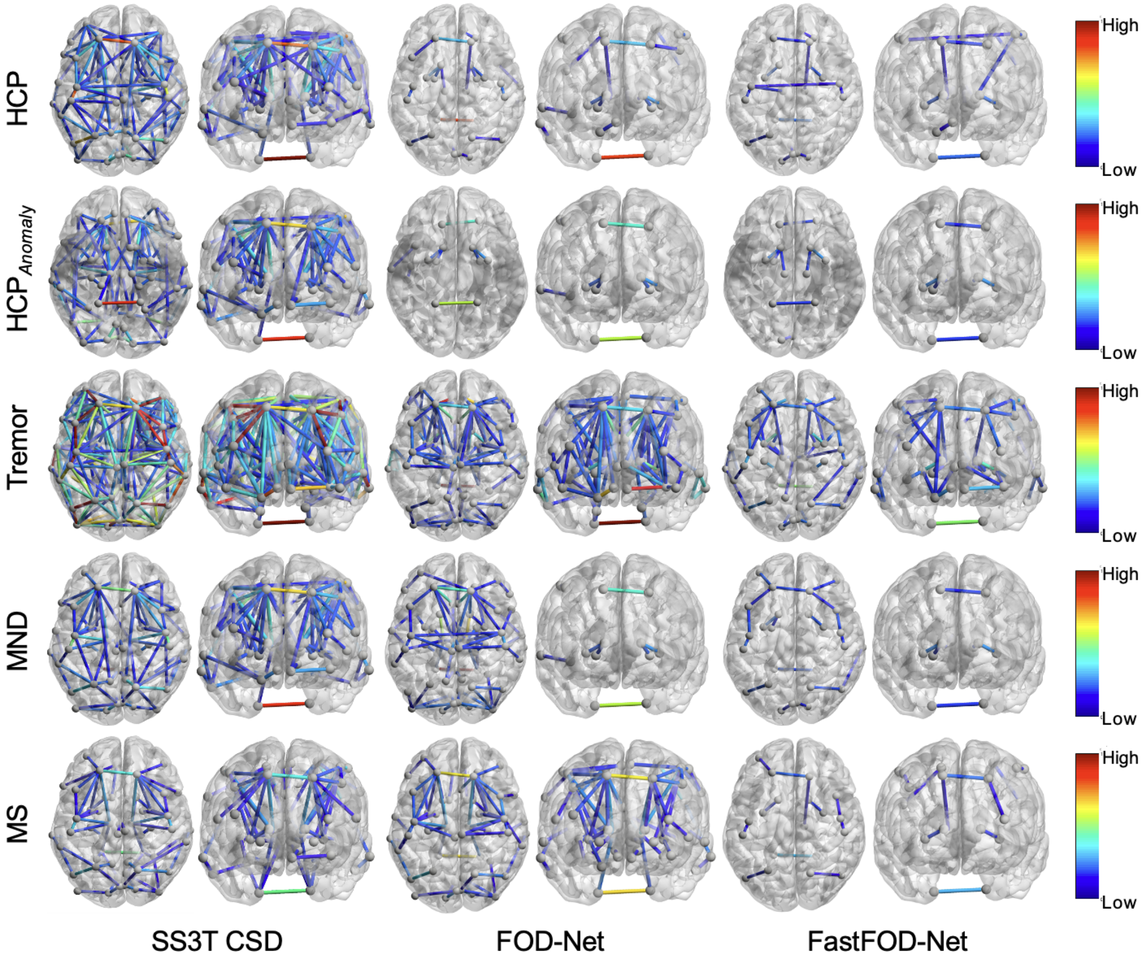


FIGURE 3.7. Quantitative analysis of disparity of connectomes. Axial and coronal views of mean edge disparity matrices between GT and each LARDI estimation method. The color scale ranges from blue to red, representing edges with a gradient from low to high disparity. For sparsity, the visualization threshold is set to 4000.

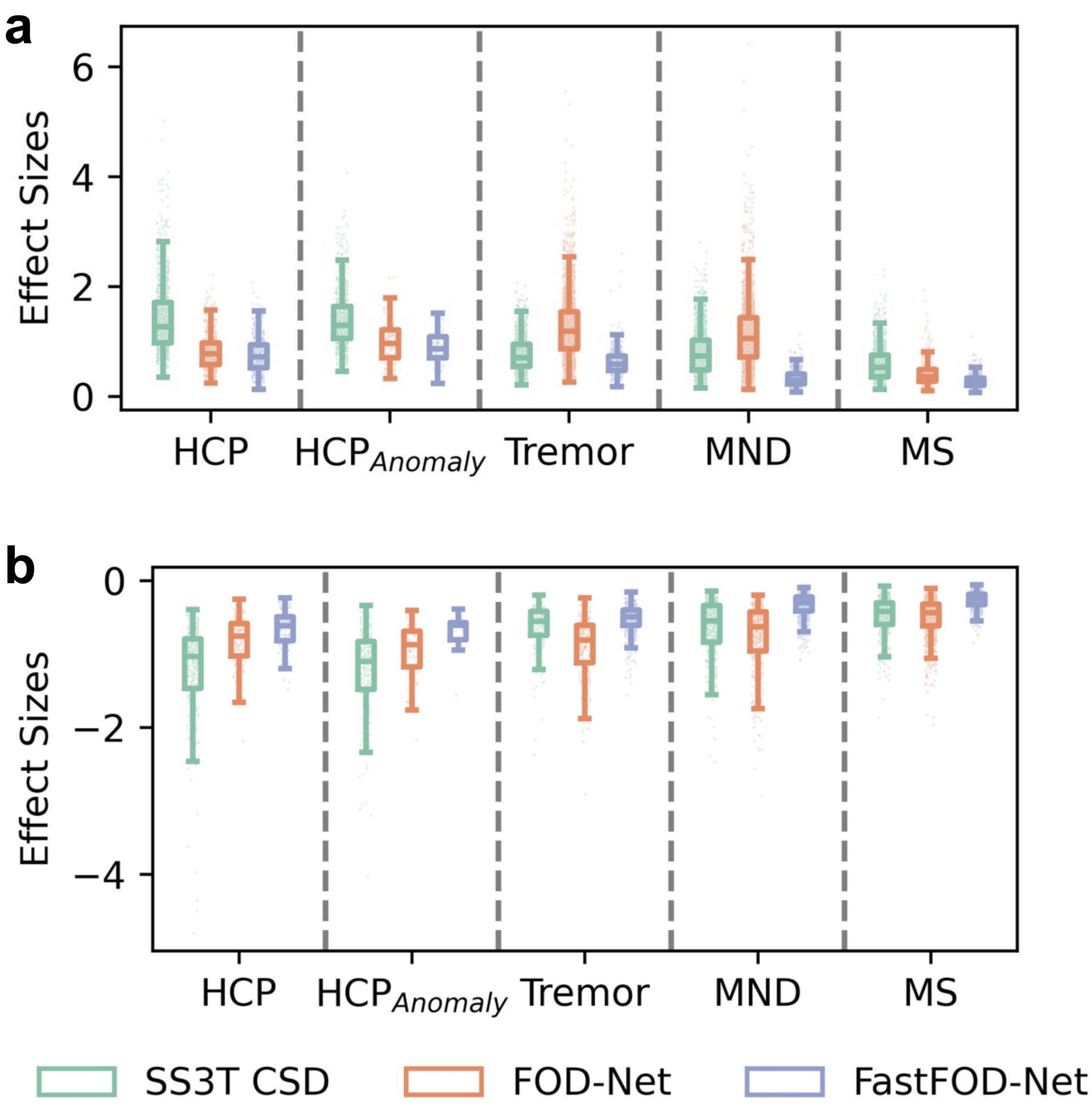


FIGURE 3.8. Effect sizes of significantly different edges in connectomes between MSMT CSD (ground truth, GT) and alternative methods (SS3T CSD, FOD-Net and FastFOD-Net). a, Positive effect sizes indicates MSMT CSD yields higher connectivity values compared to other methods. b, Negative effect sizes indicates MSMT-CSD yields lower connectivity values compared to other methods. Box plots represent the distribution of effect sizes across different datasets.

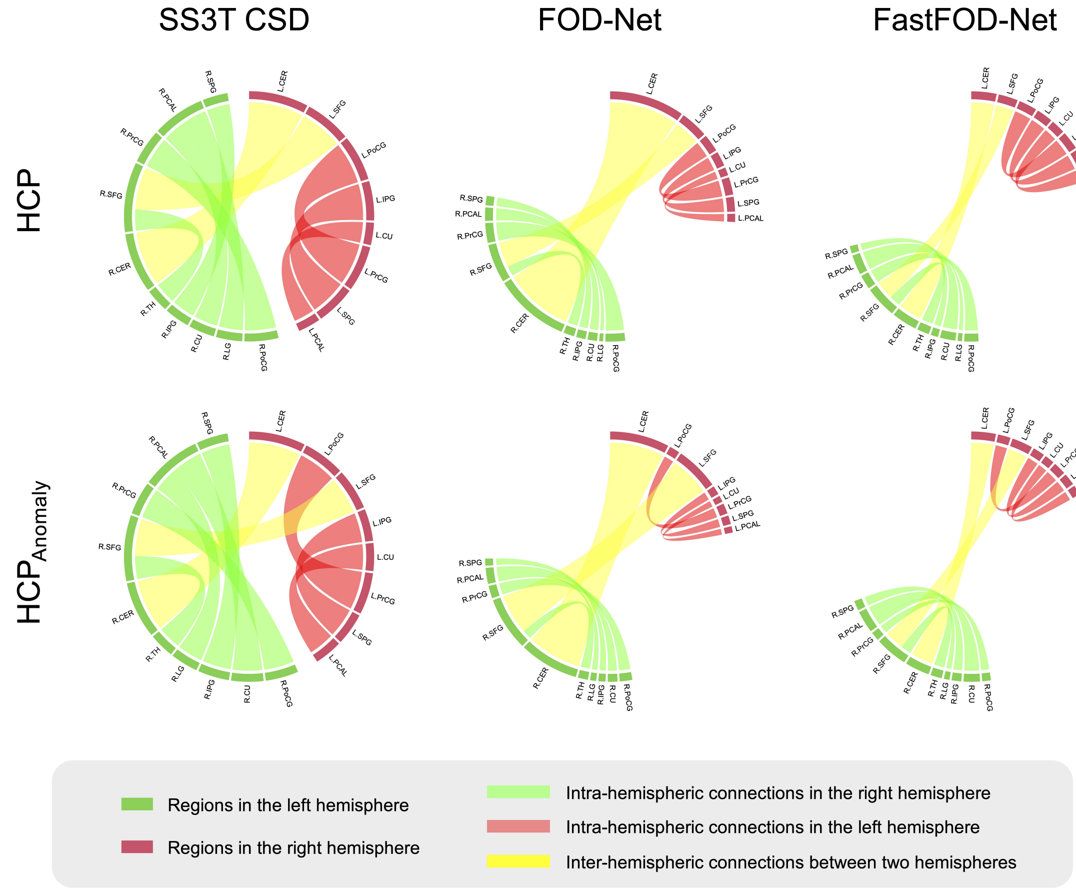


FIGURE 3.9. Chord diagrams illustrating the top 10 connections with the largest disparities between ground truth (GT, MSMT CSD) and other methods (SS3T-CSD, FOD-Net, and FastFOD-Net) on both the HCP and HCP$_{Anomaly}$ datasets. These connections were selected based on the edge disparity matrices comparing SS3T CSD to the GT. Nodes represent brain regions, with intra-hemispheric connections within the left hemisphere depicted in green, those within the right hemisphere in red, and inter-hemispheric connections in yellow. For full names of nodes, refer to Table 3.2.

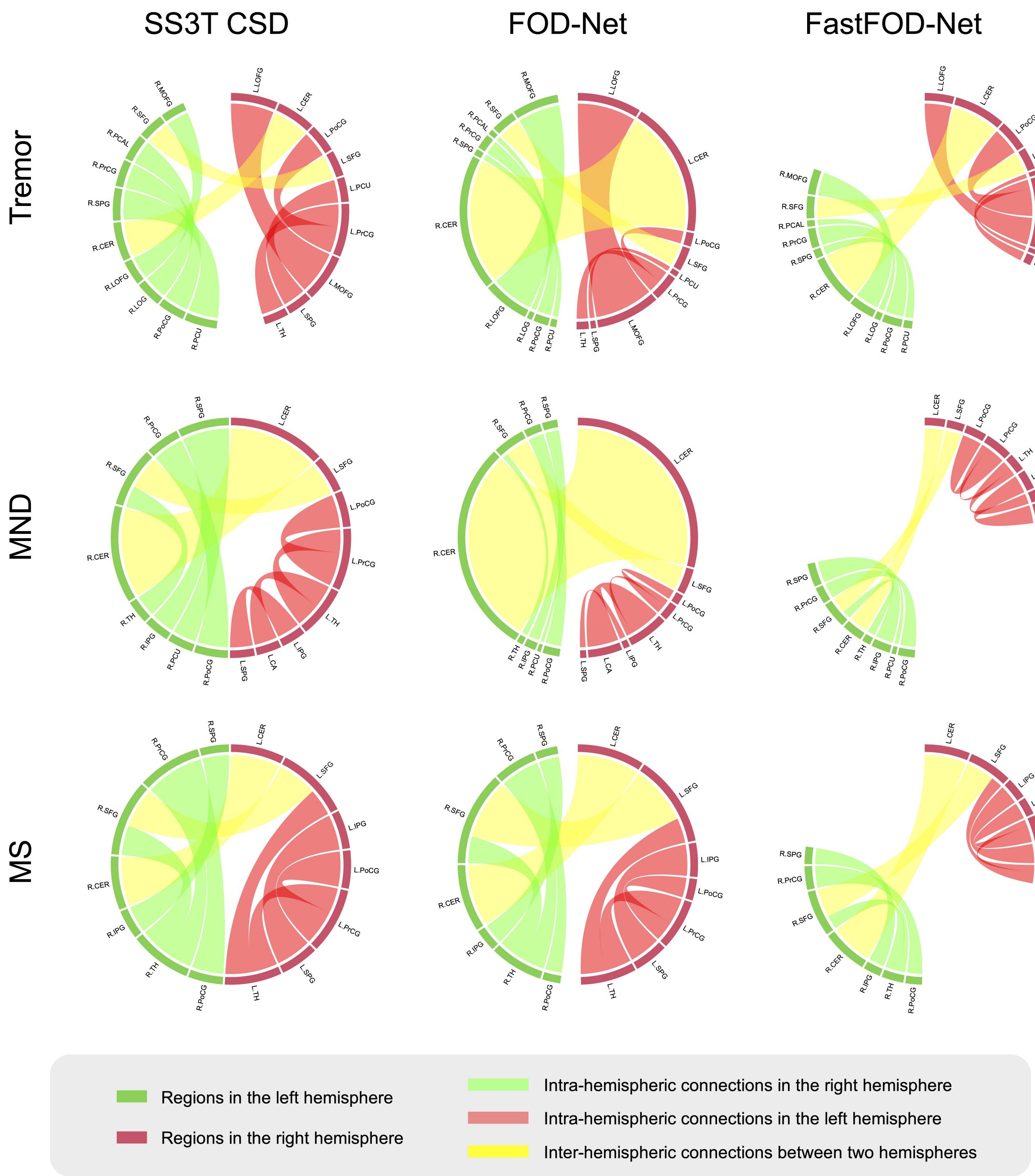


FIGURE 3.10. Chord diagrams illustrating the top 10 connections with the largest disparities between ground truth (GT, MSMT CSD) and other methods (SS3T-CSD, FOD-Net, and FastFOD-Net) on Tremor, MND and MS dataset. These connections were selected based on the edge disparity matrices comparing SS3T CSD to the GT. Nodes represent brain regions, with intra-hemispheric connections within the left hemisphere depicted in green, those within the right hemisphere in red, and inter-hemispheric connections in yellow. For full names of nodes, refer to Table 3.2.

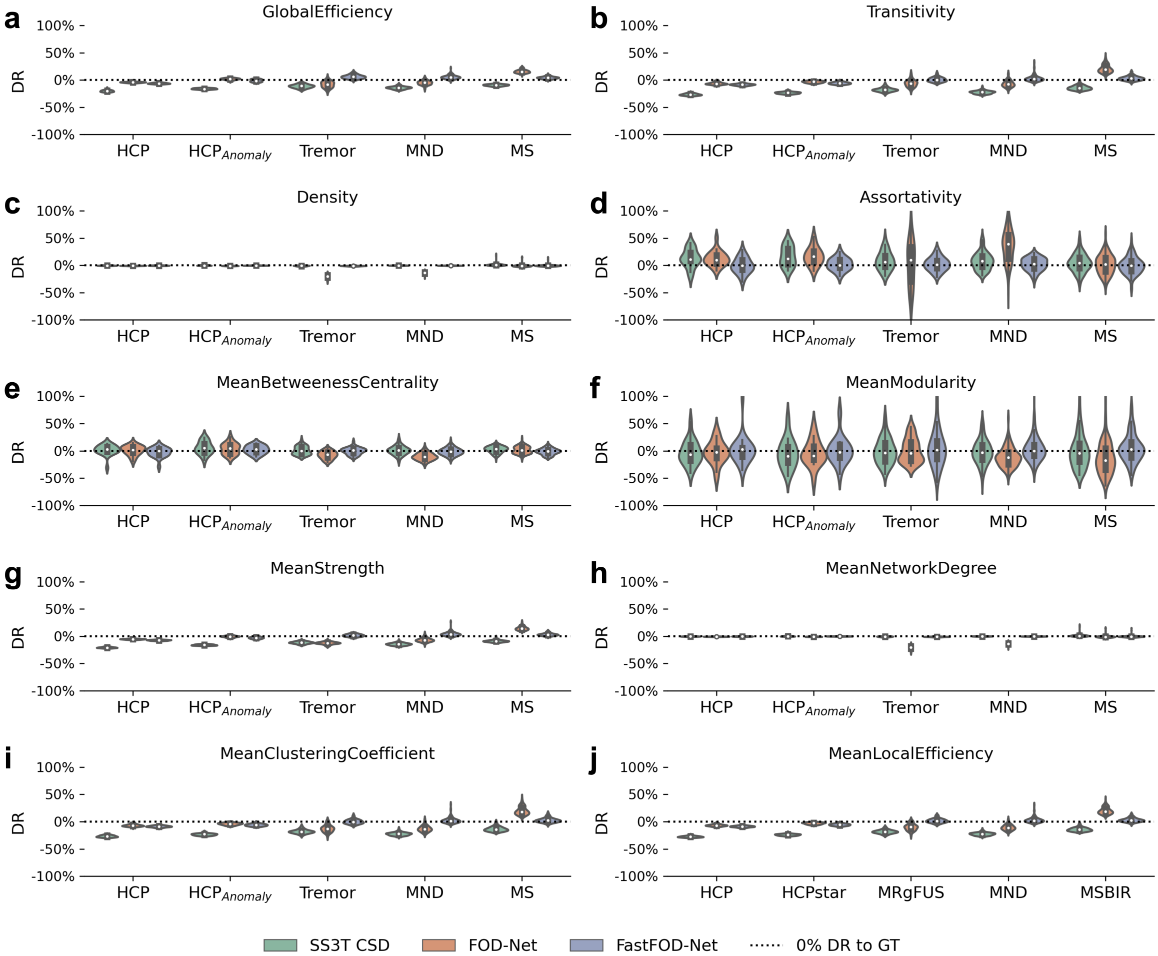


FIGURE 3.11. Violin plots of difference ratio (DR) of graph metrics between the ground truth (GT) vis MSMT CSD and other estimations via SS3T CSD, FOD-Net and FastFOD-Net on HCP, HCP$_{Anomaly}$, Tremor, MND, and MS cases. Metrics analyzed include a, Global Efficiency; b, Transitivity; c, Density; d, Assortativity; e, Mean Betweenness Centrality; f, Mean Modularity; g, Mean Strength; h, Mean Network Degree; i, Mean Clustering Coefficient; and j, Mean Local Efficiency.

TABLE 3.7. Quantitative analysis of connectome assessment between MSMT CSD-derived connectomes[1] and other LARDI-based estimates.

	τ^2 (↑)				
Methods	HCP	$HCP_{Anomaly}$[3]	Tremor	MND	MS
SS3T CSD	0.8484 ± 0.0075	0.8590 ± 0.0092	0.8739 ± 0.0179	0.8551 ± 0.0149	0.8672 ± 0.0237
FOD-Net	0.8991 ± 0.0063	0.9029 ± 0.0079	0.7415 ± 0.0198	0.7481 ± 0.0188	0.8673 ± 0.0227
FastFOD-Net	**0.9169** ± 0.0058	**0.9186** ± 0.0059	**0.9050** ± 0.0132	**0.9100** ± 0.0125	**0.9018** ± 0.0139
	μDisparity[4] (↓)				
Methods	HCP	$HCP_{Anomaly}$	Tremor	MND	MS
SS3T CSD	553.1383	488.6859	444.9234	380.1648	357.9414
FOD-Net	261.0568	230.0609	506.5477	573.3384	414.1002
FastFOD-Net	**218.0480**	**188.9277**	**219.3831**	**336.5554**	**186.4356**
	Significant Edge Difference[5] (↓)				
Methods	HCP	$HCP_{Anomaly}$[4]	Tremor	MND	MS
SS3T CSD	50.57%	36.49%	48.71%	68.42%	54.84%
FOD-Net	24.04%	6.83%	73.55%	75.96%	50.08%
FastFOD-Net	**22.69%**	**1.81%**	**30.01%**	**32.53%**	**38.49%**

[1]Connectome via MSMT CSD is ground truth (GT). [3]$HCP_{Anomaly}$ denotes HCP healthy subjects with structural abnormalities. [2]τ and [4]**μDisparity** denotes mean Kendall ranking correlation coefficient and mean disparity across all cases. [5]The number of significantly different edges (edge-wise paired t-tests with $p < 0.05$) are presented in percentage (%). The best result is highlighted in **bold**.

TABLE 3.8. Quantitative analysis of graph metrics[1] of connectomes between estimates and multi-shell HARDI data[2].

Methods	Graph Metrics[3]	HCP	$HCP_{Anomaly}$[4]	Tremor	MND	MS
SS3T CSD [117]	Assortativity	-0.0220	-0.0224	-0.0273	-0.0204	-0.0307
	DR[5]	13.04%	16.46%	8.38%	8.18%	4.36%
FOD-Net [104]	Assortativity	-0.0220	-0.0227	-0.0252	-0.0258	-0.0303
	DR	12.49%	16.85%	2.80%	36.93%	2.02%
FastFOD-Net	Assortativity	-0.0196	-0.0199	-0.0259	-0.0194	-0.0297
	DR	-0.51%	2.29%	2.47%	2.01%	-0.17%
MSMT CSD [27]	Assortativity	-0.0197	-0.0194	-0.0256	-0.0190	-0.0301
	DR	NA	NA	NA	NA	NA
SS3T CSD	Density	0.8469	0.8474	0.8387	0.8487	0.8194
	DR	-0.29%	0.08%	-0.86%	-0.19%	1.49%
FOD-Net	Density	0.8436	0.8397	0.6705	0.7312	0.8032
	DR	-0.67%	-0.84%	-20.76%	-14.01%	-0.53%
FastFOD-Net	Density	0.8463	0.8449	0.8372	0.8470	0.8055
	DR	-0.36%	-0.22%	-1.05%	-0.39%	-0.26%
MSMT CSD	Density	0.8494	0.8468	0.8461	0.8503	0.8082
	DR	NA	NA	NA	NA	NA
SS3T CSD	Global Efficiency	4447.0388	4464.7697	4319.4735	4182.7957	4400.0063
	DR	-20.08%	-16.12%	-11.39%	-14.12%	-9.29%
FOD-Net	Global Efficiency	5338.0550	5411.4844	4484.0406	4627.5225	5608.3059
	DR	-4.10%	1.68%	-8.12%	-5.01%	15.71%
FastFOD-Net	Global Efficiency	5219.9786	5265.3119	5180.2524	5132.2079	5053.7391
	DR	-6.20%	-1.06%	6.10%	5.43%	4.22%
MSMT CSD	Global Efficiency	5566.7768	5323.2690	4882.6422	4873.7886	4851.8529
	DR	NA	NA	NA	NA	NA
SS3T CSD	Mean Strength	101702.3121	101833.3209	102284.3668	95626.6352	98913.0292
	DR	-20.61%	-15.95%	-11.59%	-14.40%	-8.60%
FOD-Net	Mean Strength	121284.3588	120383.1025	101168.4562	104222.5172	124110.1985
	DR	-5.33%	-0.62%	-12.58%	-6.69%	14.77%
FastFOD-Net	Mean Strength	118756.9353	117726.1317	117652.9900	116287.9850	111721.8927
	DR	-7.30%	-2.81%	1.63%	4.14%	3.24%
MSMT CSD	Mean Strength	128106.0299	121139.1854	115762.0262	111809.4013	108339.0105
	DR	-	-	-	-	-
SS3T CSD	Transitivity	400.9858	390.9809	381.3502	368.1949	366.7435
	DR	-26.56%	-23.35%	-17.80%	-22.28%	-14.07%
FOD-Net	Transitivity	510.6750	492.3965	434.2220	440.5912	514.7142
	DR	-6.47%	-3.48%	-6.10%	-6.75%	20.82%
FastFOD-Net	Transitivity	499.1105	480.6893	468.3477	486.3475	440.6813
	DR	-8.58%	-5.74%	0.81%	2.86%	3.14%
MSMT CSD	Transitivity	545.9997	510.1657	465.3202	474.3285	428.6459
	DR	-	-	-	-	-

[3]Multi-shell HARDI data is ground truth (GT). [4]$HCP_{Anomaly}$ denotes HCP healthy subjects with structural abnormalities. For each method, [5]DR is the difference ratio of graph metrics when compared to GT.

TABLE 3.9. Quantitative analysis of graph metrics[1] of connectomes between estimates and multi-shell HARDI data[2].

Methods	Graph Metrics[3]	HCP	$HCP_{Anomaly}$[4]	Tremor	MND	MS
SS3T CSD [117]	Mean Betweeness Centrality	153.2595	156.2397	154.3462	156.5223	156.2325
	DR[5]	1.58%	5.21%	0.78%	0.69%	2.53%
FOD-Net [104]	Mean Betweeness Centrality	152.4310	154.1476	141.3419	139.2754	154.8807
	DR	1.00%	3.79%	-7.70%	-10.43%	1.67%
FastFOD-Net	Mean Betweeness Centrality	147.0679	151.9794	152.6960	153.1659	150.5734
	DR	-2.53%	2.26%	-0.35%	-1.51%	-1.22%
MSMT CSD [27]	Mean Betweeness Centrality	151.7024	149.2206	153.5751	155.8800	152.6637
	DR	NA	NA	NA	NA	NA
SS3T CSD	Mean Clustering Coefficient	366.9251	358.0020	346.6936	337.7492	330.1410
	DR	-26.63%	-23.28%	-18.07%	-22.26%	-13.34%
FOD-Net	Mean Clustering Coefficient	465.1069	447.4440	368.3558	377.9732	457.1846
	DR	-7.00%	-4.15%	-12.58%	-12.77%	20.21%
FastFOD-Net	Mean Clustering Coefficient	456.2180	439.1776	424.1803	444.6616	392.0182
	DR	-8.78%	-5.87%	0.11%	2.54%	2.76%
MSMT CSD	Mean Clustering Coefficient	500.1260	466.7638	424.4369	435.0083	382.8577
	DR	NA	NA	NA	NA	NA
SS3T CSD	Mean Local Efficiency	459.7633	450.6862	431.9841	424.8624	423.3546
	DR	-27.28%	-23.77%	-18.34%	-22.37%	-14.11%
FOD-Net	Mean Local Efficiency	589.9269	572.6218	474.5110	488.2660	592.3535
	DR	-6.69%	-3.15%	-9.99%	-10.60%	20.35%
FastFOD-Net	Mean Local Efficiency	575.9728	558.6157	539.5586	563.3202	508.4929
	DR	-8.90%	-5.50%	1.87%	3.10%	3.09%
MSMT CSD	Mean Local Efficiency	632.2549	591.2551	530.4742	547.8706	494.5935
	DR	NA	NA	NA	NA	NA
SS3T CSD	Mean Modularity	3.1036	3.0698	2.7955	3.1372	3.2706
	DR	-3.67%	-4.81%	-0.78%	-2.03%	-0.96%
FOD-Net	Mean Modularity	3.1256	3.1111	2.8303	2.8271	2.9195
	DR	-3.11%	-4.56%	0.30%	-11.53%	-12.61%
FastFOD-Net	Mean Modularity	3.2827	3.2611	2.9374	3.3290	3.5195
	DR	2.53%	0.77%	4.11%	3.51%	4.75%
MSMT CSD	Mean Modularity	3.2613	3.3198	2.9292	3.2718	3.4462
	DR	NA	NA	NA	NA	NA
SS3T CSD	Mean Network Degree	70.2893	70.3317	69.6160	70.4396	68.0142
	DR	-0.29%	0.08%	-0.86%	-0.19%	1.49%
FOD-Net	Mean Network Degree	70.0214	69.6952	55.6538	60.6919	66.6677
	DR	-0.67%	-0.84%	-20.76%	-14.01%	-0.53%
FastFOD-Net	Mean Network Degree	70.2429	70.1302	69.4841	70.2997	66.8604
	DR	-0.36%	-0.22%	-1.05%	-0.39%	-0.26%
MSMT CSD	Mean Network Degree	70.4964	70.2825	70.2253	70.5729	67.0800
	DR	-	-	-	-	-

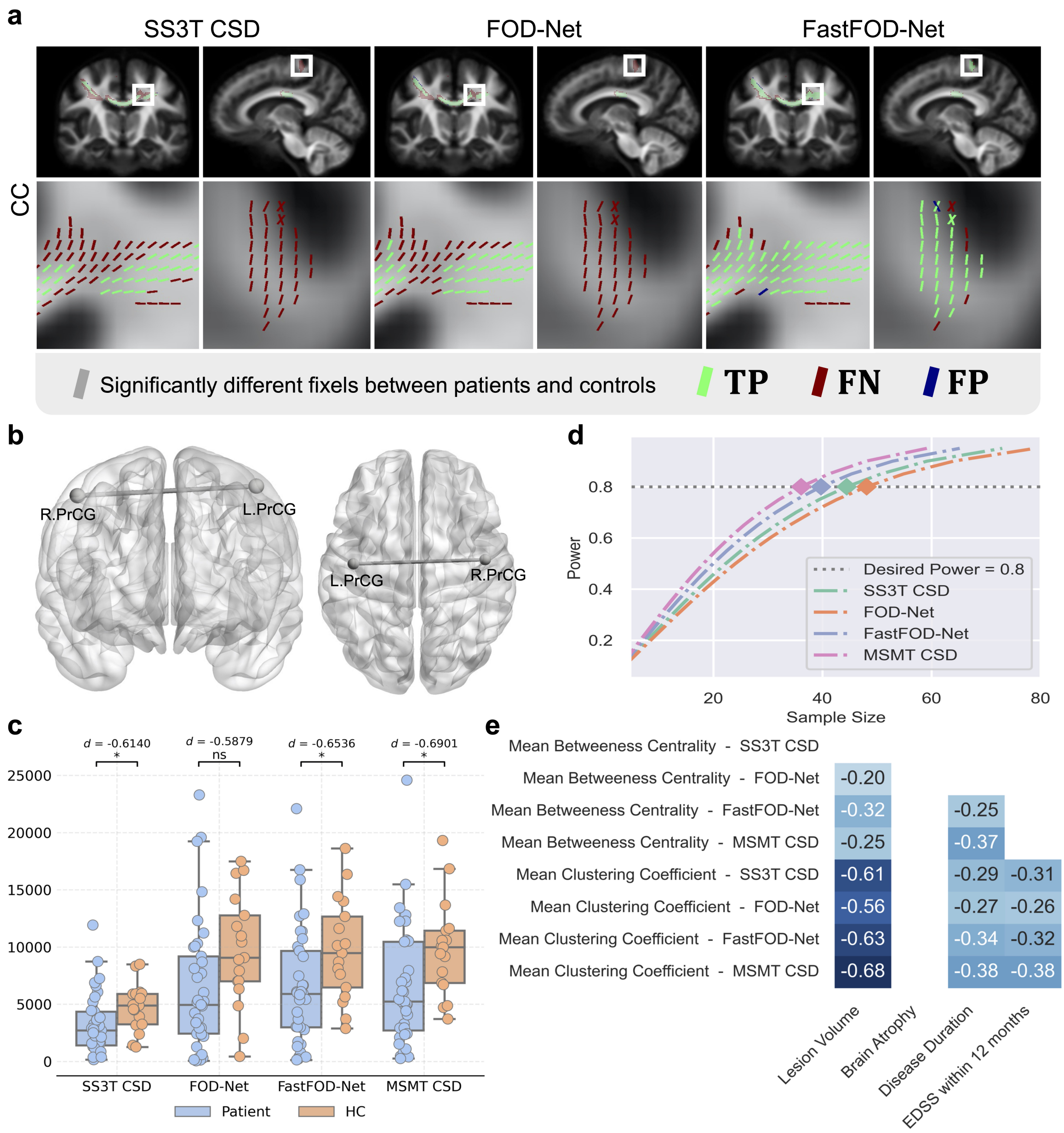
[3]Multi-shell HARDI data is ground truth (GT). [4]$HCP_{Anomaly}$ denotes HCP healthy subjects with structural abnormalities. For each method, [5]DR is the difference ratio of graph metrics when compared to GT.

3.3.4 Clinical Research Application 1: Fixel-Based Analysis

FBA [144, 59] is a powerful tool to perform group comparison analysis at the fixel level (i.e., testing for fixel-wise, rather than voxel-wise, differences). Here, we investigated differences between ALS patients and healthy controls for each FOD method, with MSMT CSD findings as GT. Details on the methodology for FBA are provided in Methods 3.2.5.4. FBA results from SS3T CSD presented a high amount of false negatives $\mathbf{FN}$ (i.e., missed fixel differences) in Figure 3.12a and 3.13. Deep learning was able to recover these missed fixels, and the figure also validated that FastFOD-Net detected a larger number of true significant differences, true positives ($\mathbf{TP}$), compared to FOD-Net and SS3T CSD, particularly along the CC. In contrast, FOD-Net produced more implausiblely significant differences, i.e., false positives $\mathbf{FP}$, on the CST as observed in coronal view. In Table 3.10, FastFOD-Net also shows $\sim 20\%$ enhancement in both $\mathbf{Sensitivity}$ and $\mathbf{F1\ score}$ along the CC compared to SS3T CSD, whereas FOD-Net could not achieve similar improvements. When looking at the CST, both learning-based methods increased $\mathbf{F1}$ by $\sim 10\%$.

3.3.5 Clinical Research Application 2: Pathological Connection Analysis

For pathology-related connections, alterations of their connectivity value between clinical groups was analyzed to further explore the practical impact of learning-based FOD enhancement in a real-world clinical study. More details on analysis methodology are provided in Methods 3.2.5.5. Figure 3.12b,c demonstrate that FastFOD-Net reconstruct FODs that better match the reference MSMT CSD in the MND patient cohort. On the one hand, similar to MSMT CSD, FastFOD-Net revealed significantly ($p < 0.05$) reduced connectivity in ALS (the Left Precentral Cortex [L.PrCG] to the Right Precentral Cortex [R.LrCG]), consistent with previously reported [150], whereas its predecessor FOD-Net failed to detect it. We further evaluated sample size in order to separate patients and HC groups based on the connection from the primary motor cortex when a desired power is achieved at 0.8 across all methods (Figure 3.12d). FOD-Net requires slightly more samples than SS3T CSD, with FastFOD-Net performing better of the two, which might suggest the true pathological connection were recovered and analyzed. For the connection between the Left Postcentral Cortex (L.PoCG)



	Lesion Volume	Brain Atrophy	Disease Duration	EDSS within 12 months
Mean Betweeness Centrality - SS3T CSD				
Mean Betweeness Centrality - FOD-Net	-0.20			
Mean Betweeness Centrality - FastFOD-Net	-0.32		-0.25	
Mean Betweeness Centrality - MSMT CSD	-0.25		-0.37	
Mean Clustering Coefficient - SS3T CSD	-0.61		-0.29	-0.31
Mean Clustering Coefficient - FOD-Net	-0.56		-0.27	-0.26
Mean Clustering Coefficient - FastFOD-Net	-0.63		-0.34	-0.32
Mean Clustering Coefficient - MSMT CSD	-0.68		-0.38	-0.38

FIGURE 3.12. Analysis of Clinical Neuroscience Applications. a, Comparisons of significantly different fixels between the fiber density (FD) of amyotrophic lateral sclerosis (ALS) patients and controls using MSMT CSD as the reference ground truth (GT) method - see more details in Figure 3.13 and Table 3.10. Notably, fixels correctly identified (**TP**) by the the single-shell methods along the corpus callosum (CC), were denoted by green coloring. Conversely, fixels that were not identified despite being significantly different with MSMT CSD (**FN**) were depicted in red, while those erroneously identified as significant (**FP**) were marked in blue. True negative fixels (**TN**) are omitted for clarity, focusing on detected and real differences. b-d, Pathology-relevant connection analysis on ALS patients, i.e, from Left Precentral Cortex (R.PrCG) to the Right Precentral Cortex (L.PrCG). e, Correlation analysis of graph metrics and clinical outcomes on MS cases.

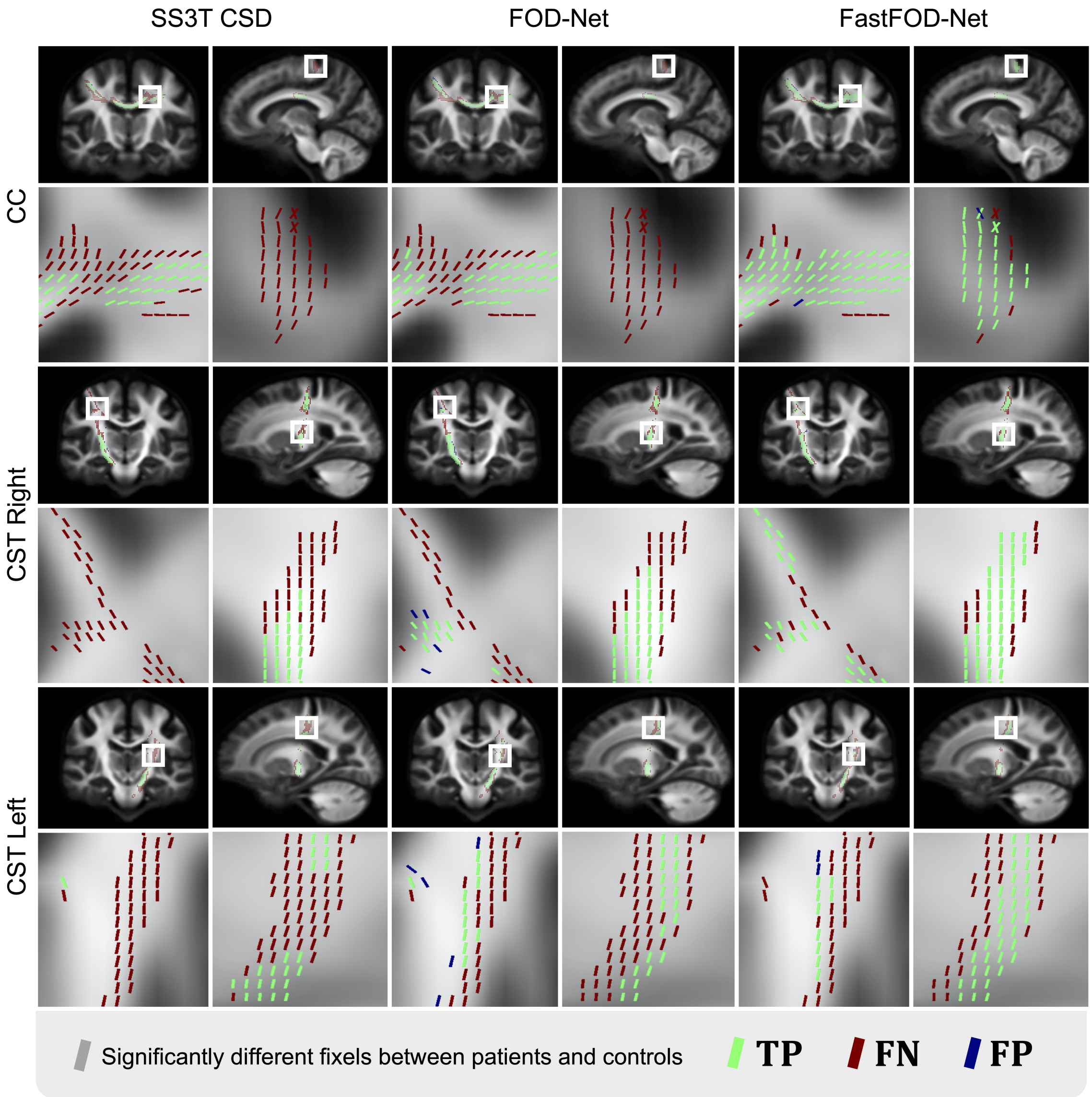


FIGURE 3.13. Comparisons of significantly different fixels between the fiber density (FD) of amyotrophic lateral sclerosis (ALS) patients and controls using MSMT CSD as the reference ground truth method. In this comparison, significant fixels derived from the reference method (MSMT CSD) were juxtaposed with those obtained from alternative methodologies (specifically, SS3T CSD, FOD-Net, and FastFOD-Net). Notably, fixels correctly identified (**TP**) by the comparison methods along various white matter tracts, namely the corpus callosum (CC), right corticospinal tract (CST), and left CST, were denoted by green coloring. Conversely, fixels that were not identified despite being significantly different with MSMT CSD (**FN**) were depicted in red, while those erroneously identified as significant (**FP**) were marked in blue. True negative fixels (**TN**) are omitted for clarity, focusing on detected and real differences.

TABLE 3.10. Quantitative comparisons of significantly different fixels along the corpus callosum (CC), right corticospinal tract (CST), and left CST between healthy subjects and patients, focusing on smoothed fiber density (FD) with negative contrast. In this comparison, significant fixels derived from the reference method (MSMT CSD) were juxtaposed with those obtained from alternative methodologies (specifically, SS3T CSD, FOD-Net, and FastFOD-Net).

	CC[2]			
	[1]**Sensitivity** (↑)	[2]**Specificity** (↑)	[3]**Precision** (↑)	[4]**F1** (↑)
SS3T CSD	0.4662	**0.9708**	0.9385	0.6229
FOD-Net	0.4659	0.9698	0.9390	0.6228
FastFOD-Net	**0.7014**	0.9586	**0.9445**	**0.8050**
	CST Right[2]			
	[1]**Sensitivity** (↑)	[2]**Specificity** (↑)	[3]**Precision** (↑)	[4]**F1** (↑)
SS3T CSD	0.3676	0.9653	0.9195	0.5253
FOD-Net	0.4320	0.9419	0.8939	0.5825
FastFOD-Net	**0.4575**	**0.9742**	**0.9560**	**0.6188**
	CST Left[2]			
	[1]**Sensitivity** (↑)	[2]**Specificity** (↑)	[3]**Precision** (↑)	[4]**F1** (↑)
SS3T CSD	0.3604	**0.9823**	**0.9599**	0.5240
FOD-Net	**0.4565**	0.9276	0.8891	**0.6032**
FastFOD-Net	0.4256	0.9569	0.9292	0.5838

to the Right Postcentral Cortex (R.LoCG) [150], all methods correctly show non-significant differences and aligned with findings by [150]. However, SS3T CSD, yielded the largest absolute effect size of 0.54. In contrast, FastFOD-Net demonstrated superior reliability in neuroscience applications by reducing the chance of finding significant differences (absolute effect size of 0.25) in Figure 3.14.

Furthermore, the FastFOD-Net-enhanced connectomes exhibit significant differences between patients and controls across the top 20 connections, comparable to those observed in single-shell LARDI connectomes, as shown in Figure 3.15 and 3.16. Notably, some of these connections that demonstrate significant differences have also been reported in [150], further supporting the validity of our findings.

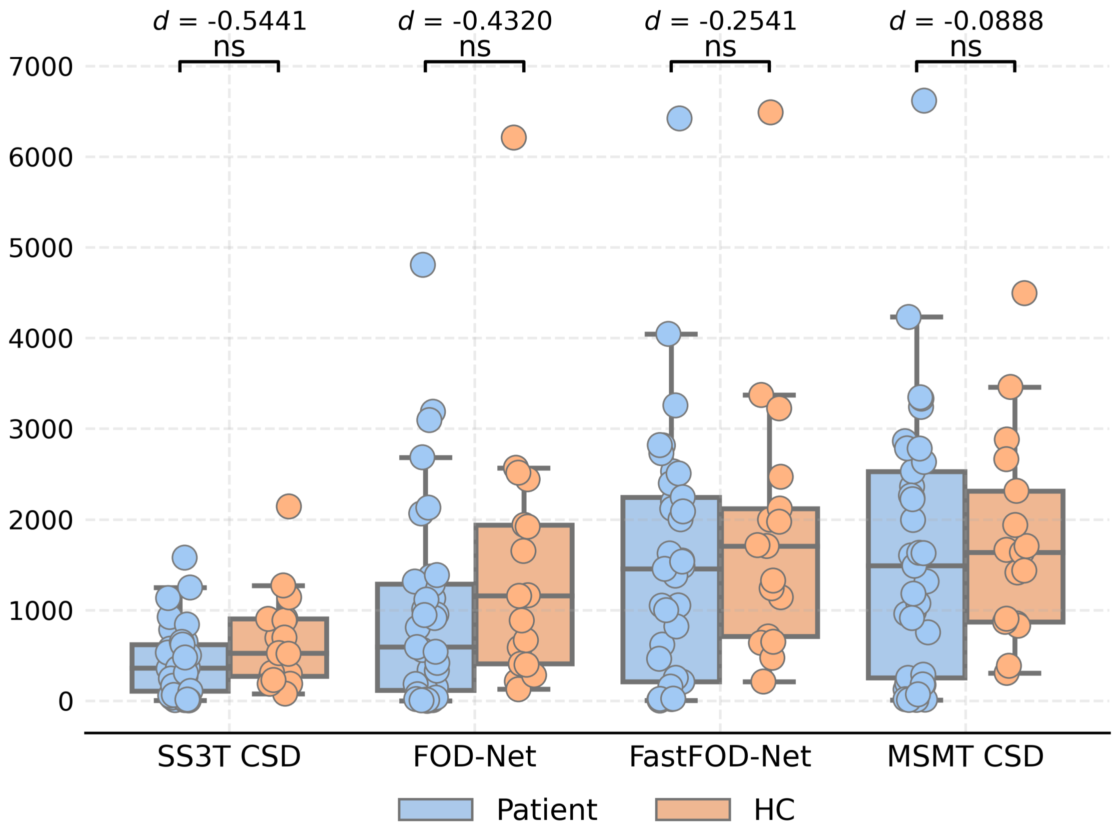


FIGURE 3.14. Box plots depicting connectivity values between amyotrophic lateral sclerosis (ALS) subjects and healthy controls (HCs) for a connection with no clinical significance: from Left Postcentral Cortex (L.PoCG) to Right Postcentral Cortex (R.PoCG) [150]. ALS samples are represented in blue, while HCs are shown in orange. Above each pair of box plots, p-values from independent t-tests indicate the statistical significance of group differences ($p < 0.05$ denoted by '*', while 'ns' indicates non-significance). Cohen's d effect sizes further quantify the magnitude of these differences.

In this study, despite the the limitation in lacking of consistent healthy control subjects in the movement disorder cohort, no significant differences between the disease groups were detected by MSMT CSD or SS3T CSD at both the fixel and connectome levels. Consequently, further biomarker analysis was not performed even though the enhancement also showed no significant differences as expected (avoid false positive connections) - see details of results in Table 3.11.

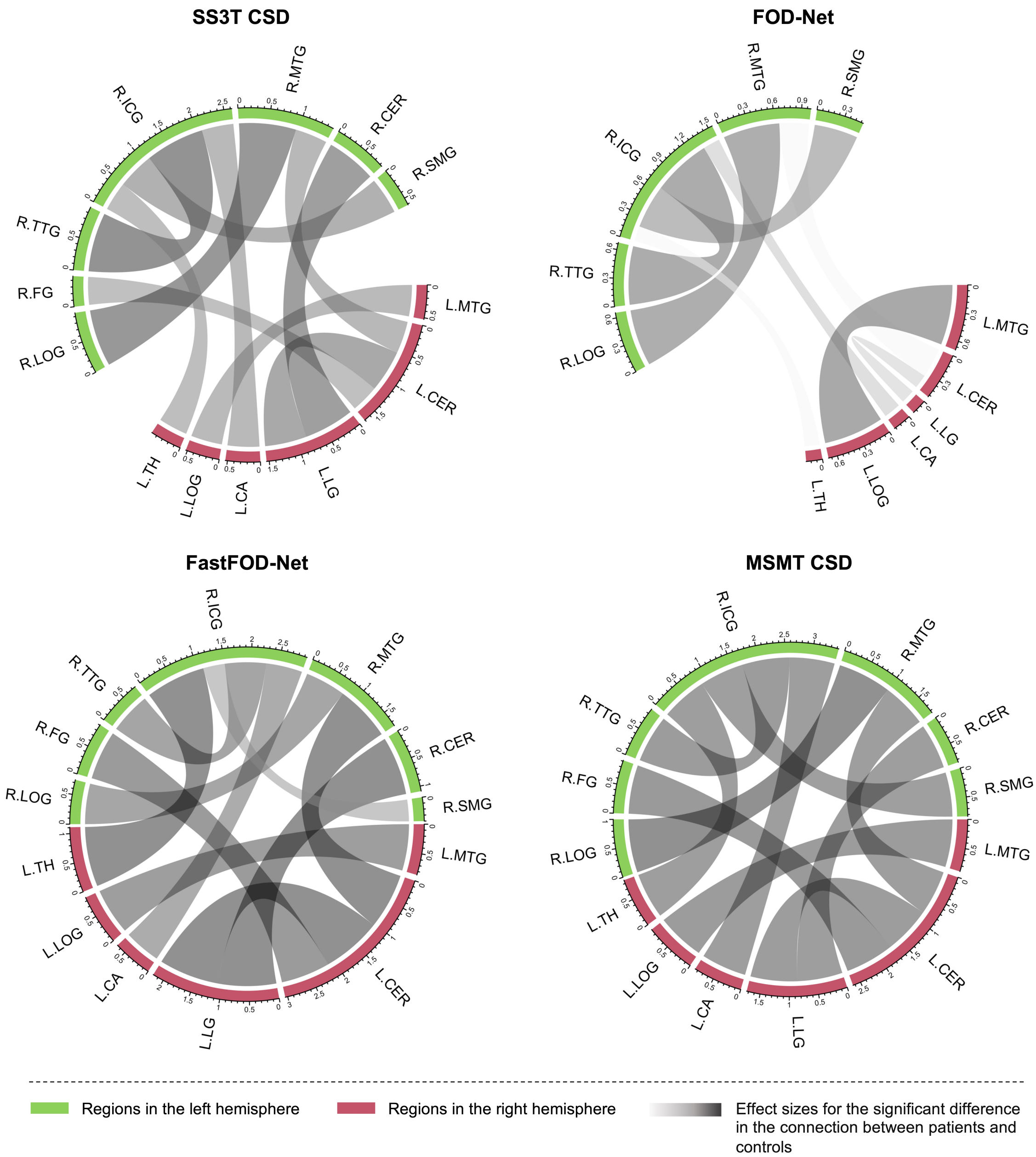


FIGURE 3.15. The chord diagram visualizes the top 10 significantly different connections with absolute value of negative effect sizes between amyotrophic lateral sclerosis (ALS) patients and healthy controls (HCs) based on multi-shell HARDI acquisitions. Red denotes GM regions in the left hemisphere, while green represents nodes in the right hemisphere. Wider bands in light black indicate smaller effect sizes, signifying decreased connectivity in ALS patients, potentially reflecting compensatory mechanisms. Thinner bands in dark black indicate smaller effect sizes.

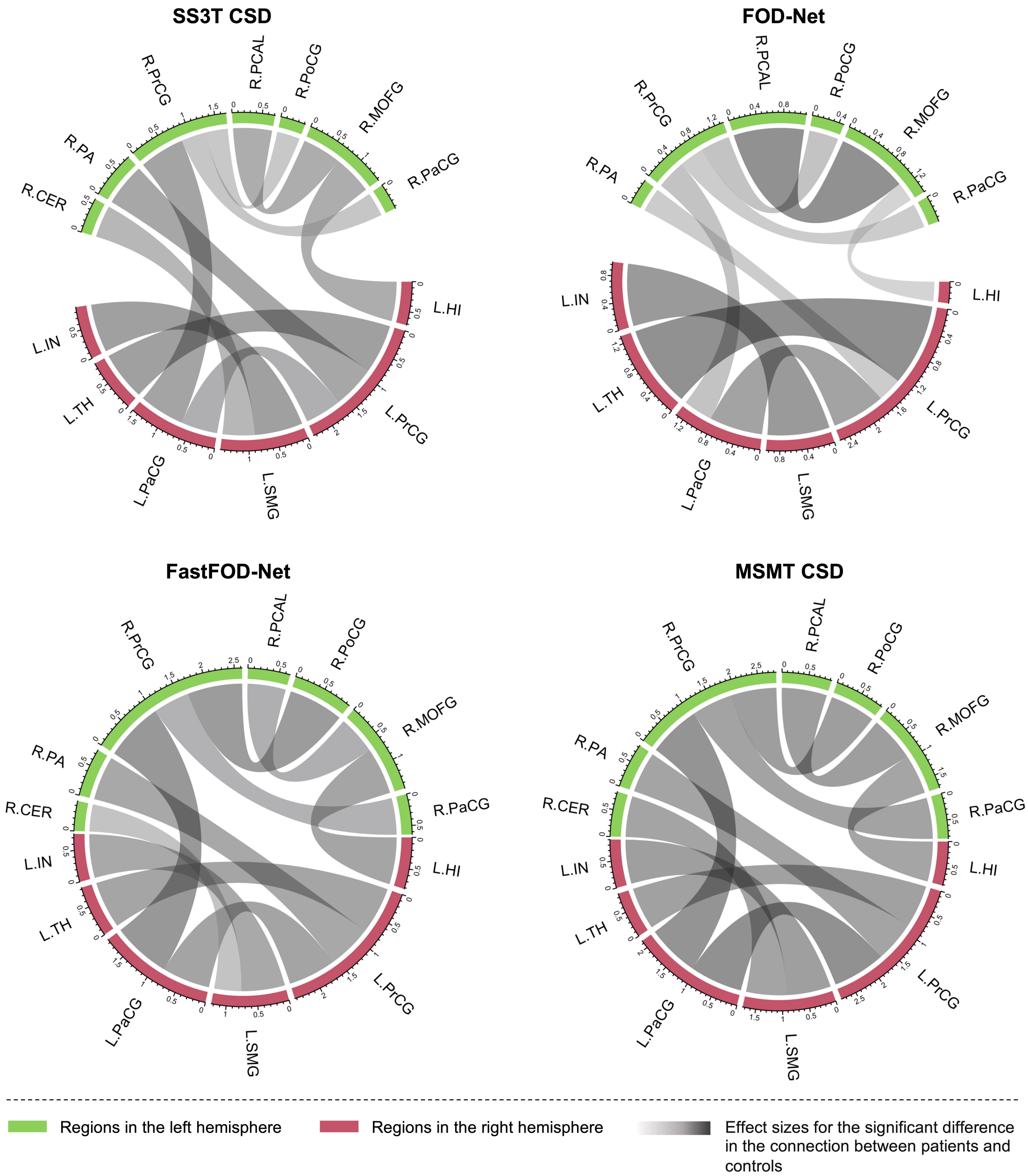


FIGURE 3.16. The chord diagram illustrates the top 10 connections with significant negative effect sizes between amyotrophic lateral sclerosis (ALS) patients and healthy controls (HCs), based on multi-shell HARDI acquisitions. In this diagram, red nodes represent GM regions in the left hemisphere, while green nodes represent those in the right hemisphere. The width of each band corresponds to the magnitude of the negative effect size: wider bands indicate larger negative effect sizes, signifying more substantial decreases in connectivity in ALS patients. This reduction in connectivity may reflect underlying disease mechanisms. Thinner bands indicate smaller negative effect sizes, suggesting less pronounced connectivity reductions.

TABLE 3.11. ANOVA analysis between tremor disorders, i.e., Parkinson's Disease (PD), Essential Tremor (ET), and Dystonic Tremor (DT), of a pathology-relevant connection (i.e., the connection between thalami) across different methods.

Method	Factor	DF[1]	SS[2]	Mean Square[3]	p
SS3T CSD	Between (Group)	2	23.8401	11.9201	0.5904
	Within (Residual)	36	802.4890	22.2913	
FOD-Net	Between	2	0.2355	0.1178	0.8488
	Within	36	25.7489	0.7152	
FastFOD-Net	Between	2	7.7358	3.8679	0.6598
	Within	36	330.9880	9.1941	
MSMT CSD	Between	2	9.0563	4.5282	0.5195
	Within	36	244.4170	6.7894	

[1]DF is the Degrees of Freedom associated with the factor. [2]Sum of Squares (SS) and [3]mean square represents the total variability and the average variation due to the factor respectively.

3.3.6 Clinical Research Application 3: Correlation Analysis

Last but not the least, the capability for recovering white matter tract from focal pathology has been raised as one of the challengs in disease such as multiple sclerosis, which may leads to false connections and biased outcome due the lesion load and distribution. In addition to group analysis, we assessed the correlation between graph metrics and clinical outcomes on MS cases shown in Figure 3.12e, including lesion volume, brain atrophy, disease duration, and Expanded Disability Status Scale (EDSS) within 12 months. Our findings suggest that SS3T CSD did not show significant correlations with lesion volume and disease duration in mean between centrality and exhibited lower correlation values in mean clustering coefficient. In contrast, FastFOD-Net demonstrated significantly higher correlations with lesion volume and disease duration in these two metrics, which are identified as significantly different between multiple sclerosis patients and controls [154]. Several graph metrics, e.g., mean strength, mean transitivity, mean global efficiency, showed statistically significant relationships with EDSS [155, 156, 157]. Consist with prior findings, FastFOD-Net captured significant relationships with EDSS in these graph metrics similarly to MSMT CSD in Figure 3.17. Assortativity, however, did not show a significant correlation with EDSS [155], and FastFOD-Net, like MSMT CSD, avoided false positive correlations in this metric.

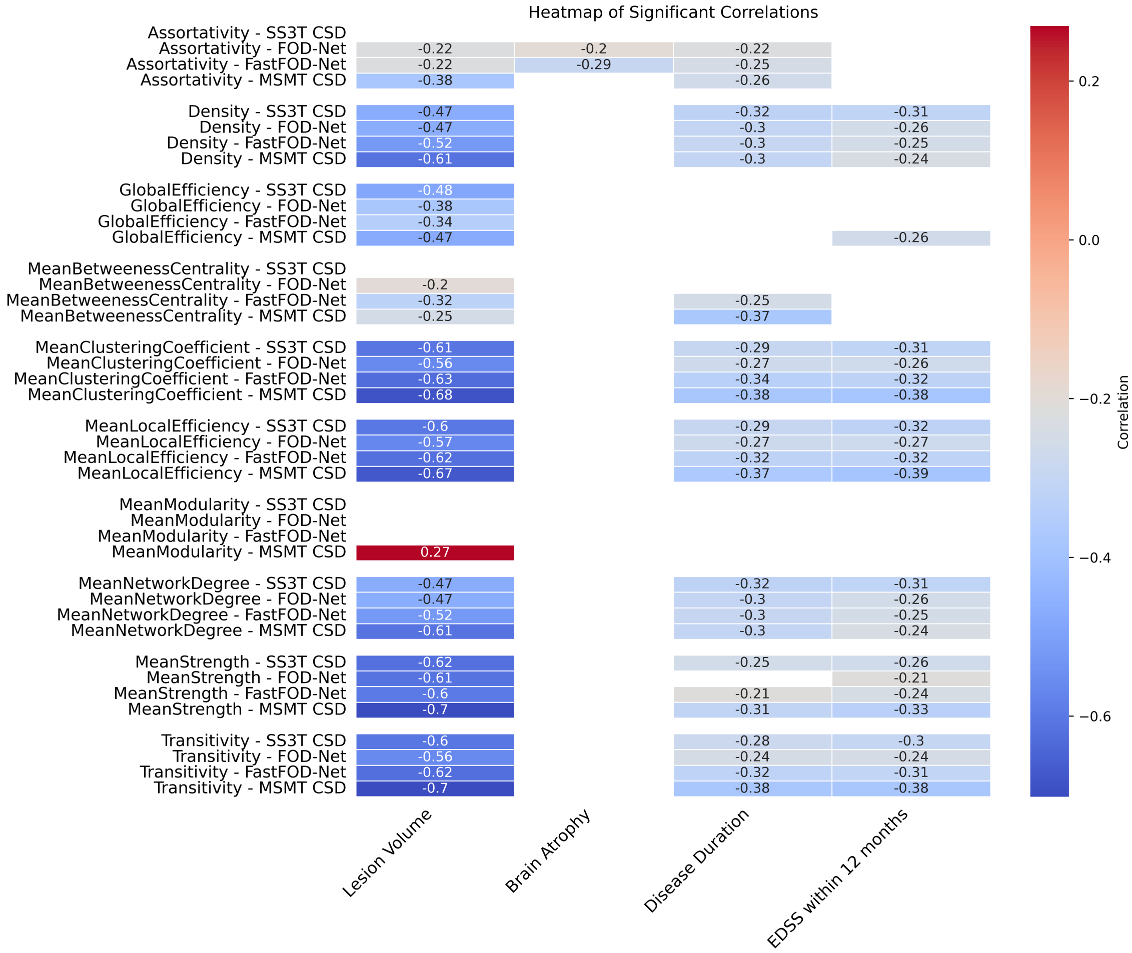


FIGURE 3.17. Heatmap of significant correlations ($p < 0.05$) between graph metrics from different FOD estimation methods and clinical outcomes in multiple sclerosis patients. Correlations are shown for total lesion volume, whole brain atrophy, disease duration, and Expanded Disability Status Scale (EDSS) within 12 months.

3.4 Discussions

While there have been major research advances in methods to exploit the wealth of information available from diffusion MRI, their reliability based on clinical quality imaging data remains a major practical challenge. Recent studies in DWI and FOD enhancement [99, 100, 105, 104, 106, 158, 159, 160] have introduced novel learning-based approaches to address this issue, yet have neglected the, arguably, essential evaluation of how enhanced image quality translates into real-world clinical advantages. This is one of the main reasons for the field of deep learning, despite endeavors to achieve greater accuracy and employ larger architectures, has seen the translation of these advancements into real-world clinical benefits remaining stubbornly rare. To address this gap, we have developed a clinical-trust evaluation pipeline to assess the real-world practical impact of learning-based diffusion MRI enhancement. This pipeline adopts commonly used FOD-level metrics and refines fixel-level measures by considering missing and extra fiber elements within a voxel, to more accurately measure the differences in complex fiber configurations between different methods. Furthermore, a more comprehensive connectome assessment was incorporated to evaluate edge disparities and differences of topological patterns in brain connectivity. This extensive battery of tests, applied to a range of neurological conditions, provide the missing piece and evidence needed for a widespread adoption of deep learning solutions for FOD enhancement in clinical applications.

Since most prior DWI enhancement methods considered a varying number of gradient directions, their computational cost was relatively high (more than one week with multiple GPUs) due to the use of recurrent CNNs (RCNNs) and the replacement of convolutions with sub-sampled MLPs [99, 100]. Using FODs as an intermediate model is a better solution, as it is more closely related to downstream tasks of interest (e.g., fiber-tracking, FBA, connectome, etc). Earlier studies on FOD enhancement [105, 104, 106, 124, 160] utilized similar deep learning methods, but mostly focused on a limited number of healthy subjects or a limited suite of metrics. From these methods, we chose FOD-Net [104] as a representative method that has been reported to have good performance. FOD-Net processed coefficients of FOD data voxel-by-voxel for each spherical harmonic order respectively. Despite its benefits, FOD-Net is a time-consuming method that overlooked the inherent attributes of such 3D imaging data and inter-order coefficient relationships. In contrast, by avoiding voxel-wise learning and eliminating high computational architectural components for independent order learning, the proposed FastFOD-Net used a fully convolutional architecture to predict the entire input patch at once with improved performance and faster inference (within a few minutes). Thus, FastFOD-Net serves as a simplified end-to-end solution for fast FOD enhancement, further facilitating its large-scale clinical evaluation and adoption.

While FOD-Net 2.0 extended the network to encode coefficients from multiple spherical harmonic orders, which is conceptually sound, we observed that this architectural change brought minimal practical improvements. One possible explanation is that a complete fiber bundle configuration may rely on the joint pattern of all 45 coefficients, and separating them by order might have represented the directional information in an overly fragmented manner. This insight motivated the design of FastFOD-Net, which adopts a more integrated and efficient architecture.

Recent benchmarks [161] further underscore that, despite the conceptual appeal of novel architectures like Transformers [162] and Mamba [163], their practical benefits over well-established models like U-Net are not always evident. Their findings suggest that a properly configured U-Net architecture, particularly when utilizing frameworks like nnU-Net and scaling models to modern hardware resources, can still achieve state-of-the-art results in 3D medical image analysis. These results indicate an ongoing innovation bias towards novel architectures in the field and underscore the need for more stringent validation standards in the quest for scientific progress.

At the FOD, fixel, and connectome levels, FastFOD-Net was found to provide superior FOD enhancement from clinical single-shell acquisitions across a range of different neurological conditions, including motor neuron diseases, tremor disorders, and multiple sclerosis. With a single GPU, the training and testing time was greatly reduced from 4-8 weeks and 23 hours, respectively (for FOD-Net), to within 1 day and 2-3 minutes, respectively (for FastFOD-Net). When compared to the state-of-the-art SS3T CSD, FOD-Net was superior for healthy subjects; however, it led to a higher number of significantly different edges (relative to the GT values) than SS3T CSD in both Tremor and MND cases (Table 3.7). This limitation was less pronounced in FOD and fixel metrics (Figure 3.4). In Figure 3.7, SS3T CSD and FOD-Net exhibited fewer edges with large disparity, suggesting that they may generate implausible connections that, despite low disparity, deviate significantly from the GT. Many of these connections were absent or near zero in the GT but appeared with nontrivial values, indicating a tendency toward false positives. In contrast, FastFOD-Net nearly halved the disparity analysis metrics on patient cases compared to SS3T CSD for disease states, indicating superior robustness in downstream tasks. Similar trends were observed in topological analyses of the connectomes. Overall, the combined analysis of these metrics supports the role of FastFOD-Net for clinical applications to recover information missing from LARDI acquisitions.

ALS is a rare neurodegenerative disorder clinically characterized by a progressive breakdown of essential everyday motor function, ultimately leading to a loss of life [164]. Electrophysiology studies have consistently identified loss of cortical motor inhibition, i.e., cortical hyperexcitability, as a prominent early marker of disease [165]. On neuroimaging, disease pathology is characterized by clear structural, functional, and metabolic changes [166]. Diffusivity abnormalities in the CC and CST hold the greatest discriminative accuracy and significantly correlates with severity of functional motor impairment in patients [167, 168]. Notably, fine-grained changes within the CST and CC have been shown to reflect clinically meaningful phenotypic variation in asymmetry of disease presentation [169, 170]. Our results (Figure 3.12, Figure 3.13 and Table 3.10) demonstrated that FastFOD-Net could effectively reproduce the detection of most significantly different fixels (over 80%) along the CC tract when compared with GT, which has the potential of aiding in understanding neural disorder pathogenicity. Both deep learning methods had comparable results along the CST, with still pronounced but lower results compared to the original LARDI acquisition (10% increase in true positives). Consequently, FastFOD-Net therefore provides an avenue to investigate the pathophysiology of ALS using routinely collected LARDI diffusion data.

Analysis of the pathology-relevant connections from P.rCG to R.rCG found that connectomes derived from MSMT CSD (GT) showed higher variability (larger standard deviation) v.s. SS3T CSD ones (Figure 3.12), potentially reflecting more faithful recovery of complex white matter fibre configurations along L.PrCG-R.PrCG connectome with the GT acquisition. Since FODs generated with FOD-Net tend to have noisier coefficients with a higher number of lobes, FOD-Net demonstrated the largest variability with the widest and most overlapping distributions. Conversely, the GT distribution of this pathological connection better approximated by FastFOD-Net, reducing the likelihood of false positive disease-related differences in clinical-quality diffusion MRI data identified in previous clinical studies [150, 149]. In terms of the control connection (Figure 3.14), FastFOD-Net lowered the differences between groups, while

SS3T CSD and FOD-Net were much less efficient in this regard. This suggests that our new method reduces the possibility of false positive disease-related differences in clinical-quality diffusion MRI data. Our results demonstrate the potential of FastFOD-Net to identify deficits in the primary motor cortex among ALS patients, consistent with biomarkers identified in previous clinical studies [150, 149]. The ability of FastFOD-Net to achieve sufficient statistical power with a smaller sample size (Figure 3.12) makes it a valuable practical tool with reduced recruitment burdens for investigating neurological disorders based on clinical quality data, and thus optimizing research efforts.

Although FastFOD-Net consistently improved FOD quality across the tremor movement disorder cohorts, small sample sizes (ET: $n = 19$, PD: $n = 7$, DT: $n = 13$) and the absence of healthy controls with a consistent MRI protocol limits the statistical power of group comparisons and the interpretation of disease-specific features. Previou studies have established that each of these disorders exhibit structural connectome changes relative to healthy individuals [171, 172, 173]. Bergamino et al. [174] used ET ($n = 17$) and PD ($n = 68$); however, direct comparisons among ET, PD, and DT are scarce in the literature, and unique diffusion signatures that distinguish these movement disorders are yet to be defined. Within this context, our analysis detected no significant differences in connectivity graph metrics between the ET, PD, and DT groups using FastFOD-Net, consistent with the GT MSMT CSD results. FastFOD-Net may potentially preserve true biological networks without introducing spurious group differences, though the limited sample size warrants cautious interpretation. Consistency between FastFOD-Net and MSMT CSD results increases confidence that FastFOD-Net's enhancements are faithful and do not distort group comparisons. Future studies with larger cohorts and matched healthy controls will be valuable in confirming subtle between-group differences (or the lack thereof) in fixel-based and graph-based metrics.

In MS, FastFOD-Net's utility for clinical applications was confirmed in a large patient cohort ($n = 111$). MS is an inflammatory and neurodegenerative disease characterized by multifocal white matter lesions and diffuse microstructural damage throughout the brain. Conventional MRI (T2-FLAIR, T1) is incorporated into the diagnostic criteria for MS, facilitating the demonstration of lesions in the CNS that are disseminated in space and time [175]. However, these sequences are less sensitive to subtler damage in the brain; and do not reveal changes associated with disease progression independent of relapse biology (i.e., new or enlarging lesions) [176]. Additionally, a particular challenge for diffusion-based tractography in MS is the presence of focal lesions that disrupt fiber organization: demyelination and axonal loss lead to low anisotropy, which may impact fiber orientation reconstructions [177]. The heterogeneous distribution and variable size of MS lesions make it difficult to obtain consistent fiber orientation information across the brain, posing a stringent test for FOD enhancement methods. FastFOD-Net performed robustly in our MS cohort, producing high-quality FODs both inside and outside lesions (Fig. 3.4 and Supplementary Fig. 3). This suggests that the deep learning model generalized well to focal tissue damage. Additionally, FastFOD-Net's FOD enhancement led to subsequent connectome metrics that were biologically meaningful and consistent with those obtained using the gold-standard MSMT CSD approach (Supplementary Fig. 10). For instance, many global network measures (e.g., efficiency, strength, network degree, etc.) computed from FastFOD-Net data showed significant correlations with conventional markers of disease burden, such as total lesion volume, as well as with EDSS (but not whole-brain atrophy). Our findings recapitulate the results of prior studies linking diffusion MRI metrics to MS disease severity. For example, global structural network disruption (e.g., lower network efficiency or density) has been associated with increased disability in people with MS [178]. These correlations were highly consistent with those derived from MSMT CSD-based connectomes, suggesting that little information was lost when using FastFOD-Net with SS3T; and that FastFOD-Net retains sensitivity to clinically relevant microstructural changes. This robustness across a large MS sample highlights the potential of FastFOD-Net longitudinal studies and therapeutic trials in MS, where single-shell dMRI could be efficiently acquired while still enabling sophisticated fixel-based analyses or connectomics to monitor disease progression.

Our study has several limitations. First, analysis of the movement disorder cohorts was limited by small sample sizes and the lack of a relevant control group, reducing our ability to detect subtle group differences. Future studies should include larger cohorts of ET, PD, and DT patients, as well as healthy controls, to validate whether FastFOD-Net can reveal network alterations specific to each condition. Additionally, extending FastFOD-Net to other clinical populations and scanner settings would be valuable. MS lesions present a challenging task in white matter for FOD enhancement, but other neurological diseases (such as stroke or traumatic brain injury) exhibit different patterns of white matter damage that warrant evaluation with FOD enhancement. Finally, ongoing improvements in network architecture (e.g., incorporating transformer-based models [107], Mamba-based models [163], vector-quantized variational autoencoders [VQ-VAE] [179], and diffusion models [88] for FOD super-resolution) may yield even better performance and introduce variability in generations. However, recent benchmarks [161, 137, 180, 181] caution that architectural novelty, on its own, does not guarantee meaningful progress; equally crucial is that large-scale, rigorously validated clinical benefit can also be demonstrated. We anticipate that continued refinement of FOD enhancement methods, combined with thorough clinical validation, will solidify their role in precision neurology—enabling robust connectome and fixel analyses from standard diffusion acquisitions to inform diagnosis, prognosis, and treatment planning.

3.5 Summary

In conclusion, FastFOD-Net is a fast deep learning method to enhance single-shell LARDI acquisitions within a few minutes, and our evaluation pipeline in a range of neurological conditions provides an in-depth exploration of the impact promised by deep learning in clinical applications. The results demonstrated FastFOD-Net's ability to recover fiber information missing from single shell clinical quality data, and its superiority for robust downstream brain connectivity analysis. FastFOD-Net could also optimize clinical research efforts due to its efficiency in using single-shell LARDI data (lower acquisition time) and its reduced sample size, while still recovering clinical significant group differences that are the most closely aligned to multi-shell HARDI quality data without introducing false positives. In summary, our study provides a missing piece to increase the widespread adoption and construct clinical trust of deep learning enhancement moving forward. The code repository has been released to further help with this big step.

CHAPTER 4

Explainability: Estimating Uncertainty on Deep Learning-Driven FOD Enhancement

---

Diffusion imaging techniques are commonly used to study water diffusivity and white matter tracts in the brain. High angular resolution acquisitions yield high-quality FODs, but their longer acquisition time makes them impractical in the most clinical settings. Single-shell low angular resolution protocols are a more feasible alternative, though less informative to reconstruct FODs for complex white matter tracts. To address that gap, recent studies have focused on enhancing FODs from single-shell low angular resolution data.

In the last two chapters, we addressed the challenges related to the inefficiency of method design, the lack of clinical evaluation, and the acquisition protocol requirements. Nonetheless, uncertainty remains a critical concern due to the complex nature of white matter tracts and the impact of FOD enhancement on downstream tasks such as tractography and human connectome mapping. Current uncertainty estimation methods show promise in classification and segmentation (classification at the pixel or voxel level) tasks. However, medical image analysis often involves regression tasks, such as super resolution and denoising, where uncertainty estimation is equally important but in general less studied. This study formulates FOD enhancement from single-shell low angular resolution data as a multivariate regression problem and estimates the uncertainty of the predictions using CNNs from a statistical basis. Our results show that the proposed FOD uncertainty correlates with FOD-level errors, providing insights into prediction reliability. This is the first work to uncover the link between improving the coefficients of a model for water diffusivity (FOD) and its true effect on fiber estimation, thereby advancing the reliability of deep learning-driven enhancement for diffusion imaging. This chapter contributes to enhancing the explainability of deep learning methods in diffusion MRI, tackling **Challenge 4: Explainability of Deep Learning** (Section 1.4.4) and further bridging the gap between deep learning-driven enhancement and its practical applicability in neuroimaging.

4.1 Introduction

dMRI is a non-invasive advanced imaging technique to explore the anisotropy of water diffusivity of brain WM tracts [110, 111]. The quality of dMRI analysis is highly dependent on the angular resolution (number of gradient directions used in acquisitions) and the models used to condense the information. When combined with CSD [27, 42, 54], multi-shell HARDI can be used to generate high quality FODs. FOD is a model employed to represent the probability of white matter fiber distribution in the brain, offering detailed insights into the orientation and density of fiber bundle elements ('fixels') within each voxel. This modelling approach accounts for the complex structure of WM when conducting tractography, which can subsequently be used to probe structural brain connectivity [112]. However, multi-shell acquisitions cannot be routinely acquired in clinical scenarios due to their extended acquisition time. Therefore, single-shell LARDI is commonly acquired at the expense of losing crucial information about crossing fiber bundles and potentially estimating misleading tracts [113]. These challenges underscore the need for enhancement methods that can produce reliable FOD coefficients from data acquired with single-shell LARDI protocols.

Deep learning methods have demonstrated robust performance in estimating basic diffusion models, such as DTI, from LARDI [182, 94]. However, the estimation of FOD, a more advanced and complex diffusion model, presents greater challenges [95]. Recent works on FOD enhancement [109, 115, 106, 159] have focused on predicting FODs from patches of single-shell LARDI acquisitions represented by their estimated coefficients of a spherical harmonic expansion [54], achieving enhancements comparable to those from multi-shell HARDI acquisitions. Despite these advancements, the inherent uncertainty of deep learning models due to the acquisition noise, the variability of the FOD data, and the model optimization process remains a critical concern and has not yet been explored. Consequently, analyzing the uncertainty of these enhanced variable fiber structures is of high interest.

Despite the prevalence of CNNs in the field of medical imaging analysis, earning the trust of clinicians and patients [183] is still an open challenge due to the black box nature of such techniques. Therefore, assessing the uncertainty of the predictions is of high interest in the field of explainable artificial intelligence. Conventional Bayesian neural networks [184] offer a mathematically-grounded strategy to directly represent weights as probability distributions. However, such methods are inviable for large networks with millions of parameters due to the computational complexity. A solution to this problem is the Monte-Carlo Dropout (MCDropout) technique that has been usually interpreted as a Bayesian approximation of a classical Gaussian probabilistic model [185]. Nonetheless, this technique has the disadvantage of requiring dropout before every layer with learnable weights and multiple forward passes during inference. Nevertheless, carefully tuned models with MCDropout are often preferable and are one of the most popular methods in diverse case studies [186]. Other studies in uncertainty in deep learning have proved that Deep Ensembles (DE) [187, 188] of multiple models are another trustworthy technique [189, 190] at the cost of a large memory consumption and training time. More recently, a scalable Layer Ensembles (LE) method [191] was proposed to estimate the uncertainty with one single forward pass and a relatively small number of additional parameters, while still following the principle of DE. These methods have shown promise in a variety of classification and segmentation (classification at the pixel or voxel level) tasks. However, medical image analysis often involves regression tasks, such as super resolution [192] and denoising [193], where uncertainty estimation is equally important but less studied.

Enhancing FODs from clinical scans of low angular resolution is a complex task that can be formulated as a multivariate regression problem. By analyzing uncertainty in this context, we aim to improve the explainability of enhancement models, gain insights into their decision-making processes, and help clinical researchers understand the confidence levels of the predicted FODs. Thus, analyzing uncertainty is essential for earning the trust of clinicians, given the complex nature of brain tracts (varying number of fiber bundle elements per voxel). To our knowledge, this is the first work using uncertainty techniques to uncover the link between improving the coefficients of a model representing the structural organisation of white matter tracts (FOD) and its true effect on downstream tasks like fiber bundle element estimation. This work has potential implications for advancing the reliability and trustworthiness of deep learning-driven enhancements in dMRI data.

4.2 Methods

In this section, we first present a deep learning method to enhance the single-shell LARDI FOD data. Subsequently, we estimate uncertainty on deep learning-driven FOD enhancement by formatting it as a multivariate regression problem and simplifying the associated differential entropy.

4.2.1 FOD Enhancement

To enhance the single-shell LARDI FODs, our study presents an end-to-end CNN based framework inspired by [104, 115, 159] in Figure 4.1. Each input FOD volume has a resolution of $Height \times Width \times Depth \times C$ regardless of the angular resolution, where C denotes the number of spherical harmonic coefficients (45 when setting $lmax = 8$). Specifically, we used a U-Net backbone [136] to predict the HARDI FODs on a patch-wise basis to address the computational costs of the previous voxel-wise estimation method. Our model consists of 4 encoding layers for feature extraction of coefficients from LARDI data, and 4 decoding layers for spatial resolution upsampling to the ultimate FOD dimension. Skip connections were introduced to allow communication of coefficient learning between the encoding and decoding phases.

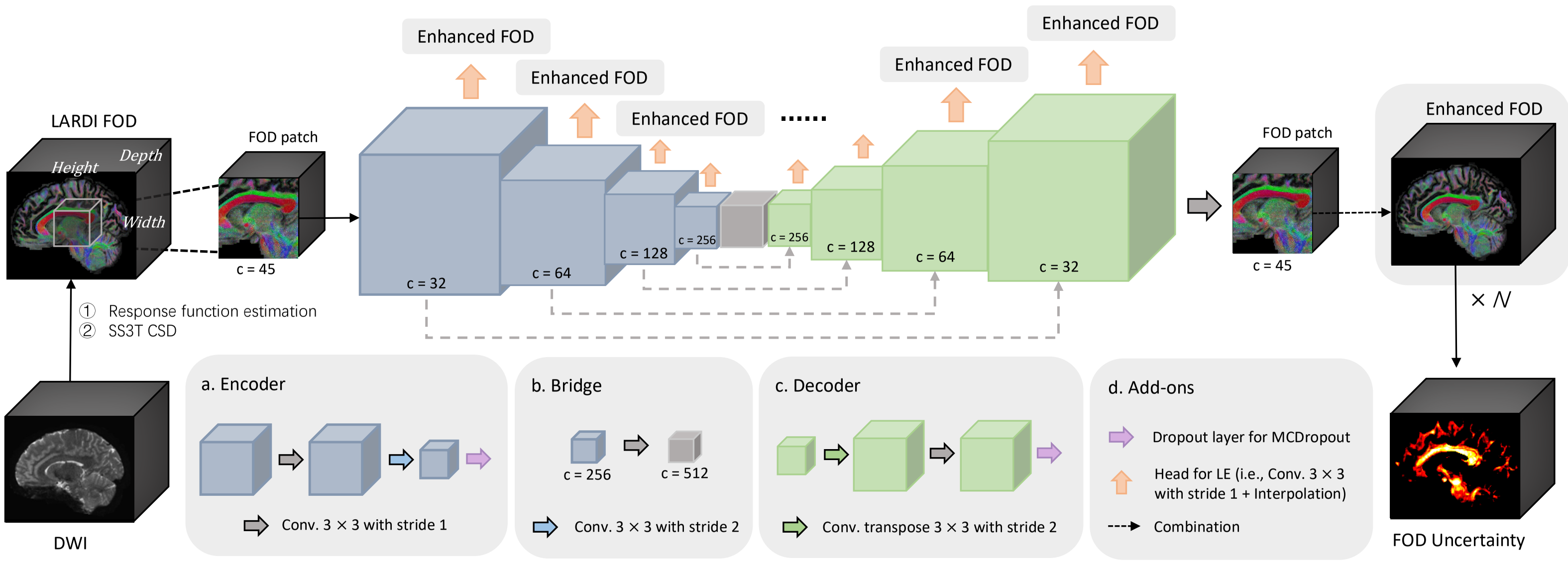


FIGURE 4.1. Overview of the framework for FOD uncertainty estimation. Purple arrows demonstrate the dropout layer added after each encoder and decoder module for deep ensemble (DE). Orange arrows denote the heads added at each layer of the backbone for layer ensemble (LE). N denotes the number of samples using different sampling methods: number of models N_{de} for DE, number of layers N_{le} for LE, and number of forward passes N_{mcd} for Monte-Carlo Dropout (MCDropout).

4.2.2 Multivariate Regression Uncertainty

Shannon's entropy [194] is one of the most common techniques to represent uncertainty in classification approaches (both pixel-wise and image-wise). Its value represents the average level of 'uncertainty' inherent to the variable's possible outcomes with a uniform distribution being the extreme case of higher uncertainty (all outcomes are equally likely). Given a discrete random variable x with a probability density function $p(x)$, Shannon's entropy can be defined as:

$$\mathbf{H}(x) = -\sum p(x) \log_b p(x), \tag{4.1}$$

where b denotes the logarithm's base for the units of the entropy (i.e., $log_e = ln$ represents natural units of information termed 'nats').

In the case of a continuous random variable, Shannon [194] extended his definition of entropy termed "differential entropy" by changing the summation to an integral over the domain of the whole distribution. For a random D-dimensional vector of all the possible predictions of a network $x = (x_1, x_2, ..., x_D)$ following a multivariate Gaussian distribution ($x \sim \mathcal{N}_D(\mu,\, \mathbf{\Sigma})$), we can express the differential entropy as:

$$\mathbf{H}(x) = -\int p(x) \log p(x)\, dx = -\mathbb{E}[\log p_{\mathcal{N}_D}(x)]. \tag{4.2}$$

Using the Gaussian probability density function, this expands to:

$$-\mathbb{E}\left[\log\left(\frac{1}{\sqrt{(2\pi)^D|\mathbf{\Sigma}|}}\exp\left(-\frac{1}{2}(x-\mu)^{\mathbf{T}}\mathbf{\Sigma}^{-1}(x-\mu)\right)\right)\right]. \tag{4.3}$$

We can simplify that expectation using logarithm rules as follows:

$$\mathbf{H}(x) = \frac{D}{2}\log(2\pi) + \frac{1}{2}\log|\mathbf{\Sigma}| + \frac{1}{2}\mathbb{E}\left[(x-\mu)^{\mathbf{T}}\mathbf{\Sigma}^{-1}(x-\mu)\right]. \tag{4.4}$$

Knowing that $(x-\mu)^{\mathbf{T}}\Sigma^{-1}(x-\mu) \sim \chi^2(D)$, the last term simplifies to:

$$\mathbb{E}\left[(x-\mu)^{\mathbf{T}}\mathbf{\Sigma}^{-1}(x-\mu)\right] = \mathbb{E}[\chi^2(D)] = D. \tag{4.5}$$

Finally, we can substitute that term into Equation 4.4 to get:

$$\begin{aligned}\mathbf{H}(x) &= \frac{D}{2}\log(2\pi) + \frac{1}{2}\log|\mathbf{\Sigma}| + \frac{1}{2}D \\ &= \frac{D}{2}(1+\log(2\pi)) + \frac{1}{2}\log|\mathbf{\Sigma}|.\end{aligned} \tag{4.6}$$

To measure the uncertainty at the FOD level, multiple predictions are obtained from three common sampling methods as outlined in Figure 4.1: (1) DE [187], which trains $N_{de} > 1$ models with different random seeds and aggregates them to simulate the random distribution of the network weights; (2) MCDropout [185], which samples multiple predictions with $N_{mcd} > 1$ forward passes through the network with the activation of stochastic dropout layers during inference; and (3) LE [191], which attaches a head after each network layer yielding a composition of N_{le} sub-networks at different depths of the network to sample multiple predictions with a single forward pass.

Each FOD estimate is represented by 45 coefficient values that can be decomposed into different orders of spherical harmonics, with higher orders having a greater number of coefficients. Given a set of N random predictions for the same subject, assuming that the predictions follow a multivariate Gaussian distribution $\mathcal{N}_D$ with $D = 45$ dimensions, we can define a covariance matrix $\mathbf{\Sigma}$ that represents the variability of the predictions for each voxel x:

$$\mathbf{\Sigma} = \begin{bmatrix}\sigma_{x_1} & & & 0 \\ & \sigma_{x_2} & & \\ & & \ddots & \\ 0 & & & \sigma_{x_D}\end{bmatrix}\begin{bmatrix}1 & \rho_{x_1,x_2} & \cdots & \rho_{x_1,x_D} \\ \rho_{x_2,x_1} & 1 & \cdots & \rho_{x_2,x_D} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{x_D,x_1} & \rho_{x_D,x_2} & \cdots & 1\end{bmatrix}\begin{bmatrix}\sigma_{x_1} & & & 0 \\ & \sigma_{x_2} & & \\ & & \ddots & \\ 0 & & & \sigma_{x_D}\end{bmatrix}. \tag{4.7}$$

Assuming that the coefficients are independent [54], the correlation between coefficients becomes 0 ($\rho_{x_i,x_j} = 0$) for any pair of coefficients x_i and x_j ($x_i \neq x_j$). In that case, $\mathbf{\Sigma}$ is a diagonal matrix with its determinant defined as the product of its diagonal elements:

$$|\mathbf{\Sigma}| = \prod_{i}^{D} \text{diag}_i(\mathbf{\Sigma}) = \prod_{i}^{D} \sigma_{x_i}^2 = \sigma_{x_1}^2 \sigma_{x_2}^2 \dots \sigma_{x_D}^2. \tag{4.8}$$

Finally, going back to Equation 4.6, the first term is defined as a constant for any FOD voxel (since it only depends on D fixed to 45) that we can ignore.

Taking into account the properties of logarithms and the fact that spherical harmonic coefficients x are usually small, we swapped the order of the operations from taking the determinant first and then applying the logarithm to computing the logarithm first and then summing the diagonal values to mitigate numerical precision issues. Considering that only the diagonal would have non-zero values, and assuming that $\log(0) = 0$, these approaches are mathematically equivalent. Consequently, we can define our differential entropy measurement as the trace of the logarithm of the covariance matrix[1] (tying the amount of uncertainty to the generalised variance of the predictions):

$$\mathbf{H}(x) \propto \log(|\mathbf{\Sigma}|) = \log(\sigma_{x_1}^2) + \log(\sigma_{x_2}^2) + \cdots + \log(\sigma_{x_D}^2) = \text{trace}(\log(\mathbf{\Sigma})). \tag{4.9}$$

4.2.3 Implementation Details

The subset of images from the Human Connectome Project (HCP) database [114], comprising 100 unrelated subjects (80 cases for training, 20 cases for testing), was used for all the experiments. The T1-weighted and DWI for each subject were acquired using a 3T Siemens 'Connectom' Skyra scanner. The HARDI volumes were captured using a resolution of 1.25 mm isotropic and 3 shells of b = 1000, 2000, 3000 s/mm^2 (90 directions/shell) together with 18 b0 volumes, while the T1-weighted images were acquired with 0.7 mm isotropic resolution, TR/TE = 2400/2.14 ms, and flip angle = 8°. To simulate a common clinical LARDI acquisition, we sub-sampled 32 directions from the most common b = 1000 s/mm^2 shell and included the mean of all the b0 volumes [109]. FOD maps for the LARDI and HARDI images were generated using the SS3T CSD [117] and MSMT CSD methods [27], respectively.

[1]Similarly to Gal [186], we abuse standard notation, and use $A \propto B$ to mean that A is a linear transformation of B.

Finally, fiber bundle elements were computed for each voxel. Each fixel is parameterised by the voxel in which it resides, the estimated mean orientation of the underlying fibers attributed to that bundle, and a fiber density (or partial volume fraction) [58].

All models followed the same training strategy with the exception of the MCDropout method that required the inclusion of dropout layers during training. The network weights were optimised for 50 epochs by minimising the L2 loss with the Adam optimiser (initial learning rate of 0.01) and a batch size of 4 on a single NVIDIA GeForce Tesla V100-SXM2 GPU (PyTorch 1.10.1 and Python 3.7.11). Before feeding the data into the models, the 4D FOD data was normalised into z-scores along each coefficient and cropped into patches of $64 \times 64 \times 64 \times 45$ with an overlap of 32. The final prediction was constructed by combining all output patches and averaging overlapping regions. $N_{de} = 5$ models were trained for DE and the outputs from individual networks were averaged for the final prediction. $N_{le} = 9$ heads were used to generate multiple intermediate outputs for LE. For MCDropout, a single model with dropout layers was used for each dropout rate (rate $\in [0.1, 0.32, 0.5, 0.7]$), with $N_{mcd} = 5$ stochastic forward passes for uncertainty estimation.

TABLE 4.1. Quantitative analysis of FOD and fixel[1] metrics between the reference HARDI data and other estimations.

	FOD Metrics		**Fixel Metrics**		
Methods	MSE[2] (↓)	ACC[3] (↑)	Angular Err.[4] (↓)	AFD Err.[5] (↓)	Peak Err.[6] (↓)
LARDI FOD	0.0590	0.7404	0.4286	0.1021	0.1116
FOD-Net [109]	-	0.8948	0.3421	0.0912	0.0719
DE ($N_{de} = 5$)	0.0126	0.9147	0.2378	0.0888	0.0738
LE ($layer = final$)	0.0180	0.8928	0.2677	0.0928	0.0898
MCDropout ($rate = 0.7$)	0.0208	0.8823	0.2728	0.0928	0.1166

[1]fiber bundle element. [2]Mean squared error. [3]Angular correlation coefficient.
[4]Angular error. [5]Apparent fiber density error. [6]Peak error.

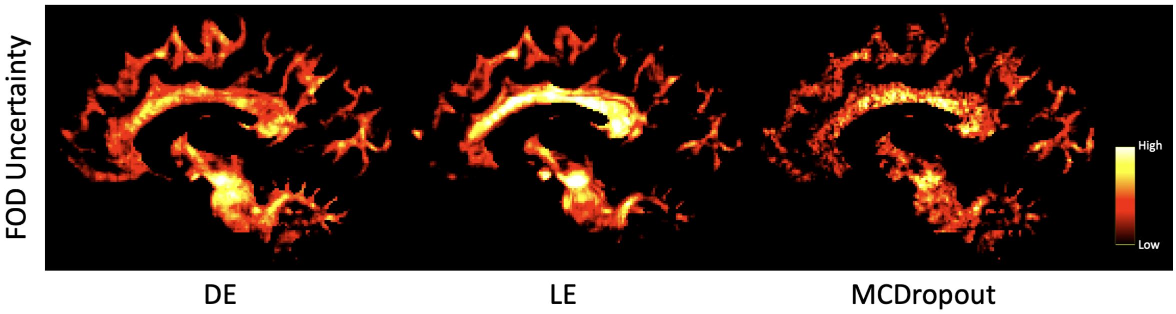


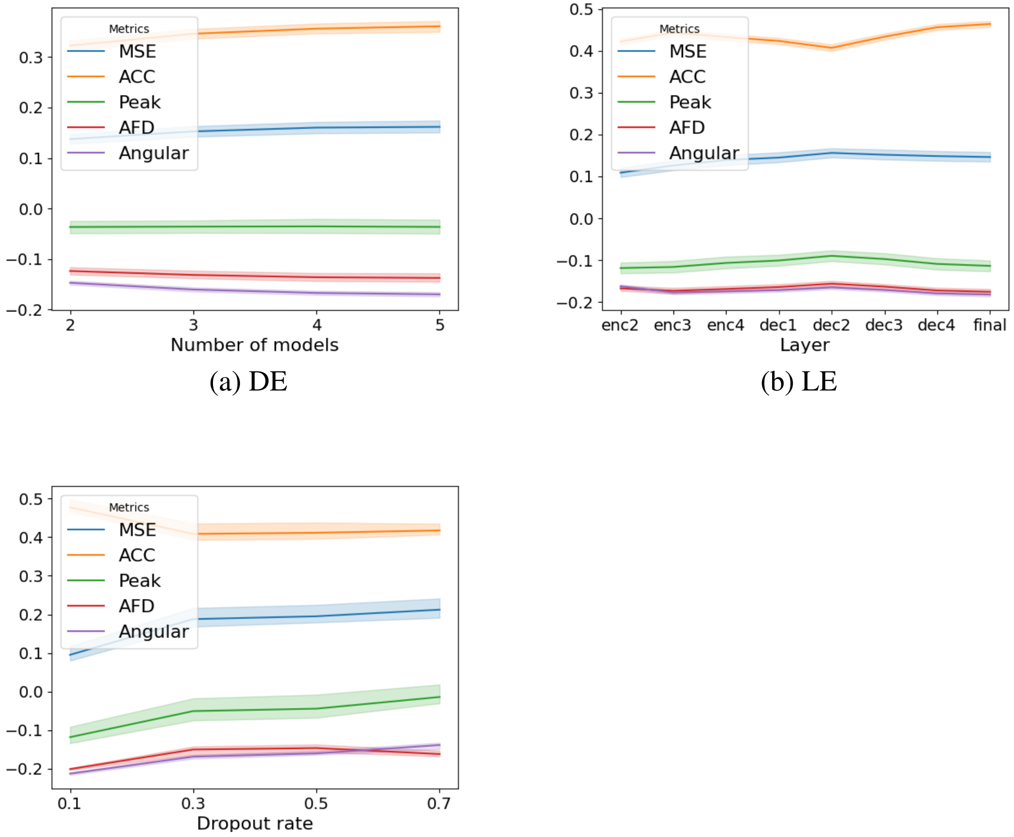
FIGURE 4.2. Visualisation of FOD uncertainty maps from different methods: DE ($N_{de} = 5$), LE ($layer = final$), and MCDropout ($rate = 0.7$).

4.3 Results

4.3.1 Model Validation

To validate the networks used for the uncertainty measurements, we computed different metrics based on the FOD and fixel estimates for the original LARDI FODs, the predictions from the state-of-the-art FOD-Net [109], and our CNNs using different sampling methods: DE, LE, and MCDropout. Two FOD metrics, mean squared error (MSE) and angular correlation coefficient (ACC), measured the reconstruction accuracy of FODs themselves. For downstream tasks including tractography and connectome mapping, fixel-level analysis served as a proxy to measure the reconstruction of fiber bundles. The mean absolute error for direction angle (most important for downstream tasks), peak (amplitude at the maximum peak per fixel), and apparent fiber density (AFD, integral of the FOD lobe) between the reference HARDI fixels and fixels from other methods were defined as the angular error, peak error, and AFD error, respectively. Table 4.1 shows that our CNN framework accurately enhances the FODs, evidenced by lower MSE, higher ACC, and reduced fixel-level errors from both FOD and fixel perspectives. These results ensured that our models were well-trained and comparable to the state-of-the-art FOD-Net. This validation is critical because any substantial underperformance could render the uncertainty estimation irrelevant.

4.3.2 FOD-level Analysis

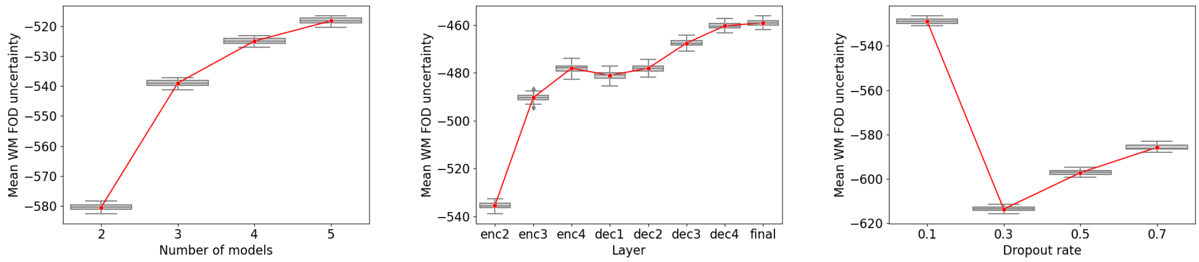


(a) DE

(b) LE

(c) MCDropout

FIGURE 4.3. Kendall correlation between FOD uncertainty and metrics against parameters of each sampling strategy. a, DE: Correlation with metrics of the final ensembled output against the number of models. b, LE: Correlation with metrics of the final layer output up to a given layer. c, MCDropout: Correlation with metrics of the prediction at each dropout rate. Lines represent FOD metrics: mean squared error (MSE) and angular correlation coefficient (ACC) between each output and the HARDI reference; and fixel metrics: angular error, peak error, and AFD error between predicted fiber bundle elements ('fixels') and the reference fixels. Error bands and mean values are determined subject-wise.



(a) DE (b) LE (c) MCDropout

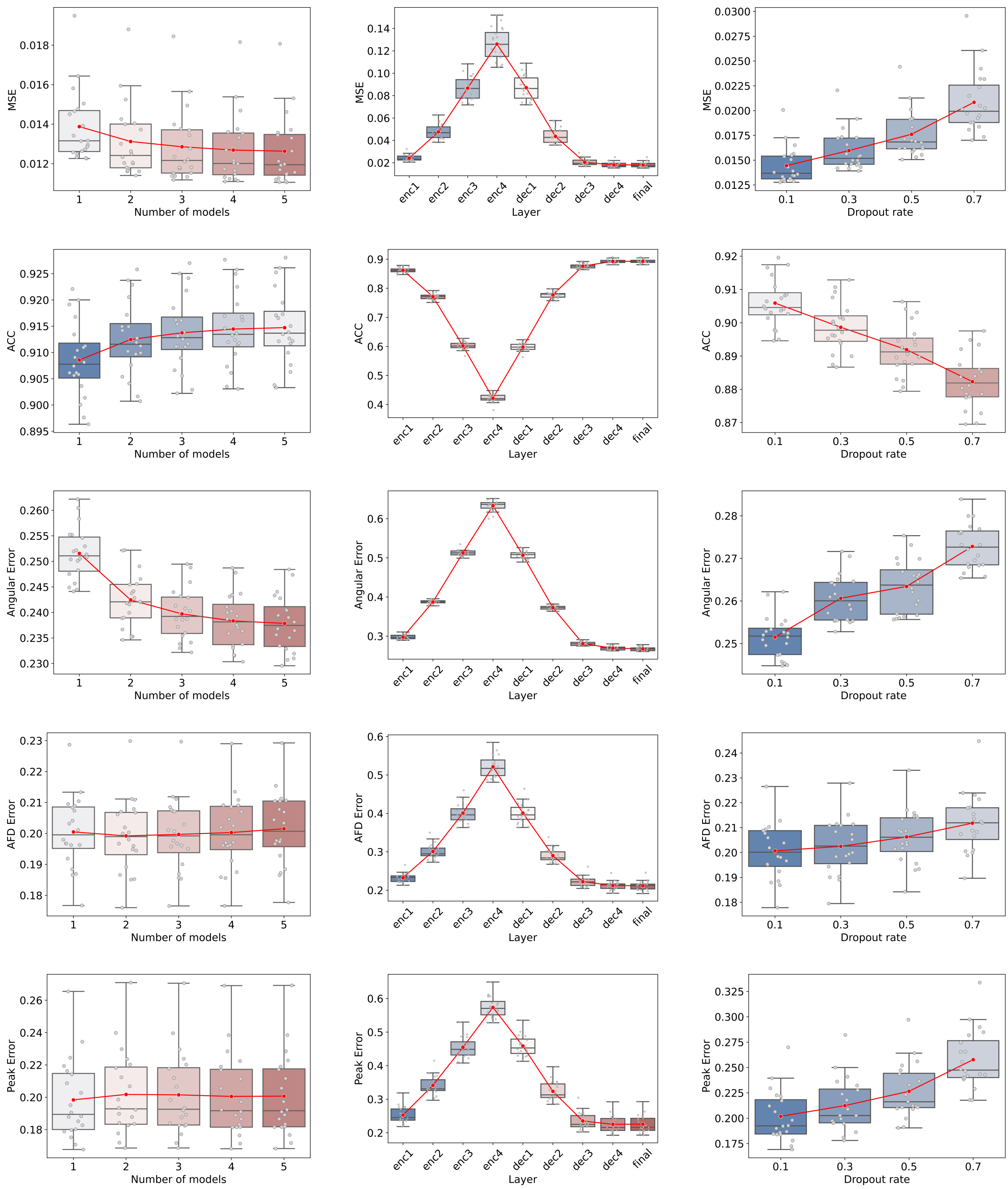
FIGURE 4.4. Box plots of mean FOD uncertainty within white matter (WM) estimated by each sampling method against their respective parameters. a, DE against different numbers of models. b, LE up to a given layer. c, MCDropout against different dropout rates. Error bands, medians, and mean values are determined subject-wise. The red line denotes the trend of the average.

4.3.2.1 Qualitative Analysis of FOD Uncertainty

Figure 4.2 illustrates the FOD uncertainty maps using different uncertainty quantification methods within WM. LE indicates a large disagreement (high uncertainty) on learned features of the CC between the encoder and decoder layers compared to other brain regions. As the network depth increases, the focus of our network shifts towards refining the feature representation of CC fibers from encoding to decoding phases. On the other hand, DE presents higher uncertainty in other WM regions outside the CC. Notably, uncertainty maps from MCDropout appear noisier. Since dropout introduces randomness by turning off neurons during training, the output variability can manifest as noisy estimates.

4.3.2.2 Correlation between FOD Uncertainty and FOD Metrics

To validate the uncertainty estimated with Equation 4.9, we analysed the Kendall correlation [195] between uncertainty and two measures of error for the predicted FODs with respect to the parameters of each strategy (number of ensembled models N_{de}, number of ensembled head predictions up to a given layer N_{le}, and dropout rate) as illustrated in Figure 4.3. Box plots of the FOD uncertainty within WM (image-wise mean) estimated by each sampling method and box plots of the FOD metrics of outputs from each sampling method against their respective parameters are illustrated in Figure 4.4 and Figure 4.5 to complement the



(a) DE (b) LE (c) MCDropout

FIGURE 4.5. Box plots of metrics from each sampling method against their respective parameters. a, DE against different numbers of models. b, LE up to a given layer. c, MCDropout against different dropout rates. Error bands, medians, and mean values are determined subject-wise. The red line denotes the trend of mean.

uncertainty analysis. The goal of this experiment is to determine if there is a monotonic relationship between uncertainty and erroneous predictions. It is essential to highlight that while it may seem intuitive to equate high uncertainty with inaccuracy, this is not necessarily the case. Uncertainty merely indicates the confidence of the model in its predictions, but not necessarily their correctness. By rigorously investigating this relationship, we may provide deeper insights into limitations of our model.

One of the first observations is that the highest correlations can be observed for the ACC, while lower correlations can be observed for the MSE metric. This could be explained by the fact that ACC, as a correlation metric, is insensitive to differences in scale between coefficients, focusing instead on the relative ordering of the values. For instance, if our predicted coefficients are consistently twice the ground truth values, MSE would be high due to the magnitude of the error, but ACC would remain unaffected, reflecting the correct directional information. Essentially, the fiber orientations would remain accurate despite the scaling error. Therefore, a higher correlation with ACC suggests that our uncertainty estimates are more relevant to the actual fiber orientation accuracy, which is more critical for downstream tasks than the absolute error magnitude captured by MSE.

Next, we focused on each method separately.

DE: The correlation between uncertainty and FOD metrics of the final ensembled output increases with the number of models (Figure 4.3a), indicating that increased uncertainty (Figure 4.4a) can better represent prediction mistakes. A larger ensemble also decreases MSE and increases ACC in Figure 4.5a. Consequently, we can infer that not only the accuracy improves with larger DEs but that with a sufficiently large number of models uncertainty can be used as a proxy for the reconstruction errors in the absence of the true HARDI FODs.

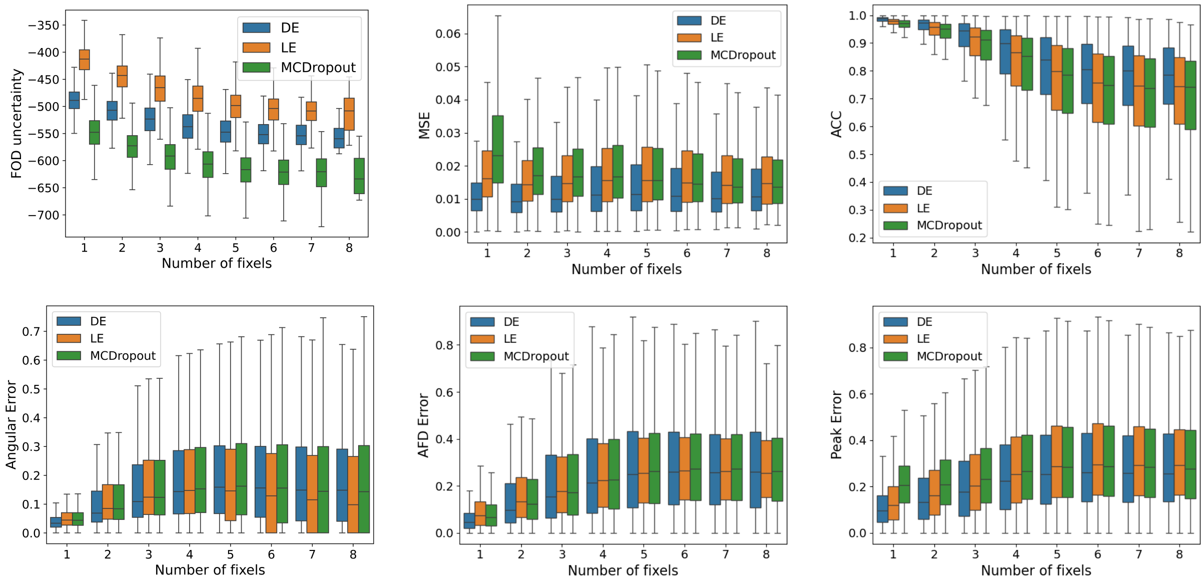
LE: In general, the correlations (Figure 4.3b) follow similar trends to the FOD metrics (Figure 4.5b). The majority of high uncertainty correlations with MSE come from the encoder, suggesting that the errors on the final prediction are related to the uncertainty of the encoder. Each 'x' coordinate on the graph represents the combination of outputs from all previous layers, showing that as more layers are pooled, the correlation between uncertainty and accuracy increases. This implies that more layers used, the better the uncertainty reflects the accuracy in the final prediction, leading to a more robust uncertainty estimation. For FOD enhancement, this suggests that larger networks with more intermediate outputs provide a better understanding of confidence levels and prediction quality, potentially outperforming DE with a single forward pass.

MCDropout: As the dropout rate increases, the correlation between both FOD-level metrics and uncertainty increases (Figure 4.3c), with the exception that the lowest dropout rate has the highest correlation with ACC. As dropout introduces noisy features to prevent overfitting, a low dropout rate leads to less noisy estimates and a network with the highest ACC and lowest MSE in Figure 4.5c and the highest FOD uncertainty in Figure 4.4c.

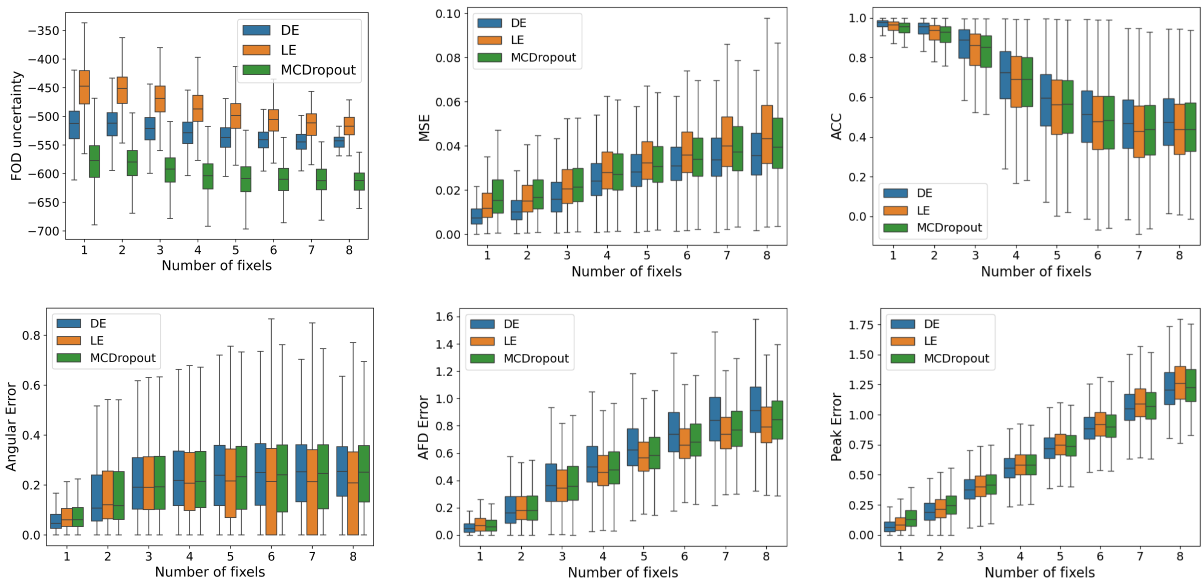
4.3.3 Fixel-level Analysis

To better probe the relationship between FOD predictions and fixels as a proxy for fiber bundles, the correlations between fixel-level errors and FOD uncertainty are also illustrated in Figure 4.3. The first observation is that all methods show a relatively low absolute correlation (similar to the correlation with MSE) and a negative tendency for peak errors. These results suggest a potential disconnect between errors in the coefficients themselves and their effect on downstream tasks (e.g. fixel analysis), and that higher uncertainty might be a positive thing in terms of accurate predictions. Examining the different methods independently reveals similar trends and correlation values.

At a first glance, these results suggest that there is less correlation between the uncertainty and the fixel-level errors than at the FOD level. A particularly important fact is that one single set of 45 coefficients can lead to multiple fiber bundle elements per voxel depending on the number of lobes defined by the spherical harmonics. Voxels with different numbers of fixels might follow different trends and comparing their uncertainty estimates without taking that into account could explain these overall low correlation values. To validate this hypothesis, we plotted the distribution of the uncertainty as box plots by grouping voxels based on their number of fixels in the input LARDI image and the real HARDI FODs as illustrated in Figure 4.6.

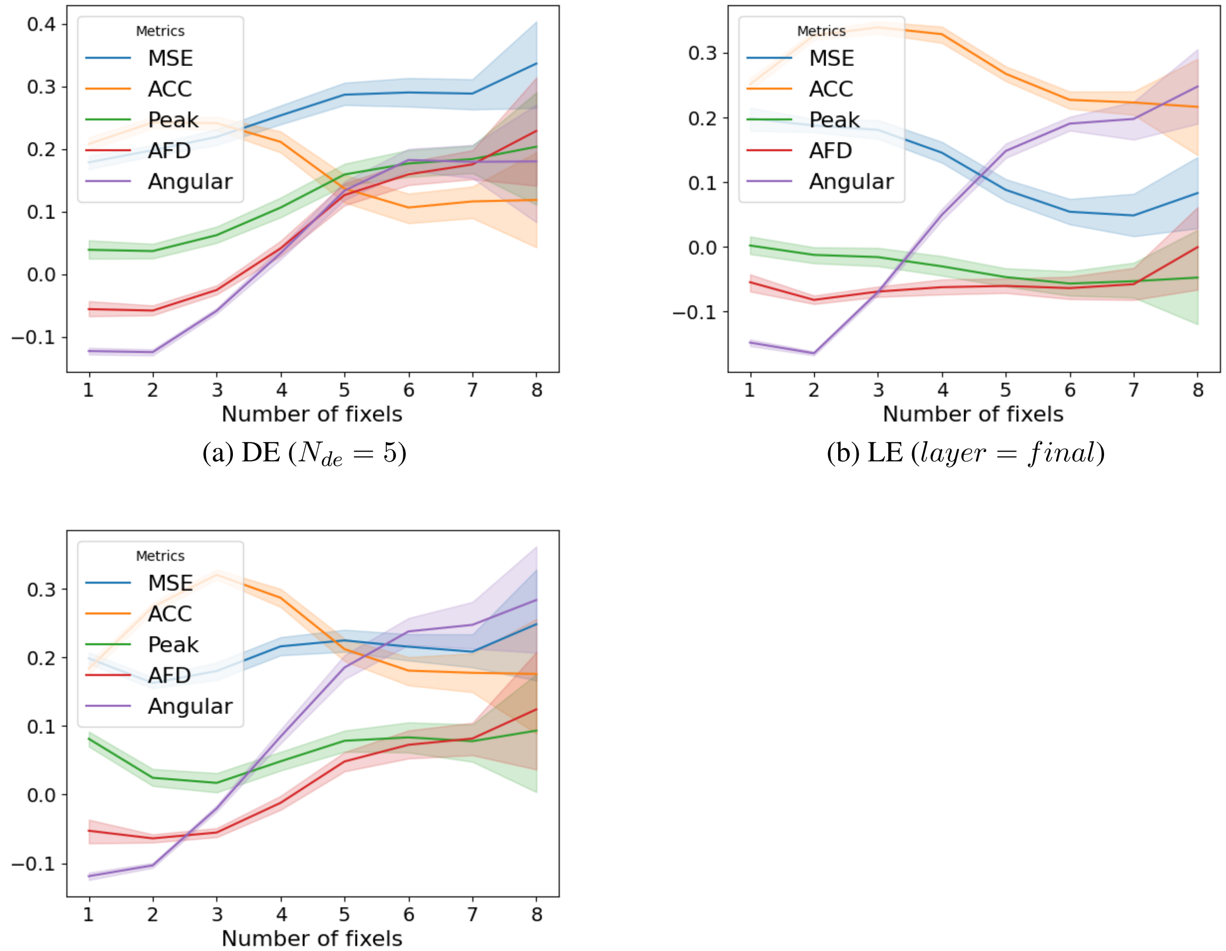


(a) Group voxels based on the number of fixels from the input data.



(b) Group voxels based on the number of fixels from the ground truth data.

FIGURE 4.6. Box plots of FOD uncertainty and metrics against voxels with different numbers of fixels. Group voxels based on the number of fixels from: a, the input LARDI image; or b, the ground truth HARDI data. Blue, orange, and green boxes represent DE ($N_{de} = 5$), LE ($layer = final$), and MCDropout ($rate = 0.7$).



(a) DE ($N_{de} = 5$)

(b) LE ($layer = final$)

(c) MCDropout ($rate = 0.7$)

FIGURE 4.7. Kendall correlation between uncertainty and metrics against the number of input fixels. Error bands and mean values are determined subject-wise.

In both cases, the tendency of the box plots remains the same. As the number of fixels increases, the differential entropy decreases as well as its dispersion (especially when considering the true fixel number on the ground truth). The figure also shows an overall lower performance (FOD and fixel level) with higher number of fixels. Higher numbers of fixels lead to more consistent network enhancement and greater prediction certainty, as voxels with many ground truth fixels are more similar to the input LARDI data, making the network more confident and resulting in lower uncertainty.

To investigate that further, we examined the monotonic correlation between the error metrics and the uncertainty by grouping voxels based on their number of fixels as illustrated in Figure 4.7. Comparing Figure 4.3 and Figure 4.7 confirms our hypothesis that the number of fixels is a confounder in terms of the correlations between FOD uncertainty estimates and fixel-based measures. Specifically, angular error is more correlated when the number of fixels increases. This fact also holds true for the AFD and peak errors to an extent, particularly with DE uncertainty estimates. Notably, the negative correlations are observed for the angular error in regions with a small number of fixels (< 3).

Another interesting detail is that the dispersion of the correlation remains small up until 4 fixels, where it starts to increase, showing a wide range of correlation values for 8 fixels in Figure 4.7. This is not necessarily a surprising fact, as in general most modeling methods assume that at most 3 fiber bundles can cross a single voxel [196, 197]. When the number of fixels exceeds this common assumption, regions with more than 3 fixels become challenging, introducing aleatoric uncertainty and potential ground truth errors. Consequently, the ground truth is likely to be incorrect beyond 3 fixels, resulting in more errors and mixed relationship between model uncertainty and prediction errors. In such regions, models might produce different predictions or worse overall results despite predicting consistent results (low uncertainty, high error). As a result, subjects could display varied patterns: easy subjects with high certainty and hard subjects with low certainty. Correlation changes might arise from worse predictions or less consistency, leading to varied accuracy-uncertainty relationships across subjects. In summary, in regions with a high number of fixels, the increased complexity (more crossing fibers) and noisier estimates may cause the network to default to a similar behavior, such as removing small coefficients, resulting in low uncertainty but different error levels per patient, thus showing a large range of correlations.

Finally, investigating the FOD-based errors reveals different trends depending on the sampling method in Figure 4.7. While DE and MCDropout have an increased correlation with MSE as the number of input fixels in a voxel increases, LE shows fluctuations, with a large drop after 3 fixels. In the case of LE, that might suggest that for less plausible regions (with more than 3 crossing fibers), the uncertainty or error estimates might not be representative of each other. In terms of ACC, all uncertainty strategies exhibit a decreasing trend after 3 fixels, reinforcing the FOD uncertainty of more complex fiber structures and their link to errors in the coefficients of FOD are disturbed by the increasing fixel complexity. This also highlights the need for future studies to improve the robustness of our uncertainty estimation method at the fixel level.

4.4 Summary

Uncertainty analysis is crucial in the enhancement of FODs as it helps to understand the reliability of enhanced predictions. This work presents a new formulation for estimating FOD uncertainty by framing the FOD enhancement problem as a multivariate regression problem. Our findings from the FOD-level analysis indicate that our proposed FOD uncertainty is representative of FOD reconstruction accuracy when employing current sampling techniques with appropriately tuned parameters, i.e., the number of trained models and layers, or the dropout rate. LE is the uncertainty method most representative of errors at the FOD level, while DE is the most representative measurement of errors at the fixel level in terms of correlation. The uncertainty also suggests the need for a mechanism (e.g. weighting) to bring greater variability and improved accuracy when handling complex fiber structures, preventing the network from defaulting to similar behaviour. Additionally, our results highlight the relationship between uncertainty and the number of fiber bundle elements per voxel, indicating the need to refine the proposed uncertainty measurements from a fixel perspective in future studies to account for regions with multiple crossing fibers, and estimate FOD uncertainty in an order-wise manner as different spherical harmonic basis might have different uncertainty. The integration of our proposed uncertainty with CNNs, along with a comparative analysis of existing learning-based uncertainty estimation techniques in other regression tasks, could be further explored to benchmark effectiveness. Additional experimental investigations could be undertaken to thoroughly assess the reliability and generalisability, exploring the impact of dataset size, testing across different brain regions and evaluating the robustness of the method under varying acquisition conditions.

CHAPTER 5

UFREE: Unified Fiber Orientation Distribution Resolution Enhancement

---

DWI is a MRI technique that enables non-invasive mapping of structural brain connectivity. FOD is a crucial intermediate model used to represent complex white matter fiber configurations within a voxel and serves as a key step for subsequent brain tractography and connectome analysis. However, obtaining reliable FOD estimates is particularly challenging due to the constraints of acquisition protocols, which often feature low spatial and angular resolutions in non-research-focused clinical settings. Current deep learning approaches, including those presented in previous chapters, have shown promise in addressing these limitations, they typically focus on enhancing either low angular or low spatial resolution inputs rather than providing a unified solution. This highlights **Challenge 5: Unified Solutions for Multiple Resolutions** in Section 1.4.5.

Here, we present a unified and robust deep learning framework, namely UFREE, for FOD enhancement from varied resolutions. UFREE can produce high quality FOD images that align closely with ground truth data from multimodal analysis of FOD derivatives. This framework holds the potential to bring research-level FOD quality to routine neuroimaging acquired in the clinical environment.

5.1 Introduction

DWI is a MRI technique that can map structural brain connectivity non-invasively. FOD is a crucial intermediate model employed to represent the probability of WM fiber distribution in the brain, offering detailed insights into the orientation and density of fiber bundle elements ('fixels') within each voxel. This modeling approach accounts for the complex structure of

WM when conducting tractography, which can subsequently be used to probe structural brain connectivity [112].

The quality of FOD is highly dependent on the angular resolution (number of gradient directions acquired) and the spatial resolution. Through CSD [27, 42, 54], high-resolution (HR) FOD estimates can be reliably computed from multi-shell acquisitions of high angular resolution (HAR) and high spatial resolution (HSR) diffusion imaging. In general clinical practice, low-resolution (LR) diffusion MRI (i.e., low angular resolution - LAR, single-shell, typically < 60 directions; low spatial resolution - LSR, $> 2mm$ isotropic) is more commonly acquired due to the constraints of acquisition time for patients [198, 104] in non-research-focused clinical sites.

Enhancing FOD angular resolution remains a key challenge in diffusion MRI, as the constrained angular detail in LAR images loses crucial information about complex fiber bundles and limits the fidelity of downstream analyses using corresponding FOD derivatives [121]. Recent efforts [104, 115, 106, 107] have focused on predicting HAR-FODs from patches of single-shell LAR FODs [54]. Low spatial resolution, meanwhile, introduces spatial blurring that could also degrade FOD quality. Another recent work [108] presented a fiber-population guided framework to address the low spatial resolution issue of DWI by directly applying super-resolution on FODs to mitigate the erroneous differences in FODs and the associated derivatives. These tailored deep learning methods have demonstrated acceptable performance in reconstructing HR FODs from their low angular resolution or low spatial resolution counterparts, respectively. However, existing methods cannot effectively enhance FODs when dealing with flexible input resolutions, whether low spatial, low angular, or a combination of both. They are limited by their task-specific focus and often require much effort of retraining for different LR datasets. These challenges underscore the need for a unified enhancement method to produce reliable FOD coefficients from data acquired with LR diffusion imaging protocols.

In this work, we introduce a Unified FOD angular and spatial Resolution Enhancement framEwork (UFREE) to estimate HR FOD from any type of LR FOD. UFREE employs a two-stage learning strategy. In the first stage, high-resolution feature learning is shared across tasks. In the second, task-specific stage, LR FOD features are aligned with a set of HR FOD features from pre-trained implicit HR FOD priors through latent feature matching. Our contributions are as follows: (1) To the best of our knowledge, UFREE is the first two-stage framework that pre-train HR information applicable to any LR-to-HR FOD enhancement task, minimizing re-training requirements when adapting to new LR datasets. (2) While we utilize VQ-VAE as our encoder-decoder architecture, this unified framework is flexible and can incorporate other advanced autoencoders, enhancing its overall potential. (3) Experiment results show the effectiveness of UFREE and demonstrate its generalizability across different input types while enhancing FOD quality.

5.2 Methods

The FOD enhancement task is framed as a conditional image generation process in this work, a single HR FOD image (denoted by Y) corresponds to multiple LR FOD images (denoted by X) derived from diverse distributions for different tasks. To maximize the utility of HR image information across various LR-to-HR tasks, we propose a two-stage FOD enhancement framework, with each stage following an encoder-decoder structure. In Stage 1, the model learns the shared HR image distribution in the encoding phase and produces an HR image generator. Stage 2 then applies task-specific transfer learning techniques through latent feature matching. The model architecture is illustrated in Fig. 5.1.

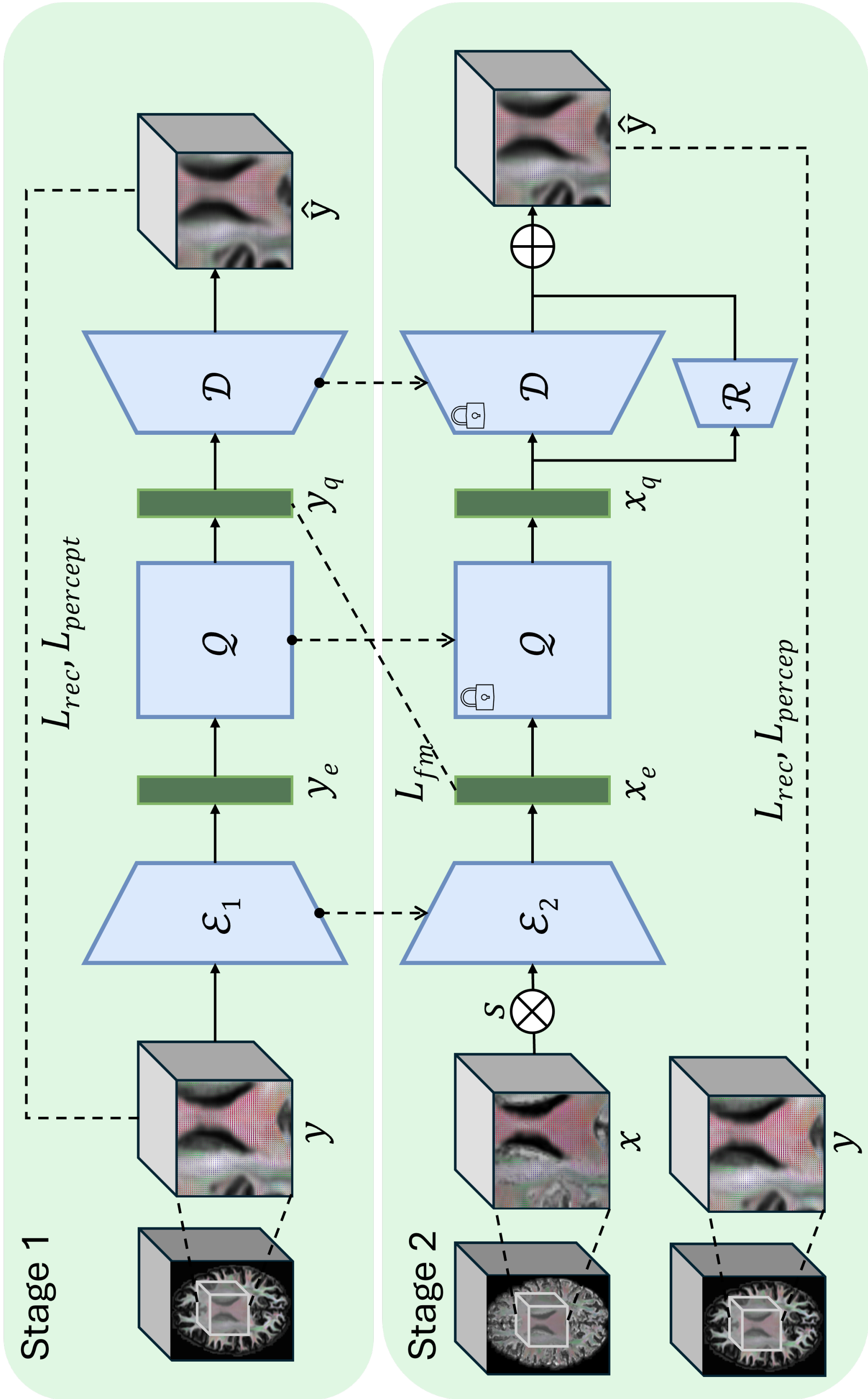


FIGURE 5.1. The overview of UFREE.

5.2.1 Stage 1: Shared HR Learning

In Stage 1, we aim to learn a decoder for high quality image generation and an encoder with strong generalization ability. Note any autoencoder structures can be applied here, a VQ-VAE autoencoder [199] with perceptual loss is chosen for implementation given the high quality of the generated image.

The forward propagation of the VQ-VAE follows:$\hat{y} = \mathcal{D}(\mathcal{Q}(\mathcal{E}_1(y)))$, where $\mathcal{E}_1$ is the Stage-1 encoder, $\mathcal{Q}$ is the Stage-1 vector quantization module, and $\mathcal{D}$ is the Stage-1 decoder. The encoder $\mathcal{E}_1$ learns a continuous latent distribution of the input image, and the quantizer $\mathcal{Q}$ converts the continuous latent vectors into discrete indices, applying a codebook to obtain the quantized latent distribution. This process enhances the generalization ability of the encoder. Finally, the decoder $\mathcal{D}$ generates the reconstructed image.

Two losses are applied to train Stage 1:

$$L = L_{rec}(\hat{y}, y) \times \text{mask} + L_{percep}(\hat{y}, y), \tag{5.1}$$

where L_{rec} is the $L1$ loss between $\hat{y}$ and y, applying a mask to focus on brain regions. L_{percep} measures the semantic distance between $\hat{y}$ and y using a pre-trained model from 'medicalnet_resnet10_23datasets' [200].

5.2.2 Stage 2: Task-specific Learning

Stage 2 also follows an encoder-decoder structure with components: $\mathcal{E}_2$ whose weights are initialized by $\mathcal{E}_1$, as well as the frozen modules $\mathcal{Q}$ and $\mathcal{D}$, to reduce the training effort for the Stage-2 model. We employ two transfer learning techniques: fine-tuning $\mathcal{E}_2$ for latent feature matching, and incorporating a residual module in $\mathcal{D}$ for better quality output.

5.2.2.1 Scale Factor before Encoder

The distributions of X and Y can be similar in terms of brain structure, with the primary difference being the scale of voxel values. Therefore, before feeding X into $\mathcal{E}_2$, we apply a simple scale factor [201] on X. The scale factor s_c is calculated per channel by obtaining a statistical ratio from the training set. Specifically, for each channel, we compute a list of $s_{x,c}$ by comparing x_c and y_c:

$$s_{x,c} = \sqrt{\frac{\sum\limits_{w,h,d} ||y_{w,h,d,c}||^2}{\sum\limits_{w,h,d} ||x_{w,h,d,c}||^2 + \epsilon}}, \text{ where } y_{w,h,d,0} \neq 0. \tag{5.2}$$

Next, we apply the interquartile range (IQR) method to remove outliers that fall outside 1.5 times the IQR from the 1st and 3rd quartiles. We then compute $s_c = \sum\limits_{x \in X} s_{x,c}/|X|$. Using this naive statistical model, we obtain a $\hat{y}_{\text{naive}} = x \times s$, which can be fed into $\mathcal{E}_2$ directly. We denote the encoding process of Stage 2 as $x_e = \mathcal{E}_2(x \times s)$, where $\mathcal{E}_2$ is initialized by $\mathcal{E}_1$ and remains trainable.

5.2.2.2 Latent Feature Matching

The training process of Stage 2 primarily focuses on latent feature matching since x_e does not always produce the exact latent vector required by $\mathcal{Q}$ and $\mathcal{D}$. To address this, we apply a feature mapping loss by calculating the L_2 distance between the encoded features x_e and the ideal vectors y_q needed to generate y, where $y_q = \mathcal{Q}(\mathcal{E}_1(y))$, inspired the codebook quantization loss in VQ-VAE.

Meanwhile, we still apply the quantizer $\mathcal{Q}$ to x_e to obtain the quantized output, denoted by $x_q = \mathcal{Q}(x_e)$. Here, $\mathcal{Q}$ is non-trainable.

5.2.2.3 Residual Module on Decoder

Although the latent vector generated from $\mathcal{Q}$ should serve as a suitable input for $\mathcal{D}$, we aim to maximize the generative capacity of $\mathcal{D}$. Therefore, we introduce a residual module for each block of the decoder, while freezing all existing parameters of $\mathcal{D}$. For each block i, this is represented as $x_{d,i} = \mathcal{D}_{i-1}(x_{d,i-1}) + \mathcal{R}(x_{d,i-1})$. Thus, the overall forward process is:

$$\hat{y} = (\mathcal{D} + \mathcal{R})(\mathcal{Q}(\mathcal{E}_2(x \times s))), \tag{5.3}$$

where s is pre-calculated, $\mathcal{Q}$ and $\mathcal{D}$ are frozen, and $\mathcal{E}_2$ and $\mathcal{R}$ are trainable. Similar to Stage 1, the loss function for Stage 2 is:

$$L = L_{rec}(\hat{y}, y) \times \text{mask} + L_{percep}(\hat{y}, y) + L_{fm}(x_e, y_q). \tag{5.4}$$

5.2.3 Datasets and Data Processing

The HCP database[1] [114] provided the T1-weighted (0.7 mm isotropic resolution, TR/TE = 2400/2.14 ms, flip angle = 8°) and DWI (b = 1000, 2000, and 3000 s/mm^2, 90 directions each and 18 b0 volumes, 1.25 mm isotropic, TR/TE = 5520/89.5 ms, flip angle = 78°) data used in this study, which includes 100 unrelated subjects (80 for training, 20 for testing). After multiple correction steps [128], the processed DWI data HR-DWI was used to produce the reference HR FODs using MSMT CSD in MRtrix [116, 63]. To simulate clinical LAR acquisition, 32 directions were sub-sampled from the b = 1000 s/mm^2 shell [104], and SS3T CSD [117] was applied to generate the LAR-FODs. To emulate lower spatial resolution [108], HR-DWI data were spatially down-sampled to 2.5 mm isotropic, resampled back using interpolation, and then MSMT CSD was applied to create LSR-FOD inputs.

To further analyze the practical consequences, whole-brain tractograms were performed with the HCP pipeline in MRtrix, applying ACT driven by iFOD2 using a 5TT mask [64]. Key parameters included a step size of 0.5 × voxel size, max curvature of 45° per step, and an amplitude threshold of 0.06, with randomized seeding from the GM-WM interface to produce 10 million streamlines. SIFT2 weighting was then applied to assign weights to each streamline [69]. Subsequently, T1 images were parcellated into 84 brain regions using the Desikan-Killiany atlas [70]. The connectome matrix for each FOD-based tractogram was computed by summing SIFT2-weighted streamlines between each pair of nodes.

[1] https://www.humanconnectome.org/

5.2.4 Experimental Details

We adopted a patch-based learning approach. During the model learning phase, patches were randomly cropped from the original images in each epoch, with size of $96 \times 96 \times 96 \times 45$, where 45 being the number of spherical harmonic coefficients (lmax = 8). In the inference phase, a grid-based patch extraction with overlap was applied, and the patches were merged to reconstruct the final output. The framework was trained for 200 epochs with early stopping by a learning rate of $1e^{-4}$ on a single NVIDIA GeForce Tesla V100-SXM2 GPU using PyTorch and a customized MONAI package. We compared UFREE with state-of-the-art FOD-Net [104]. Since FOD-Net is specifically designed for angular resolution enhancement and requires large training efforts, we compared UFREE with VQ-VAE for spatial resolution enhancement.

5.3 Results

5.3.1 FOD Assessment

Two key metrics, Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR), were employed to quantify the impact of enhancement on image quality and the similarity of FODs. Additionally, the angular correlation coefficient ($\mathbf{r_{Angular}}$) [139] was used to measure the correlation of spherical harmonic coefficients.

We began by analyzing the differences between two LR datasets: LSR-FODs and LAR-FODs in Table 1, identifying a big gap in their distribution. Notably, angular resolution enhancement emerged as a more challenging task. However, after applying UFREE, the disparity between LSR and LAR estimations is much smaller, indicating UFREE's ability to generate comparable quality in predictions regardless of input type, resulting from the shared HR learning in Stage 1.

Method	Tissue Type	MSE (↓)	PSNR (↑)	$\mathbf{r_{Ang.}}$ (↑)
LR Diff.	Brain	0.0872	75.9207	0.4658
UFREE Diff.	Brain	**0.0029**	**95.3977**	**0.9506**
LAR	WM	0.0568	70.7430	0.7590
	JCWM	0.0809	60.2084	0.5265
	WM-SGM	0.1060	55.2939	0.4534
VQ-VAE	WM	0.0730	69.7225	0.6826
	JCWM	0.0333	64.6214	0.4330
	WM-SGM	0.0541	58.1384	0.3713
FOD-Net	WM	0.0238	74.3551	0.8614
	JCWM	0.0187	66.6154	0.7013
	WM-SGM	0.0361	59.8562	0.6115
UFREE	WM	**0.0218**	**75.0527**	**0.8859**
	JCWM	**0.0132**	**68.2915**	**0.7528**
	WM-SGM	**0.0300**	**60.7776**	**0.6709**

TABLE 5.1. Quantitative FOD assessment. The top two rows show differences between LSR and LAR estimations. The bottom rows compare metrics between reference HR FOD images and LAR-based estimations.

Next, we examined whether GM partial volume affected the efficacy of FOD enhancement methods. The VQ-VAE in the table indicates its application in learning an LR-to-HR task. Experimental results illustrate that UFREE tended to outperform baseline methods in enhancing FOD, suggesting a promising direction warranting further investigation. This performance is evidenced by the lowest MSE, highest PSNR, and optimal $\mathbf{r_{Angular}}$ across all brain regions, specifically: (1) deep WM regions, and (2) WM/GM boundary regions, which include juxtacortical (JCWM) or subcortical (WM-SGM) areas containing both WM and GM. These findings underscore the effectiveness of UFREE in mitigating partial volume effects and demonstrate its generalizability across different input types while enhancing FOD quality.

5.3.2 Fiber Bundle Element Assessment

The impact of the FOD enhancement was then assessed using fixel-wise metrics [104], reflecting the angular and amplitude accuracy. The mean absolute error for direction angle (most critical for tractography), peak (amplitude at the maximum peak per fixel), and fiber density (FD, the integral of the FOD lobe) between the reference HR fixels and fixels from other methods were defined as the angular error ($\mathbf{E_{Angular}}$), peak error ($\mathbf{E_{Peak}}$), and FD error ($\mathbf{E_{FD}}$), respectively.

To evaluate the performance on regions with varying fiber complexity, three main anatomical regions of interest (ROIs) were defined via TractSeg [196, 104]: a single fiber region (ROI 1), 2-way fiber crossing region (ROI 2), and 3-way fiber crossing region (ROI 3). The results shown in Table 2 demonstrate that the proposed UFREE achieves better performance for all reported metrics in ROI 1. Although improvements were reduced in ROI with two fibers crossing, they became more significant on more complex the fiber configurations.

Method	ROIs	$\mu\mathbf{E_{Ang.}}$ ($\downarrow$)	$\mathbf{E_{FD}}$ ($\downarrow$)	$\mu\mathbf{E_{Peak}}$ ($\downarrow$)
FOD-Net	ROI 1	0.1297	0.1668	0.1455
	ROI 2	**0.1384**	**0.1939**	**0.2581**
	ROI 3	0.1528	0.3151	0.3940
UFREE	ROI 1	**0.1163**	**0.1267**	**0.1271**
	ROI 2	0.1438	0.2585	0.3635
	ROI 3	**0.1140**	**0.2630**	**0.3092**
VQ-VAE	ROI 1	0.2310	0.3330	0.3769
	ROI 2	0.2115	0.4646	0.6867
	ROI 3	0.1927	0.4976	0.6594
UFREE	ROI 1	**0.1153**	**0.1333**	**0.1384**
	ROI 2	**0.1433**	**0.2634**	**0.3620**
	ROI 3	**0.1135**	**0.2751**	**0.3228**

TABLE 5.2. Quantitative fixel assessment between the reference HR FOD images and other LR based estimations for angular resolution (top) and spatial resolution (bottom) enhancement.

5.3.3 Connectome Assessment

A structural connectome [20] serves as a graph matrix with nodes denoting grey matter regions and edges denoting the brain connections between them. To assess the overall difference between a LR estimate and its corresponding HR matrix, we defined a disparity matrix [104] and computed the subject-wise average of mean edge disparities ($\mu\mathbf{Disparity}$). The percentage of significantly different edges was calculated using edge-wise paired t-tests ($p < 0.05$) with false discovery rate (FDR) correction to further analyze disparity. As shown

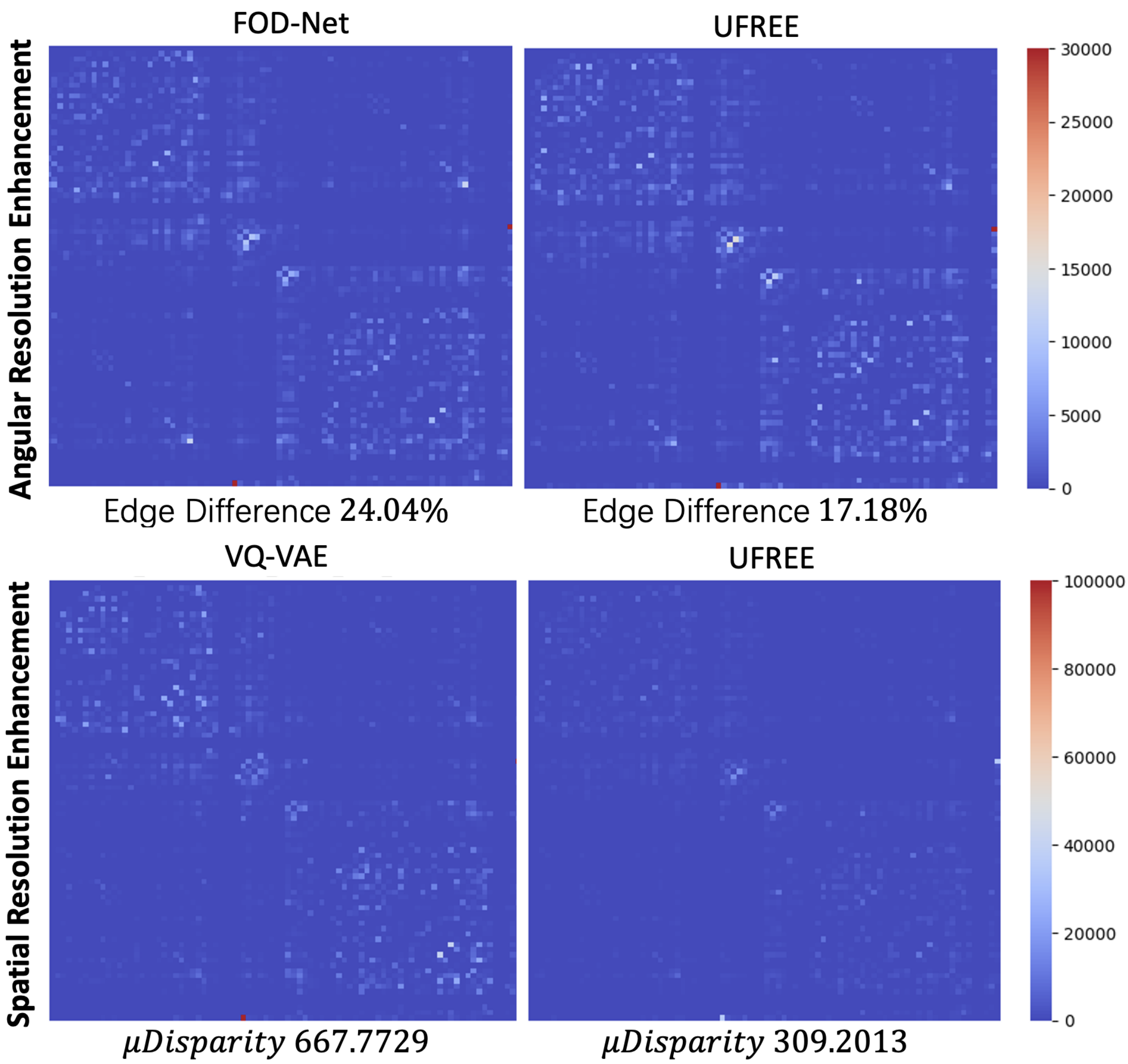


FIGURE 5.2. Connectome assessment between the reference HR connectomes and other LR based estimations.

in Fig. 5.2, UFREE produced high quality connectomes, with the least number of significant different edges and the lowest disparity relative to the ground truth.

5.3.4 Discussion

While recent methods such as FOD-Net and its improved version FastFOD-Net have demonstrated strong performance for angular resolution enhancement, they are specifically designed for LARDI settings and do not generalize to other types of resolution degradation. In contrast, UFREE is proposed as the first unified framework capable of simultaneously enhancing both angular and spatial resolution. This generalizability introduces certain trade-offs, and UFREE may not always outperform task-specific models in scenarios narrowly optimized for angular resolution only. However, its ability to operate across heterogeneous input settings without retraining or reconfiguration represents a significant step toward practical, flexible diffusion MRI enhancement. We believe this unified design fills a key gap in the current literature, and paves the way for more adaptable deep learning-based tools in clinical diffusion imaging. As such, direct comparisons with task-specific models like FastFOD-Net are not entirely appropriate, as performance may be confounded by differences in target domains and optimization goals. Instead, this work focuses on demonstrating the versatility and extensibility of UFREE as a general solution, rather than pursuing specialized performance in one narrowly defined setting.

5.4 Summary

In this work, we introduced UFREE, a deep learning framework for Unified Fiber Orientation Distribution Resolution Enhancement. By leveraging implicit high-resolution priors, UFREE effectively reconstructs high-resolution FODs with minimum effort of task-specific re-training. The flexible encoder-decoder structure also opens opportunities for future improvements, allowing for the integration of more advanced image encoders and generators to further enhance performance. UFREE holds the potential to generalize to more low quality FOD images obtained across diverse clinical protocols or with different types of artifacts, supporting broader applicability in future clinical and research settings.

CHAPTER 6

Conclusions and Future Outlook

6.1 Conclusions

In conclusion, this thesis has made significant strides in bridging the gap between the promise of deep learning methods and their practical application in clinical neuroscience, particularly in the context of diffusion MRI data. Through the development of innovative methods and frameworks, we have addressed several critical challenges that have hindered the broader adoption of deep learning solutions in this field.

Chapter 2 introduced FOD-Net 2.0, an enhanced version of FOD-Net that improves the angular resolution of FODs, demonstrating its potential on both healthy and clinical datasets, and laying the foundation for method design in clinical settings. Chapter 3 advanced this work further by introducing FastFOD-Net, a more practical solution for real-world clinical application, and establishing a comprehensive clinical evaluation platform. This include These two chapters address the first challenge of clinical rationale and method design by demonstrating its effectiveness and efficiency on both healthy and clinical datasets.

In Chapter 2, we also explored the optimal acquisition protocols for deep learning solutions, focusing on gradient direction counts, and identified the best trade-off between acquisition quality and model performance. Chapter 4 then tackled the challenge of enhancing FODs from low-angular resolution clinical scans, introducing a novel multivariate regression approach with uncertainty techniques that improved the reliability and transparency of deep learning-driven enhancements.

Finally, Chapter 5 introduced the Unified FOD Angular and Spatial Resolution Enhancement Framework (UFREE), a unified solution for enhancing resolution across diverse FODs, showcasing the power of deep learning to transform clinical imaging across different data disruptions.

Together, these chapters address critical barriers to the integration of deep learning in clinical neuroscience, including efficiency, clinical evaluation, acquisition protocols, explainability, and handling data disruptions. This thesis provides novel solutions that transform deep learning from a theoretical promise to a practical reality in diffusion MRI analysis, with potential for significant clinical impact. In summary, this work advances the field by addressing key challenges, offering both technical and clinical contributions, and laying the foundation for future research to further integrate deep learning into clinical practice.

Throughout this thesis, we have directly addressed key barriers to the broader integration of deep learning into clinical neuroscience, including inefficiencies, lack of comprehensive evaluation platforms, the underexplored role of acquisition protocols, limited explainability, and fragmented approaches to data disruptions. This thesis provides novel solutions that transform deep learning from a theoretical promise to a practical reality in diffusion MRI analysis, with potential for significant clinical impact. The proposed methods advances the state of the art in deep learning-driven diffusion MRI analysis, but also provide a solid foundation for future research to further integrate deep learning into clinical practice.

6.2 Future Outlook

Looking forward, several exciting directions for future research and method development emerge from this work. For example, the integration of geometric constraints, such as those derived from spherical harmonics, is of high interest for enhancing FODs. The incorporation of latent diffusion models (LDM) is another promising direction, as these models are gaining popularity in image synthesis tasks. Given the high-dimensional nature of dMRI data, compressing this data into a latent space holds significant potential for improving deep learning models. By incorporating geometry priors, this compression could be made more meaningful, facilitating better performance in downstream tasks.

Moreover, our use of Vector Quantized Variational Autoencoders (VQ-VAE) in the UFREE demonstrated its utility in feature matching across different datasets. However, there are many other challenges in dMRI enhancement beyond just resolution improvement. One such challenge is the complexity and time-consuming nature of DWI preprocessing, which involves crucial steps like image registration, motion correction, eddy-current distortion, susceptibility correction, and spatial normalization. Deep learning-based methods could potentially streamline these processes, improving both efficiency and accuracy.

Additionally, the application of large language models (LLMs) and multimodal integration in clinical neuroscience presents an exciting opportunity. By combining text-based commands with image generation capabilities, clinicians could interact with deep learning models in a more intuitive manner, querying the system in natural language and receiving enhanced dMRI results. Specialized foundation models tailored for FODs could become a key area of focus, advancing computational neuroscience by providing more accurate, robust, and scalable tools for analyzing dMRI data. Another important challenge in clinical practice is the harmonization of data from different MRI acquisition protocols and scanners. Variations in scanner types, protocols, and clinical settings can lead to inhomogeneous datasets, complicating analysis. A foundation model, built with domain adaptation techniques, could be used to harmonize these datasets, making it easier to draw consistent and reliable conclusions across diverse clinical environments.

The integration of deep learning techniques into clinical neuroscience has the potential to transform the field, and the work presented in this thesis provides a strong foundation for future developments. By addressing key technical and clinical challenges, we are positioning deep learning to make a significant impact on diffusion MRI analysis, with the potential to greatly enhance clinical decision-making and patient outcomes.

Bibliography

[1] M. J. Ehrhardt, K. Thielemans, L. Pizarro, D. Atkinson, S. Ourselin, B. F. Hutton and S. R. Arridge. 'Joint reconstruction of PET-MRI by exploiting structural similarity'. In: *Inverse Problems* 31.1 (2014), p. 015001.

[2] D. McRobbie, E. Moore, M. Graves and M. Prince. *MRI From Picture to Proton*. 2008.

[3] R. Smith-Bindman, D. L. Miglioretti, E. Johnson, C. Lee, H. S. Feigelson, M. Flynn, R. T. Greenlee, R. L. Kruger, M. C. Hornbrook, D. Roblin et al. 'Use of diagnostic imaging studies and associated radiation exposure for patients enrolled in large integrated health care systems, 1996-2010'. In: *Jama* 307.22 (2012), pp. 2400–2409.

[4] O. Indicators and O. Hagvísar. *Health at a glance 2019: OECD indicators*. Paris: OECD Publishing, 2019.

[5] K.-D. Ebel and G. Benz-Bohm. 'Differential diagnosis in pediatric radiology'. In: *(No Title)* (1999).

[6] W. Bradley, M. Brant-Zawadzki and J. Cambray-Forker. *MRI of the Brain*. LWW MRI teaching file series v. 2. Lippincott Williams & Wilkins, 2001. ISBN: 9780781725682. URL: https://books.google.com.hk/books?id=40f2WPdivA8C.

[7] M. Clinic. *Magnetic Resonance Imaging (MRI)*. Accessed: 2023-12-22. n.d. URL: https://www.mayoclinic.org/tests-procedures/mri/about/pac-20384768.

[8] B. Heim, F. Krismer, R. De Marzi and K. Seppi. 'Magnetic resonance imaging for the diagnosis of Parkinson's disease'. In: *Journal of Neural Transmission* 124.8 (2017), pp. 915–964. ISSN: 1435-1463. DOI: 10.1007/s00702-017-1717-8.

[9] G. B. Frisoni, N. C. Fox, C. R. Jack, P. Scheltens and P. M. Thompson. 'The clinical use of structural MRI in Alzheimer disease'. In: *Nature Reviews Neurology* 6.2 (2010), pp. 67–77. ISSN: 1759-4766. DOI: 10.1038/nrneurol.2009.215.

[10] A. Segato, A. Marzullo, F. Calimeri and E. De Momi. 'Artificial intelligence for brain diseases: A systematic review'. In: *APL Bioengineering* 4.4 (2020), p. 041503. ISSN: 2473-2877. DOI: 10.1063/5.0011697.

[11] Sea438. *Spin Orientations During Relaxation*. Accessed: 2025-03-11. 2023. URL: https://commons.wikimedia.org/wiki/File:Spin_Orientations_During_Relaxation.jpg.

[12] Case Western Reserve University. *Description of the specific image*. Accessed: 2025-03-11. 2025. URL: https://case.edu/med/neurology/NR/MRI%20Basics.htm.

[13] G. V. Hirsch, C. M. Bauer and L. B. Merabet. 'Using structural and functional brain imaging to uncover how the brain adapts to blindness'. In: *Annals of neuroscience and psychology* 2 (2015), p. 5.

[14] B. A. Holland, D. K. Haas, D. Norman, M. Brant-Zawadzki and T. H. Newton. 'MRI of normal brain maturation.' In: *American Journal of Neuroradiology* 7.2 (1986), pp. 201–208.

[15] A. C. Evans, B. D. C. Group et al. 'The NIH MRI study of normal brain development'. In: *Neuroimage* 30.1 (2006), pp. 184–202.

[16] A. Gallagher, C. Bulteau, D. Cohen and J. L. Michaud. *Neurocognitive Development: Disorders and Disabilities*. Vol. 174. Elsevier, 2020.

[17] H. West, J. L. Leach, B. V. Jones, M. Care, R. Radhakrishnan, A. C. Merrow, E. Alvarado and S. D. Serai. 'Clinical validation of synthetic brain MRI in children: initial experience'. In: *Neuroradiology* 59 (2017), pp. 43–50.

[18] DMRITool. *Single-shell sampling scheme illustration*. Accessed: 2025-03-11. 2025. URL: https://diffusionmritool.github.io/tutorial_qspacesampling.html.

[19] V. Baliyan, C. J. Das, R. Sharma and A. K. Gupta. 'Diffusion weighted imaging: technique and applications'. In: *World journal of radiology* 8.9 (2016), p. 785.

[20] C.-H. Yeh, R. E. Smith, T. Dhollander, F. Calamante and A. Connelly. 'Connectomes from streamlines tractography: Assigning streamlines to brain parcellations is not trivial but highly consequential'. In: *Neuroimage* 199 (2019), pp. 160–171.

[21] M. E. Moseley, Y. Cohen, J. Kucharczyk, J. Mintorovitch, H. Asgari, M. Wendland, J. Tsuruda and D. Norman. 'Diffusion-weighted MR imaging of anisotropic water diffusion in cat central nervous system.' In: *Radiology* 176.2 (1990), pp. 439–445.

[22] R. Turner, D. Le Bihan, J. Maier, R. Vavrek, L. K. Hedges and J. Pekar. 'Echo-planar imaging of intravoxel incoherent motion.' In: *Radiology* 177.2 (1990), pp. 407–414.

[23] T. Chenevert, J. Brunberg and J. Pipe. 'Anisotropic diffusion within human white matter'. In: *Seventy sixth scientific assembly and annual meeting of the Radiological Society of North America*. 1990.

[24] P. Douek, R. Turner, J. Pekar, N. Patronas and D. Le Bihan. 'MR color mapping of myelin fiber orientation'. In: *J Comput Assist Tomogr* 15.6 (1991), pp. 923–929.

[25] G. Yang, Q. Tian, C. Leuze, M. Wintermark and J. A. McNab. 'Double diffusion encoding MRI for the clinic'. In: *Magnetic resonance in medicine* 80.2 (2018), pp. 507–520.

[26] C. Martín-Martín, Á. Planchuelo-Gómez, Á. L. Guerrero, D. García-Azorín, A. Tristán-Vega, R. de Luis-García and S. Aja-Fernández. 'Viability of AMURA biomarkers from single-shell diffusion MRI in clinical studies'. In: *Frontiers in Neuroscience* 17 (2023), p. 1106350.

[27] B. Jeurissen, J.-D. Tournier, T. Dhollander, A. Connelly and J. Sijbers. 'Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data'. In: *NeuroImage* 103 (2014), pp. 411–426.

[28] J. N. Morelli, M. R. Saettele, R. A. Rangaswamy, L. Vu, C. M. Gerdes, W. Zhang and F. Ai. 'Echo planar diffusion-weighted imaging: possibilities and considerations with 12-and 32-channel head coils'. In: *Journal of clinical imaging science* 2 (2012), p. 31.

[29] C. M. Tax, M. Bastiani, J. Veraart, E. Garyfallidis and M. O. Irfanoglu. 'What's new and what's next in diffusion MRI preprocessing'. In: *NeuroImage* 249 (2022), p. 118830.

[30] C. Pierpaoli. 'Artifacts in diffusion MRI'. In: *Diffusion MRI: theory, methods and applications* (2010), pp. 303–318.

[31] C. M. Tax, S. B. Vos and A. Leemans. 'Checking and correcting DTI data'. In: *Diffusion Tensor Imaging: A Practical Handbook* (2016), pp. 127–150.

[32] M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich and S. M. Smith. 'FSL'. In: *Neuroimage* 62.2 (2012), pp. 782–790.

[33] V. Popescu, M. Battaglini, W. Hoogstrate, S. C. Verfaillie, I. Sluimer, R. A. van Schijndel, B. W. van Dijk, K. S. Cover, D. L. Knol, M. Jenkinson et al. 'Optimizing parameter choice for FSL-Brain Extraction Tool (BET) on 3D T1 images in multiple sclerosis'. In: *Neuroimage* 61.4 (2012), pp. 1484–1494.

[34] L. I. Rudin, S. Osher and E. Fatemi. 'Nonlinear total variation based noise removal algorithms'. In: *Physica D: nonlinear phenomena* 60.1-4 (1992), pp. 259–268.

[35] J. W. Gibbs. 'Fourier's series'. In: *Nature* 59.1522 (1898), pp. 200–200.

[36] D. K. Jones. 'The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: a Monte Carlo study'. In: *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* 51.4 (2004), pp. 807–815.

[37] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich and J. C. Gee. 'N4ITK: improved N3 bias correction'. In: *IEEE Transactions on Medical Imaging* 29.6 (2010), pp. 1310–1320.

[38] P. Jezzard and R. S. Balaban. 'Correction for geometric distortion in echo planar images from B0 field variations'. In: *Magnetic resonance in medicine* 34.1 (1995), pp. 65–73.

[39] C. B. Hansen, V. Nath, A. E. Hainline, K. G. Schilling, P. Parvathaneni, R. G. Bayrak, J. A. Blaber, O. Irfanoglu, C. Pierpaoli, A. W. Anderson et al. 'Characterization and correlation of signal drift in diffusion weighted MRI'. In: *Magnetic resonance imaging* 57 (2019), pp. 133–142.

[40] A. Leemans and D. K. Jones. 'The B-matrix must be rotated when correcting for subject motion in DTI data'. In: *Magnetic Resonance in Medicine: An Official Journal of*

the International Society for Magnetic Resonance in Medicine 61.6 (2009), pp. 1336–1349.

[41] D. Le Bihan, J.-F. Mangin, C. Poupon, C. A. Clark, S. Pappata, N. Molko and H. Chabriat. 'Diffusion tensor imaging: concepts and applications'. In: *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine* 13.4 (2001), pp. 534–546.

[42] J.-D. Tournier, F. Calamante and A. Connelly. 'Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution'. In: *Neuroimage* 35.4 (2007), pp. 1459–1472.

[43] P. J. Basser, J. Mattiello and D. LeBihan. 'MR diffusion tensor spectroscopy and imaging'. In: *Biophysical journal* 66.1 (1994), pp. 259–267.

[44] J. M. Soares, P. Marques, V. Alves and N. Sousa. 'A hitchhiker's guide to diffusion tensor imaging'. In: *Frontiers in neuroscience* 7 (2013), p. 31.

[45] P. J. Basser. 'Inferring microstructural features and the physiological state of tissues from diffusion-weighted images'. In: *NMR in Biomedicine* 8.7 (1995), pp. 333–344.

[46] H. Ni, V. Kavcic, T. Zhu, S. Ekholm and J. Zhong. 'Effects of number of diffusion gradient directions on derived diffusion tensor imaging indices in human brain'. In: *American Journal of Neuroradiology* 27.8 (2006), pp. 1776–1781.

[47] H. M. Feldman, J. D. Yeatman, E. S. Lee, L. H. Barde and S. Gaman-Bean. 'Diffusion tensor imaging: a review for pediatric researchers and clinicians'. In: *Journal of Developmental & Behavioral Pediatrics* 31.4 (2010), pp. 346–356.

[48] B. J. Jellison, A. S. Field, J. Medow, M. Lazar, M. S. Salamat and A. L. Alexander. 'Diffusion tensor imaging of cerebral white matter: a pictorial review of physics, fiber tract anatomy, and tumor imaging patterns'. In: *American Journal of Neuroradiology* 25.3 (2004), pp. 356–369.

[49] W.-S. Tae, B.-J. Ham, S.-B. Pyun, S.-H. Kang and B.-J. Kim. 'Current clinical applications of diffusion-tensor imaging in neurological disorders'. In: *Journal of clinical neurology (Seoul, Korea)* 14.2 (2018), p. 129.

[50] P. J. Basser and C. Pierpaoli. 'Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI'. In: *Journal of magnetic resonance* 213.2 (2011), pp. 560–570.

[51] Y. Assaf and O. Pasternak. 'Diffusion tensor imaging (DTI)-based white matter mapping in brain research: a review'. In: *Journal of molecular neuroscience* 34 (2008), pp. 51–61.

[52] E. Sbardella, F. Tona, N. Petsas and P. Pantano. 'DTI measurements in multiple sclerosis: evaluation of brain damage and clinical implications'. In: *Multiple sclerosis international* 2013.1 (2013), p. 671730.

[53] O. Ciccarelli, D. Werring, C. Wheeler–Kingshott, G. Barker, G. Parker, A. Thompson and D. Miller. 'Investigation of MS normal-appearing brain using diffusion tensor MRI with clinical correlations'. In: *Neurology* 56.7 (2001), pp. 926–933.

[54] J.-D. Tournier, F. Calamante, D. G. Gadian and A. Connelly. 'Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution'. In: *Neuroimage* 23.3 (2004), pp. 1176–1185.

[55] J. D. Tournier, F. Calamante, A. Connelly et al. 'Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions'. In: *Proceedings of the international society for magnetic resonance in medicine*. Vol. 1670. Wiley Online Library. 2010.

[56] D. M. Healy Jr, H. Hendriks and P. T. Kim. 'Spherical deconvolution'. In: *Journal of Multivariate Analysis* 67.1 (1998), pp. 1–22.

[57] D. Raffelt, J.-D. Tournier, S. Rose, G. R. Ridgway, R. Henderson, S. Crozier, O. Salvado and A. Connelly. 'Apparent fibre density: a novel measure for the analysis of diffusion-weighted magnetic resonance images'. In: *Neuroimage* 59.4 (2012), pp. 3976–3994.

[58] R. E. Smith, J.-D. Tournier, F. Calamante and A. Connelly. 'SIFT: Spherical-deconvolution informed filtering of tractograms'. In: *Neuroimage* 67 (2013), pp. 298–312.

[59] D. A. Raffelt, J.-D. Tournier, R. E. Smith, D. N. Vaughan, G. Jackson, G. R. Ridgway and A. Connelly. 'Investigating white matter fibre density and morphology using fixel-based analysis'. In: *Neuroimage* 144 (2017), pp. 58–73.

[60] S. Genc, C. M. Tax, E. P. Raven, M. Chamberland, G. D. Parker and D. K. Jones. 'Impact of b-value on estimates of apparent fibre density'. In: *Human brain mapping* 41.10 (2020), pp. 2583–2595.

[61] I. Benou and T. Riklin Raviv. 'Deeptract: A probabilistic deep learning framework for white matter fiber tractography'. In: *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part III 22*. Springer. 2019, pp. 626–635.

[62] F. Zhang, A. Daducci, Y. He, S. Schiavi, C. Seguin, R. E. Smith, C.-H. Yeh, T. Zhao and L. J. O'Donnell. 'Quantitative mapping of the brain's structural connectivity using diffusion MRI tractography: A review'. In: *Neuroimage* 249 (2022), p. 118870.

[63] J.-D. Tournier, R. Smith, D. Raffelt, R. Tabbara, T. Dhollander, M. Pietsch, D. Christiaens, B. Jeurissen, C.-H. Yeh and A. Connelly. 'MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation'. In: *Neuroimage* 202 (2019), p. 116137.

[64] R. E. Smith, J.-D. Tournier, F. Calamante and A. Connelly. 'Anatomically-constrained tractography: improved diffusion MRI streamlines tractography through effective use of anatomical information'. In: *Neuroimage* 62.3 (2012), pp. 1924–1938.

[65] P. J. Basser, S. Pajevic, C. Pierpaoli, J. Duda and A. Aldroubi. 'In vivo fiber tractography using DT-MRI data'. In: *Magnetic resonance in medicine* 44.4 (2000), pp. 625–632.

[66] S. Mori, B. J. Crain, V. P. Chacko and P. C. Van Zijl. 'Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging'. In: *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society* 45.2 (1999), pp. 265–269.

[67] J. D. Tournier, F. Calamant, D. Gadian and A. Connelly. 'A novel fiber-tracking technique: front evolution using a fiber orientation probability density function'. In: *Proc. of the International Society of Magnetic Resonance in Medicine*. 2002.

[68] B. Jeurissen, A. Leemans, J.-D. Tournier, D. K. Jones and J. Sijbers. 'Estimating the number of fiber orientations in diffusion MRI voxels: a constrained spherical

deconvolution study'. In: *International Society for Magnetic Resonance in Medicine (ISMRM)*. Vol. 573. 2010.

[69] R. E. Smith, J.-D. Tournier, F. Calamante and A. Connelly. 'SIFT2: Enabling dense quantitative assessment of brain white matter connectivity using streamlines tractography'. In: *Neuroimage* 119 (2015), pp. 338–351.

[70] R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman et al. 'An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest'. In: *Neuroimage* 31.3 (2006), pp. 968–980.

[71] R. E. Smith, J.-D. Tournier, F. Calamante and A. Connelly. 'The effects of SIFT on the reproducibility and biological accuracy of the structural connectome'. In: *Neuroimage* 104 (2015), pp. 253–265.

[72] C. Guillemot and O. Le Meur. 'Image inpainting: Overview and recent advances'. In: *IEEE signal processing magazine* 31.1 (2013), pp. 127–144.

[73] G. Liu, F. A. Reda, K. J. Shih, T.-C. Wang, A. Tao and B. Catanzaro. 'Image inpainting for irregular holes using partial convolutions'. In: *Proceedings of the European conference on computer vision (ECCV)*. 2018, pp. 85–100.

[74] C. Chen, X. Shi, Y. Qin, X. Li, X. Han, T. Yang and S. Guo. 'Real-world blind super-resolution via feature matching with implicit high-resolution priors'. In: *ACM Multimedia*. 2022, pp. 1329–1338.

[75] C. R. Turner, A. Fuggetta, L. Lavazza and A. L. Wolf. 'A conceptual basis for feature engineering'. In: *Journal of Systems and Software* 49.1 (1999), pp. 3–15.

[76] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard and L. D. Jackel. 'Backpropagation applied to handwritten zip code recognition'. In: *Neural computation* 1.4 (1989), pp. 541–551.

[77] J. Sanders and E. Kandrot. *CUDA by example: an introduction to general-purpose GPU programming*. Addison-Wesley Professional, 2010.

[78] A. Krizhevsky, I. Sutskever and G. E. Hinton. 'Imagenet classification with deep convolutional neural networks'. In: *Advances in neural information processing systems* 25 (2012).

[79] V. Nair and G. E. Hinton. 'Rectified linear units improve restricted boltzmann machines'. In: *Proceedings of the 27th international conference on machine learning (ICML-10)*. 2010, pp. 807–814.

[80] G. E. Dahl, T. N. Sainath and G. E. Hinton. 'Improving deep neural networks for LVCSR using rectified linear units and dropout'. In: *2013 IEEE international conference on acoustics, speech and signal processing*. IEEE. 2013, pp. 8609–8613.

[81] K. Simonyan and A. Zisserman. 'Very deep convolutional networks for large-scale image recognition'. In: *arXiv preprint arXiv:1409.1556* (2014).

[82] O. Ronneberger, P. Fischer and T. Brox. 'U-net: Convolutional networks for biomedical image segmentation'. In: *MICCAI*. Springer. 2015, pp. 234–241.

[83] J. Lozej, B. Meden, V. Struc and P. Peer. 'End-to-end iris segmentation using u-net'. In: *2018 IEEE international work conference on bioinspired intelligence (IWOBI)*. IEEE. 2018, pp. 1–6.

[84] X. Yang, X. Li, Y. Ye, R. Y. Lau, X. Zhang and X. Huang. 'Road detection and centerline extraction via deep recurrent convolutional neural network U-Net'. In: *IEEE Transactions on Geoscience and Remote Sensing* 57.9 (2019), pp. 7209–7220.

[85] X. Ma, X. Deng, L. Qi, Y. Jiang, H. Li, Y. Wang and X. Xing. 'Fully convolutional network for rice seedling and weed image segmentation at the seedling stage in paddy fields'. In: *PloS one* 14.4 (2019), e0215676.

[86] T. Xiang, C. Zhang, D. Liu, Y. Song, H. Huang and W. Cai. 'BiO-Net: learning recurrent bi-directional connections for encoder-decoder architecture'. In: *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part I 23*. Springer. 2020, pp. 74–84.

[87] X. Wang, T. Xiang, C. Zhang, Y. Song, D. Liu, H. Huang and W. Cai. 'BiX-NAS: Searching efficient bi-directional architecture for medical image segmentation'. In: *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part I 24*. Springer. 2021, pp. 229–238.

[88] R. Rombach, A. Blattmann, D. Lorenz, P. Esser and B. Ommer. 'High-resolution image synthesis with latent diffusion models'. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022, pp. 10684–10695.

[89] A. Radford, L. Metz and S. Chintala. 'Unsupervised representation learning with deep convolutional generative adversarial networks'. In: *arXiv preprint arXiv:1511.06434* (2015).

[90] P. Isola, J.-Y. Zhu, T. Zhou and A. A. Efros. 'Image-to-image translation with conditional adversarial networks'. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 1125–1134.

[91] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu and T. S. Huang. 'Free-form image inpainting with gated convolution'. In: *ICCV*. 2019, pp. 4471–4480.

[92] M. Battaglini, M. Jenkinson and N. De Stefano. 'Evaluating and reducing the impact of white matter lesions on brain volume measurements'. In: *Human brain mapping* 33.9 (2012), pp. 2062–2071.

[93] G. K. Rohde, A. Barnett, P. Basser, S. Marenco and C. Pierpaoli. 'Comprehensive approach for correction of motion and distortion in diffusion-weighted MRI'. In: *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* 51.1 (2004), pp. 103–114.

[94] Z. Tang, X. Wang, L. Zhu, M. Cabezas, D. Liu, M. Barnett, W. Cai and C. Wang. 'TW-BAG: Tensor-wise Brain-aware Gate Network for Inpainting Disrupted Diffusion Tensor Imaging'. In: *Proceedings of the International Conference on Digital Image Computing: Techniques and Applications (DICTA)*. IEEE. 2022, pp. 1–8.

[95] Z. Tang, X. Wang, M. Cabezas, A. D'Souza, F. Calamante, D. Liu, M. Barnett, C. Wang and W. Cai. 'Diffusion MRI Fibre Orientation Distribution Inpainting'. In: *Computational Diffusion MRI*. Springer. 2022, pp. 65–76.

[96] B. Scherrer, A. Gholipour and S. K. Warfield. 'Super-resolution in diffusion-weighted imaging'. In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011: 14th International Conference, Toronto, Canada, September 18-22, 2011, Proceedings, Part II 14*. Springer. 2011, pp. 124–132.

[97] B. Scherrer, A. Gholipour and S. K. Warfield. 'Super-resolution reconstruction to increase the spatial resolution of diffusion weighted images from orthogonal anisotropic acquisitions'. In: *Medical image analysis* 16.7 (2012), pp. 1465–1476.

[98] Z. Lin, T. Gong, K. Wang, Z. Li, H. He, Q. Tong, F. Yu and J. Zhong. 'Fast learning of fiber orientation distribution function for MR tractography using convolutional neural network'. In: *Medical physics* 46.7 (2019), pp. 3101–3116.

[99] M. Lyon, P. Armitage and M. A. Álvarez. 'Angular super-resolution in diffusion MRI with a 3D recurrent convolutional autoencoder'. In: *International Conference on Medical Imaging with Deep Learning*. PMLR. 2022, pp. 834–846.

[100] M. Lyon, P. Armitage and M. A. Álvarez. 'Spatio-Angular Convolutions for Super-resolution in Diffusion MRI'. In: *arXiv preprint arXiv:2306.00854* (2023).

[101] S. Chen, Z. Tang, M. Cabezas, X. Wang, A. D'Souza, M. Barnett, F. Calamante, W. Cai and C. Wang. 'Enhancing Angular Resolution via Directionality Encoding and Geometric Constraints in Brain Diffusion Tensor Imaging'. In: *arXiv e-prints* (2024), arXiv–2409.

[102] S. Chen, Z. Tang, X. Wang, C. Wang and W. Cai. 'Unsupervised dMRI Artifact Detection via Angular Resolution Enhancement and Cycle Consistency Learning'. In: *Australasian Joint Conference on Artificial Intelligence*. Springer. 2024, pp. 417–431.

[103] X. Wang, Z. Liu, S. Chen, C. Wang, M. Barnett, W. Cai and Z. Tang. 'SdADC: Semi-supervised diffusion MRI artefact detection and classification'. In: *Annual Meeting of the International Society for Magnetic Resonance in Medicine (ISMRM).* 2025.

[104] R. Zeng, J. Lv, H. Wang, L. Zhou, M. Barnett, F. Calamante and C. Wang. 'FOD-Net: A deep learning method for fiber orientation distribution angular super resolution'. In: *Medical Image Analysis* 79 (2022), p. 102431.

[105] O. Lucena, S. B. Vos, V. Vakharia, J. Duncan, K. Ashkan, R. Sparks and S. Ourselin. 'Enhancing the estimation of fiber orientation distributions using convolutional neural networks'. In: *Computers in Biology and Medicine* 135 (2021), p. 104643.

[106] A. Rauland and D. Merhof. 'Using Synthetic Training Data in Neural Networks for the Estimation of Fiber Orientation Distribution Functions from Single Shell Data'.

In: *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*. IEEE. 2023, pp. 1–5.

[107] M. O. D. Silva, C. P. Santana, D. S. D. Carmo and L. Rittner. 'FOD-Swin-Net: angular super resolution of fiber orientation distribution using a transformer-based deep model'. In: *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*. IEEE. 2024, pp. 1–5.

[108] Z. Tang, X. Wang, M. Cabezas, A. D'Souza, M. Barnett, F. Calamante, W. Cai and C. Wang. 'Fibre Population-guided Pre-training for 3D Spatial Super-Resolution on Multimodal Brain Diffusion MR Imaging'. In: *ACM Multimedia Asia 2024*. 2024.

[109] R. Zeng, J. Lv, H. Wang, L. Zhou, M. Barnett, F. Calamante and C. Wang. 'FOD-Net: A deep learning method for fiber orientation distribution angular super resolution'. In: *Medical Image Analysis* 79 (2022), p. 102431.

[110] D. C. Alexander, D. Zikic, A. Ghosh, R. Tanno, V. Wottschel, J. Zhang, E. Kaden, T. B. Dyrby, S. N. Sotiropoulos, H. Zhang et al. 'Image quality transfer and applications in diffusion MRI'. In: *NeuroImage* 152 (2017), pp. 283–298.

[111] R. Tanno, D. E. Worrall, E. Kaden, A. Ghosh, F. Grussu, A. Bizzi, S. N. Sotiropoulos, A. Criminisi and D. C. Alexander. 'Uncertainty modelling in deep learning for safer neuroimage enhancement: Demonstration in diffusion MRI'. In: *NeuroImage* 225 (2021), p. 117366.

[112] F. Dell'Acqua and J.-D. Tournier. 'Modelling white matter with spherical deconvolution: How and why?' In: *NMR in Biomedicine* 32.4 (2019), e3945.

[113] S. Farquharson, J.-D. Tournier, F. Calamante, G. Fabinyi, M. Schneider-Kolsky, G. D. Jackson and A. Connelly. 'White matter fiber tractography: why we need to move beyond DTI'. In: *Journal of neurosurgery* 118.6 (2013), pp. 1367–1377.

[114] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. J. Behrens, E. Yacoub and K. Ugurbil. 'The WU-Minn human connectome project: an overview'. In: *Neuroimage* 80 (2013), pp. 62–79.

[115] X. Wang, Z. Tang, M. Cabezas, A. D'Souza, F. Calamante, D. Liu, M. Barnett, S. Tu, W. Cai and C. Wang. 'FOD-Net 2.0: End-to-end FOD enhancement for low angular

diffusion acquisitions using deep learning'. In: *Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM)*. 2023.

[116] J. D. Tournier, F. Calamante and A. Connelly. 'MRtrix: diffusion tractography in crossing fiber regions'. In: *International journal of imaging systems and technology* 22.1 (2012), pp. 53–66.

[117] T. Dhollander, D. Raffelt and A. Connelly. 'Unsupervised 3-tissue response function estimation from single-shell or multi-shell diffusion MR data without a co-registered T1 image'. In: *ISMRM workshop on breaking the barriers of diffusion MRI*. Vol. 5. 5. Lisbon, Portugal. 2016.

[118] K. Breuer, M. Stommel and W. Korte. 'Analysis and evaluation of fiber orientation reconstruction methods'. In: *Journal of Composites Science* 3.3 (2019), p. 67.

[119] J. Hu, L. Shen and G. Sun. 'Squeeze-and-excitation networks'. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2018, pp. 7132–7141.

[120] J.-D. Tournier, S. Mori and A. Leemans. 'Diffusion tensor imaging and beyond'. In: *Magnetic resonance in medicine* 65.6 (2011), p. 1532.

[121] F. Calamante. 'The seven deadly sins of measuring brain structural connectivity using diffusion MRI streamlines fibre-tracking'. In: *Diagnostics* 9.3 (2019), p. 115.

[122] S. G. Mueller, M. W. Weiner, L. J. Thal, R. C. Petersen, C. R. Jack, W. Jagust, J. Q. Trojanowski, A. W. Toga and L. Beckett. 'Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI)'. In: *Alzheimer's & Dementia* 1.1 (2005), pp. 55–66.

[123] K. Marek, D. Jennings, S. Lasch, A. Siderowf, C. Tanner, T. Simuni, C. Coffey, K. Kieburtz, E. Flagg, S. Chowdhury et al. 'The Parkinson progression marker initiative (PPMI)'. In: *Progress in neurobiology* 95.4 (2011), pp. 629–635.

[124] J. J. Bartlett, C. E. Davey, L. A. Johnston and J. Duan. 'Recovering high-quality fiber orientation distributions from a reduced number of diffusion-weighted images using a model-driven deep learning architecture'. In: *Magnetic Resonance in Medicine* (2024).

[125] T. Dhollander, A. Clemente, M. Singh, F. Boonstra, O. Civier, J. D. Duque, N. Egorova, P. Enticott, I. Fuelscher, S. Gajamange et al. 'Fixel-based analysis of diffusion MRI:

methods, applications, challenges and opportunities'. In: *Neuroimage* 241 (2021), p. 118417.

[126] O. Civier, R. E. Smith, C.-H. Yeh, A. Connelly and F. Calamante. 'Is removal of weak connections necessary for graph-theoretical analysis of dense weighted structural connectomes from diffusion MRI?' In: *NeuroImage* 194 (2019), pp. 68–81.

[127] K. Kyle, J. Maller, Y. Barnett, B. Jonker, M. Barnett, A. D'Souza, F. Calamante, J. Maamary, J. Peters, C. Wang et al. 'Tremor suppression following treatment with MRgFUS: skull density ratio consistency and degree of posterior dentatorubrothalamic tract lesioning predicts long-term clinical outcomes in essential tremor'. In: *Frontiers in Neurology* 14 (2023), p. 1129430.

[128] S. N. Sotiropoulos, S. Jbabdi, J. Xu, J. L. Andersson, S. Moeller, E. J. Auerbach, M. F. Glasser, M. Hernandez, G. Sapiro, M. Jenkinson et al. 'Advances in diffusion MRI acquisition and processing in the Human Connectome Project'. In: *Neuroimage* 80 (2013), pp. 125–143.

[129] J. D. Power. 'A simple but useful way to assess fMRI scan qualities'. In: *Neuroimage* 154 (2017), pp. 150–158.

[130] E. Kellner, B. Dhital, V. G. Kiselev and M. Reisert. 'Gibbs-ringing artifact removal based on local subvoxel-shifts'. In: *Magnetic resonance in medicine* 76.5 (2016), pp. 1574–1581.

[131] B. Fischl. 'FreeSurfer'. In: *Neuroimage* 62.2 (2012), pp. 774–781.

[132] R. Smith, A. Skoch, C. J. Bajada, S. Caspers and A. Connelly. 'Hybrid surface-volume segmentation for improved anatomically-constrained tractography'. In: *OHBM*. 2020.

[133] R. W. Kennard and L. A. Stone. 'Computer aided design of experiments'. In: *Technometrics* 11.1 (1969), pp. 137–148.

[134] T. Dhollander and A. Connelly. 'A novel iterative approach to reap the benefits of multi-tissue CSD from just single-shell (+ b= 0) diffusion MRI data'. In: *Proceedings of the international society for magnetic resonance in medicine*. Vol. 24. Wiley Online Library. 2016, p. 3010.

[135] J. Wasserthal, P. Neher and K. H. Maier-Hein. 'TractSeg-Fast and accurate white matter tract segmentation'. In: *NeuroImage* 183 (2018), pp. 239–253.

[136] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox and O. Ronneberger. '3D U-Net: learning dense volumetric segmentation from sparse annotation'. In: *MICCAI*. Springer. 2016, pp. 424–432.

[137] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen and K. H. Maier-Hein. 'nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation'. In: *Nature methods* 18.2 (2021), pp. 203–211.

[138] T. Dhollander, R. Tabbara, J. Rosnarho-Tornstrand, J.-D. Tournier, D. Raffelt and A. Connelly. 'Multi-tissue log-domain intensity and inhomogeneity normalisation for quantitative apparent fibre density'. In: *Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM)*. Vol. 29. 2021, p. 2472.

[139] A. W. Anderson. 'Measurement of fiber orientation distributions using high angular resolution diffusion imaging'. In: *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* 54.5 (2005), pp. 1194–1206.

[140] M. G. Kendall. 'A new measure of rank correlation'. In: *Biometrika* 30.1/2 (1938), pp. 81–93.

[141] B. Zhang and S. Horvath. 'A general framework for weighted gene co-expression network analysis'. In: *Statistical applications in genetics and molecular biology* 4.1 (2005).

[142] A. Fornito, A. Zalesky and E. Bullmore. *Fundamentals of brain network analysis*. Academic press, 2016.

[143] M. Rubinov and O. Sporns. 'Complex network measures of brain connectivity: uses and interpretations'. In: *Neuroimage* 52.3 (2010), pp. 1059–1069.

[144] D. A. Raffelt, R. E. Smith, G. R. Ridgway, J.-D. Tournier, D. N. Vaughan, S. Rose, R. Henderson and A. Connelly. 'Connectivity-based fixel enhancement: Whole-brain statistical analysis of diffusion MRI measures in the presence of crossing fibres'. In: *Neuroimage* 117 (2015), pp. 40–55.

[145] M. Grossman, C. Anderson, A. Khan, B. Avants, L. Elman and L. McCluskey. 'Impaired action knowledge in amyotrophic lateral sclerosis'. In: *Neurology* 71.18 (2008), pp. 1396–1401.

[146] E. Kasper, C. Schuster, J. Machts, J. Kaufmann, D. Bittner, S. Vielhaber, R. Benecke, S. Teipel and J. Prudlo. 'Microstructural white matter changes underlying cognitive and behavioural impairment in ALS–an in vivo study using DTI'. In: *PloS one* 9.12 (2014), e114543.

[147] D. J. Libon, C. McMillan, B. Avants, A. Boller, B. Morgan, L. Burkholder, K. Chandrasekaran, L. Elman, L. McCluskey and M. Grossman. 'Deficits in concept formation in amyotrophic lateral sclerosis.' In: *Neuropsychology* 26.4 (2012), p. 422.

[148] L. D. Pettit, M. E. Bastin, C. Smith, T. H. Bak, T. H. Gillingwater and S. Abrahams. 'Executive deficits, not processing speed relates to abnormalities in distinct prefrontal tracts in amyotrophic lateral sclerosis'. In: *Brain* 136.11 (2013), pp. 3290–3304.

[149] D. Bäumer, K. Talbot and M. R. Turner. 'Advances in motor neurone disease'. In: *Journal of the Royal Society of Medicine* 107.1 (2014), pp. 14–21.

[150] S. Rose, K. Pannek, C. Bell, F. Baumann, N. Hutchinson, A. Coulthard, P. McCombe and R. Henderson. 'Direct evidence of intra-and interhemispheric corticomotor network degeneration in amyotrophic lateral sclerosis: an automated MRI structural connectivity study'. In: *Neuroimage* 59.3 (2012), pp. 2661–2669.

[151] A. Fornito, A. Zalesky, C. Pantelis and E. T. Bullmore. 'Schizophrenia, neuroimaging and connectomics'. In: *Neuroimage* 62.4 (2012), pp. 2296–2314.

[152] A. Fornito, A. Zalesky and M. Breakspear. 'The connectomics of brain disorders'. In: *Nature Reviews Neuroscience* 16.3 (2015), pp. 159–172.

[153] C.-H. Yeh, D. K. Jones, X. Liang, M. Descoteaux and A. Connelly. 'Mapping structural connectivity using diffusion MRI: Challenges and opportunities'. In: *Journal of Magnetic Resonance Imaging* 53.6 (2021), pp. 1666–1682.

[154] E. Abad, J. Sepulcre, E. H. Martinez-Lapiscina, I. Zubizarreta, J. Garcia-Ojalvo and P. Villoslada. 'The analysis of semantic networks in multiple sclerosis identifies preferential damage of long-range connectivity'. In: *Multiple sclerosis and related disorders* 4.5 (2015), pp. 387–394.

[155] E. Pagani, M. A. Rocca, E. De Meo, M. A. Horsfield, B. Colombo, M. Rodegher, G. Comi and M. Filippi. 'Structural connectivity in multiple sclerosis and modeling of disconnection'. In: *Multiple Sclerosis Journal* 26.2 (2020), pp. 220–232.

[156] N. Shu, Y. Liu, K. Li, Y. Duan, J. Wang, C. Yu, H. Dong, J. Ye and Y. He. 'Diffusion tensor tractography reveals disrupted topological efficiency in white matter structural networks in multiple sclerosis'. In: *Cerebral cortex* 21.11 (2011), pp. 2565–2577.

[157] G. Kocevar, C. Stamile, S. Hannoun, F. Cotton, S. Vukusic, F. Durand-Dubief and D. Sappey-Marinier. 'Graph theory-based brain connectivity for automatic classification of multiple sclerosis clinical courses'. In: *Frontiers in neuroscience* 10 (2016), p. 478.

[158] 'ECTRIMS 2024 – Poster'. In: *Multiple Sclerosis Journal* 30.3_suppl (2024), pp. 125–680. DOI: 10.1177/13524585241269219. eprint: https://doi.org/10.1177/13524585241269219. URL: https://doi.org/10.1177/13524585241269219.

[159] X. Wang, Z. Tang, M. Cabezas, A. D'Souza, D. Liu, M. Barnett, F. Calamante, C. Wang and W. Cai. 'Deep Learning-Driven Enhancement of Fiber Orientation Distribution: Effect of Choice of Gradient Direction Number'. In: *Annual Meeting of the International Society for Magnetic Resonance in Medicine (ISMRM)*. 2024.

[160] X. Wang, M. Cabezas, Z. Tang, A. D'Souza, D. Liu, C. Wang and W. Cai. 'Estimating Uncertainty on Deep Learning-Driven Fibre Orientation Distribution Enhancement'. In: *International Conference on Neural Information Processing*. Springer. 2024, pp. 361–376.

[161] F. Isensee, T. Wald, C. Ulrich, M. Baumgartner, S. Roy, K. Maier-Hein and P. F. Jaeger. 'nnU-Net revisited: A call for rigorous validation in 3d medical image segmentation'. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2024, pp. 488–498.

[162] Y. Tang, D. Yang, W. Li, H. R. Roth, B. Landman, D. Xu, V. Nath and A. Hatamizadeh. 'Self-supervised pre-training of swin transformers for 3d medical image analysis'. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022, pp. 20730–20740.

[163] A. Gu and T. Dao. 'Mamba: Linear-time sequence modeling with selective state spaces'. In: *arXiv preprint arXiv:2312.00752* (2023).

[164] S. Tu and M. C. Kiernan. ‘Amyotrophic lateral sclerosis’. In: *Advances in magnetic resonance technology and applications*. Vol. 9. Amsterdam: Elsevier, 2023, pp. 363–385.

[165] M. A. van den Bos, N. Geevasinga, M. Higashihara, P. Menon and S. Vucic. ‘Pathophysiology and diagnosis of ALS: insights from advances in neurophysiological techniques’. In: *International Journal of Molecular Sciences* 20.11 (2019), p. 2818.

[166] S. Tu, S. Vucic and M. C. Kiernan. ‘Pathological insights derived from neuroimaging in amyotrophic lateral sclerosis: emerging clinical applications’. In: *Current Opinion in Neurology* 37.5 (2024), pp. 577–584.

[167] H. Almgren, C. J. Mahoney, W. Huynh, A. D’Souza, S. Berte, J. Lv, C. Wang, M. C. Kiernan, F. Calamante and S. Tu. ‘Quantifying neurodegeneration within subdivisions of core motor pathways in amyotrophic lateral sclerosis using diffusion MRI’. In: *Journal of Neurology* 272.3 (2025), p. 215.

[168] J. Caga, S. Tu, T. Dharmadasa, N. Y. Tse, M. C. Zoing, W. Huynh, C. Mahoney, R. M. Ahmed and M. C. Kiernan. ‘Apathy is associated with parietal cortical-subcortical dysfunction in ALS’. In: *Cortex* 145 (2021), pp. 341–349.

[169] S. Tu, C. Wang, R. A. Menke, K. Talbot, M. Barnett, M. C. Kiernan and M. R. Turner. ‘Regional callosal integrity and bilaterality of limb weakness in amyotrophic lateral sclerosis’. In: *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* 21.5-6 (2020), pp. 396–402.

[170] W. Huynh, S. Tu, C. J. Mahoney, R. Schwartz and M. C. Kiernan. ‘Mills syndrome: clinical and radiologic asymmetry’. In: *Neurology* 96.14 (2021), pp. 677–678.

[171] S. Sharifi, A. J. Nederveen, J. Booij and A.-F. van Rootselaar. ‘Neuroimaging essentials in essential tremor: a systematic review’. In: *NeuroImage: Clinical* 5 (2014), pp. 217–231.

[172] K. Kamagata, A. Zalesky, T. Hatano, M. A. Di Biase, O. El Samad, S. Saiki, K. Shimoji, K. K. Kumamaru, K. Kamiya, M. Hori et al. ‘Connectome analysis with diffusion MRI in idiopathic Parkinson’s disease: evaluation using multi-shell, multi-tissue, constrained spherical deconvolution’. In: *NeuroImage: Clinical* 17 (2018), pp. 518–529.

[173] P. Bédard, P. Panyakaew, H.-J. Cho, M. Hallett and S. G. Horovitz. 'Multimodal imaging of essential tremor and dystonic tremor'. In: *NeuroImage: Clinical* 36 (2022), p. 103247.

[174] M. Bergamino, S. Aslam, J. J. Knittel, L. Alhilali and A. M. Stokes. 'White matter microstructural differences between essential tremor and Parkinson disease, evaluated using advanced diffusion MRI biomarkers'. In: *Journal of Integrative Neuroscience* 22.5 (2023), p. 114.

[175] A. J. Thompson, B. L. Banwell, F. Barkhof, W. M. Carroll, T. Coetzee, G. Comi, J. Correale, F. Fazekas, M. Filippi, M. S. Freedman et al. 'Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria'. In: *The Lancet Neurology* 17.2 (2018), pp. 162–173.

[176] M. Caranova, J. F. Soares, S. Batista, M. Castelo-Branco and J. V. Duarte. 'A systematic review of microstructural abnormalities in multiple sclerosis detected with NODDI and DTI models of diffusion-weighted magnetic resonance imaging'. In: *Magnetic Resonance Imaging* 104 (2023), pp. 61–71.

[177] I. Lipp, G. D. Parker, E. C. Tallantyre, A. Goodall, S. Grama, E. Patitucci, P. Heveron, V. Tomassini and D. K. Jones. 'Tractography in the presence of multiple sclerosis lesions'. In: *NeuroImage* 209 (2020), p. 116471.

[178] T. Charalambous, C. Tur, F. Prados, B. Kanber, D. T. Chard, S. Ourselin, J. D. Clayden, C. A. G. Wheeler-Kingshott, A. J. Thompson and A. T. Toosy. 'Structural network disruption markers explain disability in multiple sclerosis'. In: *Journal of Neurology, Neurosurgery & Psychiatry* 90.2 (2019), pp. 219–226.

[179] X. Wang, K. Wang, Z. Tang, M. Cabezas, C. Wang and W. Cai. 'UFREE: Unified Fiber Orientation Distribution Resolution Enhancement'. In: *2025 IEEE International Symposium on Biomedical Imaging (ISBI)*. IEEE. 2025, pp. 1–5.

[180] M. Barnett, D. Wang, H. Beadnall, A. Bischof, D. Brunacci, H. Butzkueven, J. W. L. Brown, M. Cabezas, T. Das, T. Dugal et al. 'A real-world clinical validation for AI-based MRI monitoring in multiple sclerosis'. In: *npj Digital Medicine* 6.1 (2023), p. 196.

[181] P. Goebl, J. Wingrove, O. Abdelmannan, B. Brito Vega, J. Stutters, S. D. G. Ramos, O. Kenway, T. Rossor, E. Wassmer, D. L. Arnold et al. ‘Enabling new insights from old scans by repurposing clinical MRI archives for multiple sclerosis research’. In: *Nature Communications* 16.1 (2025), p. 3149.

[182] Z. Tang, S. Chen, A. D’Souza, D. Liu, F. Calamante, M. Barnett, W. Cai, C. Wang and M. Cabezas. ‘High angular diffusion tensor imaging estimation from minimal evenly distributed diffusion gradient directions’. In: *Frontiers in Radiology* 3 (2023).

[183] A. T. Young, D. Amara, A. Bhattacharya and M. L. Wei. ‘Patient and general public attitudes towards clinical artificial intelligence: a mixed methods systematic review’. In: *The Lancet Digital Health* 3.9 (2021), e599–e611.

[184] L. Pinheiro Cinelli, M. Araújo Marins, E. A. Barros da Silva and S. Lima Netto. ‘Bayesian neural networks’. In: *Variational Methods for Machine Learning with Applications to Deep Networks*. Springer, 2021, pp. 65–109.

[185] Y. Gal and Z. Ghahramani. ‘Dropout as a bayesian approximation: Representing model uncertainty in deep learning’. In: *international conference on machine learning*. PMLR. 2016, pp. 1050–1059.

[186] Y. Gal. ‘Uncertainty in deep learning’. PhD thesis. University of Cambridge, 2016.

[187] B. Lakshminarayanan, A. Pritzel and C. Blundell. ‘Simple and scalable predictive uncertainty estimation using deep ensembles’. In: *Advances in neural information processing systems* 30 (2017).

[188] R. Baldock, H. Maennel and B. Neyshabur. ‘Deep learning through the lens of example difficulty’. In: *Advances in Neural Information Processing Systems* 34 (2021), pp. 10876–10889.

[189] W. H. Beluch, T. Genewein, A. Nürnberger and J. M. Köhler. ‘The power of ensembles for active learning in image classification’. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 9368–9377.

[190] M. Abdar, F. Pourpanah, S. Hussain, D. Rezazadegan, L. Liu, M. Ghavamzadeh, P. Fieguth, X. Cao, A. Khosravi, U. R. Acharya et al. ‘A review of uncertainty quantification in deep learning: Techniques, applications and challenges’. In: *Information Fusion* 76 (2021), pp. 243–297.

[191] K. Kushibar, V. Campello, L. Garrucho, A. Linardos, P. Radeva and K. Lekadir. 'Layer Ensembles: A Single-Pass Uncertainty Estimation in Deep Learning for Segmentation'. In: *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*. Springer. 2022, pp. 514–524.

[192] C.-M. Feng, H. Fu, S. Yuan and Y. Xu. 'Multi-contrast mri super-resolution via a multi-stage integration network'. In: *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part VI 24*. Springer. 2021, pp. 140–149.

[193] J. Xu and E. Adalsteinsson. 'Deformed2self: Self-supervised denoising for dynamic medical imaging'. In: *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*. Springer. 2021, pp. 25–35.

[194] C. E. Shannon. 'A mathematical theory of communication'. In: *The Bell system technical journal* 27.3 (1948), pp. 379–423.

[195] M.-T. Puth, M. Neuhäuser and G. D. Ruxton. 'Effective use of Spearman's and Kendall's correlation coefficients for association between two measured traits'. In: *Animal Behaviour* 102 (2015), pp. 77–84.

[196] J. Wasserthal, P. Neher and K. H. Maier-Hein. 'TractSeg - Fast and accurate white matter tract segmentation'. In: *Neuroimage* 183 (2018), pp. 239–253.

[197] B. Jeurissen, A. Leemans, J.-D. Tournier, D. K. Jones and J. Sijbers. 'Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging'. In: *Human brain mapping* 34.11 (2013), pp. 2747–2766.

[198] J. S. Elam, M. F. Glasser, M. P. Harms, S. N. Sotiropoulos, J. L. Andersson, G. C. Burgess, S. W. Curtiss, R. Oostenveld, L. J. Larson-Prior, J.-M. Schoffelen et al. 'The human connectome project: a retrospective'. In: *NeuroImage* 244 (2021), p. 118543.

[199] A. V. D. Oord, O. Vinyals et al. 'Neural discrete representation learning'. In: *NIPS* 30 (2017).

[200] S. Chen, K. Ma and Y. Zheng. 'Med3d: Transfer learning for 3d medical image analysis'. In: *arXiv preprint arXiv:1904.00625* (2019).

[201] X. Zhu, W. Zhang, Y. Li, L. J. O'Donnell and F. Zhang. 'When Diffusion MRI Meets Diffusion Model: A Novel Deep Generative Model for Diffusion MRI Generation'. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2024, pp. 530–540.