

Field-of-view extension for brain diffusion MRI via deep generative models

Chenyu Gao¹,^a Shunxing Bao¹,^a Michael E. Kim¹,^b Nancy R. Newlin¹,^b
Praitayini Kanakaraj¹,^b Tianyuan Yao¹,^b Gaurav Rudravaram,^a Yuankai Huo¹,^b
Daniel Moyer,^b Kurt Schilling,^c Walter A. Kukull,^d Arthur W. Toga,^e
Derek B. Archer,^{f,g} Timothy J. Hohman¹,^{f,g} Bennett A. Landman¹,^{a,b,c}
and Zhiyuan Li¹,^{a,*}

^aVanderbilt University, Department of Electrical and Computer Engineering, Nashville, Tennessee, United States

^bVanderbilt University, Department of Computer Science, Nashville, Tennessee, United States

^cVanderbilt University Medical Center, Department of Radiology and Radiological Sciences, Nashville, Tennessee, United States

^dUniversity of Washington, Department of Epidemiology, Seattle, Washington, United States

^eUniversity of Southern California, Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, Laboratory of Neuro Imaging, Los Angeles, California, United States

^fVanderbilt University Medical Center, Vanderbilt Memory and Alzheimer's Center, Nashville, Tennessee, United States

^gVanderbilt University Medical Center, Vanderbilt Genetics Institute, Nashville, Tennessee, United States

ABSTRACT. Purpose: In brain diffusion magnetic resonance imaging (dMRI), the volumetric and bundle analyses of whole-brain tissue microstructure and connectivity can be severely impeded by an incomplete field of view (FOV). We aim to develop a method for imputing the missing slices directly from existing dMRI scans with an incomplete FOV. We hypothesize that the imputed image with a complete FOV can improve whole-brain tractography for corrupted data with an incomplete FOV. Therefore, our approach provides a desirable alternative to discarding the valuable brain dMRI data, enabling subsequent tractography analyses that would otherwise be challenging or unattainable with corrupted data.

Approach: We propose a framework based on a deep generative model that estimates the absent brain regions in dMRI scans with an incomplete FOV. The model is capable of learning both the diffusion characteristics in diffusion-weighted images (DWIs) and the anatomical features evident in the corresponding structural images for efficiently imputing missing slices of DWIs in the incomplete part of the FOV.

Results: For evaluating the imputed slices, on the Wisconsin Registry for Alzheimer's Prevention (WRAP) dataset, the proposed framework achieved $PSNR_{b_0} = 22.397$, $SSIM_{b_0} = 0.905$, $PSNR_{b_{1300}} = 22.479$, and $SSIM_{b_{1300}} = 0.893$; on the National Alzheimer's Coordinating Center (NACC) dataset, it achieved $PSNR_{b_0} = 21.304$, $SSIM_{b_0} = 0.892$, $PSNR_{b_{1300}} = 21.599$, and $SSIM_{b_{1300}} = 0.877$. The proposed framework improved the tractography accuracy, as demonstrated by an increased average Dice score for 72 tracts ($p < 0.001$) on both the WRAP and NACC datasets.

Conclusions: Results suggest that the proposed framework achieved sufficient imputation performance in brain dMRI data with an incomplete FOV for improving whole-brain tractography, thereby repairing the corrupted data. Our approach achieved more accurate whole-brain tractography results with an extended and complete FOV and reduced the uncertainty when analyzing bundles associated with Alzheimer's disease.

*Address all correspondence to Zhiyuan Li, zhiyuan.li@vanderbilt.edu

© The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JMI.11.4.044008](https://doi.org/10.1117/1.JMI.11.4.044008)]

Keywords: medical image synthesis; diffusion MRI; imputation; generative model

Paper 24094GRR received Mar. 25, 2024; revised Aug. 1, 2024; accepted Aug. 1, 2024; published Aug. 24, 2024.

1 Introduction

Diffusion magnetic resonance imaging (dMRI) offers a non-invasive, *in vivo* approach for measuring the diffusion of water molecules in biological tissues and has become a well-established technique for studying human white matter microstructure and connectivity.¹⁻⁴ The movement of water molecules is often restricted by biological structures such as cell membranes and axonal fibers, resulting in a preferred direction of movement that reflects the properties of tissues. A standard dMRI scan is designed to acquire multiple volumes under varying magnetic fields (i.e., by applying diffusion-encoding magnetic gradient pulse from a number of non-collinear directions), such that each volume selectively captures the propensity of water diffusivity in a particular direction, thereby yielding diffusion-weighted images (DWIs). The effect of the gradient pulse, both in terms of time and strength, is characterized by a parameter known as the *b*-value. In addition, the orientation of the gradient is commonly specified as a unit-length vector known as the *b*-vector, and the high diffusivity of water molecules along the gradient orientation yields high signal attenuation. Reference volumes with no diffusion signal attenuation, i.e., with a *b*-value equal to 0 s/mm² and a *b*-vector equal to (0, 0, 0), are also required to be acquired during a dMRI scan and are often referred to as *b*0 images. To quantify the properties of water diffusion in brain tissues, voxel-wise scalar metrics such as mean diffusivity and fractional anisotropy are derived from an assumed diffusion tensor (ellipsoidal) model.⁵ In addition, to study whole-brain physical connections, fiber tractography methods delineate the white matter fiber pathways connecting regions of the brain.^{6,7} In the last decade, dMRI and its related diffusion measures have become the method of choice to study brain tissue properties and changes associated with Alzheimer's disease, stroke, schizophrenia, and aging.^{6,8-11}

Despite the unique clinical capabilities and potential, the whole-brain volumetric and tractography analyses brought by dMRI can be severely impeded by an incomplete field of view (FOV), commonly caused by patient misalignment, suboptimal scan plan selection, or necessity in protocol design. A major limitation of dMRI is the extended acquisition time compared with traditional structural MRI due to the acquisition of volumes with varying diffusion-encoding gradient directions. Typically, protocols with more than 31 directions are recommended for longitudinal studies of disease progression or treatment effects.¹² The long acquisition time further amplifies clinical constraints and imaging artifacts in dMRI such as inter-volume motion and eddy-current-induced artifacts.¹³⁻¹⁵ As a result, the FOV may be incomplete for whole-brain scans in suboptimal dMRI acquisition. This then leads to corrupt data with a sequence of brain slices missing in the incomplete part of the FOV, which is one of the most common issues identified during quality assurance of dMRI data.¹⁶ In a recent study of dMRI datasets, we found 103 cases with incomplete FOVs out of a total of 1057 cases that failed quality assurance of dMRI preprocessing. The estimated thickness for the missing regions ranged from 1 to 32 mm (Fig. 1). The loss of information from the missing slices not only prevents analyses in those missing regions but may also affect dMRI-derived analyses of acquired regions (Fig. 2) as the global patterns based on the whole brain are impacted. Furthermore, corrupted data with missing slices introduce bias and inaccuracies for whole-brain analyses, posing significant challenges for longitudinal studies in diagnosing and monitoring neurological developments, including Alzheimer's disease.^{17,18}

As reacquiring the data is not a feasible solution, imputing the missing slices directly from existing scans with an incomplete FOV provides a desirable alternative to discarding the affected but valuable data or re-engineering all downstream methods to accommodate the effects of missing data. Many works have been dedicated to alleviating the impact of missing dMRI data. RESTORE¹⁹ is among the pioneering efforts that introduced an iteratively reweighted least-squares regression for robust estimation of the diffusion tensor model by outlier rejection.

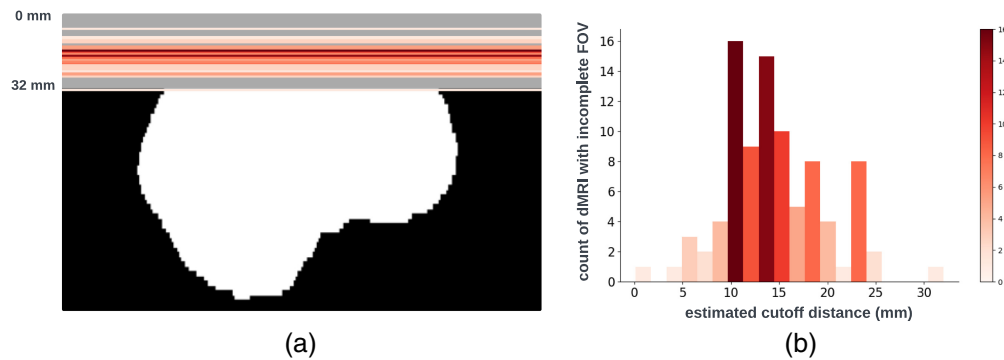


Fig. 1 Visualization (a) and histogram (b) of 103 real cases of dMRI scans with an incomplete FOV that failed quality assurance. In panel (a), horizontal red lines and background gray areas indicate where the reduced FOV ends and its corresponding missing regions, respectively, with the estimated position of a brain mask. The total cutoff distance from the reduced FOV to the top of the brain is estimated using a corresponding and registered T1w image.

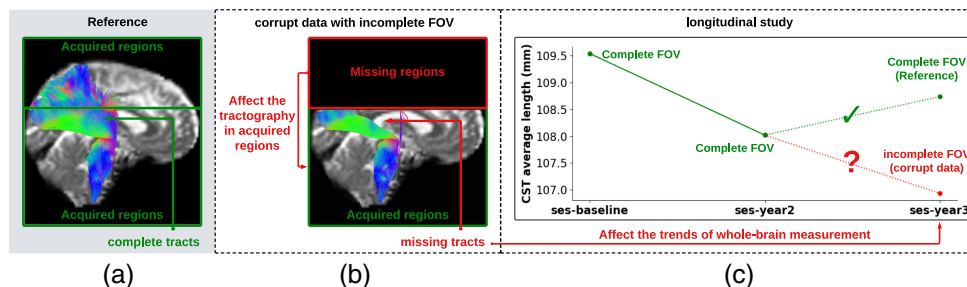


Fig. 2 Missing regions resulting from an incomplete FOV not only render analyses of those areas impossible but can also impact the tractography performed in the acquired regions [as shown in panel (b)], e.g., yielding missing streamlines of corticospinal tract (CST) compared with the reference [as shown in panel (a)]. Furthermore, whole-brain measurements derived from corrupted data can lead to incorrect interpretations in longitudinal studies [as shown in panel (c)]: the measurement from corrupted data (represented by the red dot for the “year3” session) might suggest that the average length of the CST for this subject continues to decrease. This, however, may contradict when correct measurements are considered (represented by green dots).

Recently, TW-BAG²⁰ is an inpainting neural network method for repairing the diffusion tensors in cropped regions. For the diffusion kurtosis model,²¹ which further quantifies the non-Gaussianity of water diffusion in the brain, REKINDLE²² was proposed as a robust estimation procedure to address the increased sensitivity to artifacts and model complexity. However, designing specific methods for each of the numerous and rapidly evolving diffusion and microstructural models would be challenging and inefficient. As an alternative, researchers have also put efforts into repairing the raw DWI signals directly. FSL’s “eddy”²³ and SHORE-based method²⁴ were developed to detect signal dropout and to impute the affected measurements across acquired DWI volumes. However, these methods focus on the imputation of dropout slices based on reference slice signals computed from multiple volumes and cannot be applied to the FOV extension task in which no signals are available in the incomplete part of the FOV. A reliable imputation of raw DWI signals for a contiguous sequence of regions in the incomplete part of the FOV remains an unresolved task.

To propose a first solution for this task, we turn to the recent rapid advancements in deep learning, which have shown great potential in image synthesis tasks for dMRI, such as distortion correction,²⁵ denoising,^{26–28} and registration.^{29,30} Directly generating a sequence of dMRI slices in the incomplete part of the FOV, similar to the in-painting task in computer vision, can be challenging, and how to maintain and improve the consistency between the synthesized and observed regions remains an open question.^{31,32} Moreover, in medical image synthesis, it is of greater significance for the synthesized regions to conform to the subject’s authentic

anatomical structures rather than being merely visually realistic. This restrictive requirement of anatomical alignment makes it difficult to naively adapt in-painting models, in which the outputs are often diverse.^{33,34} However, advantageously, high-quality T1-weighted images are commonly acquired as a default alongside a dMRI scan and can be utilized as an anatomical reference. Existing works have shown promising results for integrating the additional anatomical information from T1-weighted images into image synthesis methods for dMRI, such as correcting diffusion distortion by synthesized b_0 image,²⁵ synthesizing high angular resolution dMRI data,³⁵ and tractography estimation.³⁶ Inspired by these findings, in this work, we propose a deep generative model framework that imputes the missing brain regions of a DWI in the incomplete part of the FOV with extra information from the corresponding T1-weighted image. The proposed model integrates both the diffusion information within the DWI and the structural information of T1-weighted images for accurate imputation of missing slices. A combination of 2.5-dimensional neural networks is proposed for efficient graphics processing unit (GPU) usage and reduced application time. Cross-plane prediction corrections are further applied to improve the spatial consistency.

We first train and evaluate our methods on one dMRI dataset with 343 subjects from the same site. To assess generalizability and robustness, we subsequently perform an evaluation on another dMRI dataset with 50 subjects from another site. We reported the missing DWI slice imputation performance using the peak signal-to-noise ratio (PSNR) and the structural similarity index measure (SSIM). We demonstrate that our approach can improve tractography accuracy for both imputed and acquired brain regions to reduce uncertainty when analyzing bundles associated with Alzheimer’s disease.

The primary contributions of this paper are as follows: (1) we propose a framework that imputes DWI conditioned on T1-weighted (T1w) images using a deep generative model. We investigate the possibility of synthesizing multi-volume DWI in the incomplete parts of the FOV, as an advancement of existing work that only synthesizes b_0 images based on T1w images. The deep generative model fills the gap that traditional imputation methods fail to address, specifically in imputing DWI slices in the incomplete parts of the FOV. (2) We demonstrate that the imputation achieved by our work significantly increases the accuracy of whole-brain tractography, thereby repairing corrupted DWI data and making it available for conducting downstream tract analysis tasks.

2 Methods

2.1 Problem Setting

Given a diffusion-weighted image $x \in \mathbb{R}^4$ that may have an incomplete FOV, we want to learn a mapping from observed image x to output image $y \in \mathbb{R}^4$, $G: x \rightarrow y$, such that y will have a complete whole-brain FOV with imputed slices if necessary. To tackle the mapping of the DWI with V volumes, we map each volume $x_v \in \mathbb{R}^3$ ($v = 1, 2, 3, \dots, V$) separately to its corresponding output volume $y_v \in \mathbb{R}^3$ ($v = 1, 2, 3, \dots, V$). Then, the output image y is obtained by combining each output volume y_v with the corresponding b -value and b -vector in the gradient table. Directly predicting y_v from x_v can be difficult, given that there are infinite possible gradient directions, each requiring unique feature learning and altogether making the representation learning from x_v complex. We utilize an available T1w image with a complete FOV $x_{T1} \in \mathbb{R}^3$ as an extra input, aiming to provide additional information on anatomical structures within x_{T1} . Furthermore, given the input pair $\{x_{T1}, x_v\}$, the same x_{T1} shared across all DWI volumes could benefit the optimization of G because it allows the model to leverage a consistent structural reference x_{T1} while learning to predict various missing slices in the DWI, focusing on their unique and inherent contrast and directional characteristics within x_v . Following the ideas described above, Fig. 3 illustrates the comprehensive processing pipeline for the proposed framework of imputing DWI volumes.

2.2 Datasets and Data Preprocessing

In this study, we initially selected the Wisconsin Registry for Alzheimer’s Prevention (WRAP)³⁷ dataset as the primary source for training and evaluating our methodologies. The rationale behind this choice is twofold. First, the WRAP dataset contains one of the most extensively corrupted

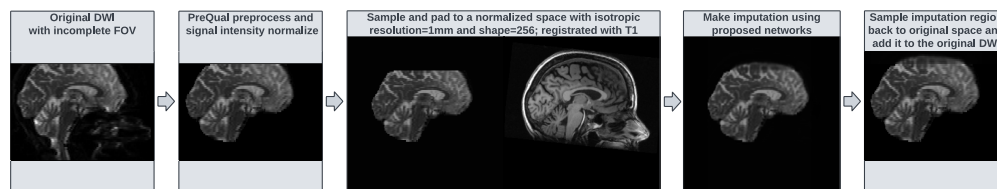


Fig. 3 Pipeline of the proposed FOV extension framework for imputing missing slices in the incomplete part of the FOV begins with PreQual preprocessing and intensity normalization for the DWI in its original space. This is followed by processing the DWI to a normalized space, including resampling and registration with its corresponding T1 image. Subsequently, the proposed 2.5D pix2pix networks are employed to impute the missing slices in the normalized space, utilizing both the DWI (incomplete FOV) and the corresponding T1 (complete FOV). Finally, the imputed regions are resampled back to their original space and added to the original DWI. Sagittal views of a b_0 volume with an incomplete FOV at each pipeline stage are visualized.

dMRI data in terms of the significant missing regions of the brain close to 30 mm due to an incomplete FOV. Second, WRAP was collected from a single site, making it an ideal starting point for training and evaluating models without the concerns of variations across multiple sites. Our first cohort on WRAP comprised 343 subjects, each possessing T1w image and single-shell dMRI scans with a b -value of 1300 s/mm^2 , the most frequent b -value acquired in WRAP. These subjects were split into three distinct groups: 245 subjects for the training set, 49 for the validation set, and 49 for the testing set. Next, to evaluate the robustness and generalizability of the proposed method, we extended our analysis to include the National Alzheimer’s Coordinating Center (NACC)³⁸ dataset, which has a large number of dMRI scans sharing the same b -value of 1300 s/mm^2 . Our second cohort comprised 49 testing subjects from the same site within NACC, each possessing T1w image and single-shell dMRI scans with a b -value of 1300 s/mm^2 . Table 1 presents the diagnosis information about the cohorts included in our study.

All DWIs were first preprocessed using the PreQual³⁹ pipeline for correction of susceptibility-induced and eddy-current-induced artifacts, slice-wise imputation of mid-brain slices, inter-volume motion, and denoising. Quality assurance checks were performed on PreQual preprocessing reports and output images to ensure valid inputs and successful preprocessing of the data. Next, intensity normalization was performed for each DWI separately, with the maximum value set to the 99.9th percentile intensity and the minimum value set to 0. All volumes of one DWI shared the same normalization parameters. The corresponding T1w image was normalized with a maximum value of its 99.9th percentile intensity and a minimum value of 0. Then, the T1w image was registered to the DWI by applying an affine transformation computed between the T1w image and the average b_0 image of the DWI using FSL’s *epi reg*.⁴⁰ Then, both the T1w image and all DWI volumes were resampled to $1 \times 1 \times 1 \text{ mm}$ resolution and padded or cropped to $256 \times 256 \times 256$ voxels.

2.3 Model

The proposed neural networks for DWI imputation are presented in Fig. 4. For tackling the large GPU memory required by learning the 3D mapping $G: \{x_{T1}, x_v\} \rightarrow y_v$, we propose a 2.5D

Table 1 Subjects’ diagnosis on training, validation, and testing sessions for WRAP and NACC datasets. The names of diagnosis follow the original subject’s demographic files.

	WRAP			NACC	
	Alzheimer’s dementia	Mild cognitive impairment	No cognitive impairment	Dementia	Normal
Train	2	2	241	N/A	N/A
Val	0	1	48	N/A	N/A
Test	0	1	48	13	36

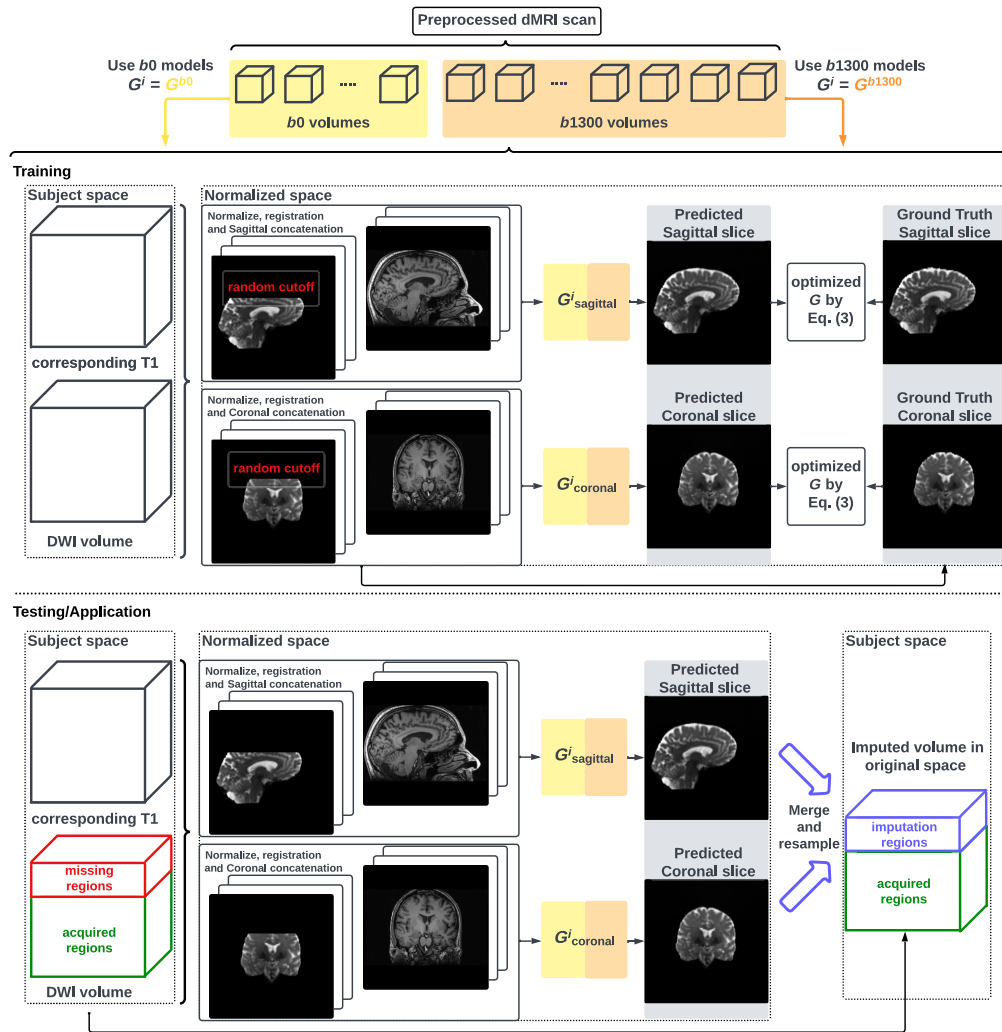


Fig. 4 Whole DWI imputation task is divided into four sub-tasks: imputing b_0 volumes' sagittal slices, imputing b_0 volumes' coronal slices, imputing b_{1300} volumes' sagittal slices, and imputing b_{1300} volumes' coronal slices. The proposed 2.5D networks contain four sub-networks that share the same pix2pix network architecture, and each subnetwork is designed to process a specific sub-task for b_0 or b_{1300} images with their sagittal or coronal slices. During training, random regions are cut off from either top or bottom of brain to obtain training DWI data with an incomplete FOV, and the subnetworks are optimized to make the imputation of the cutoff regions using Eq. (3). In the testing or application case, for each DWI volume with an incomplete FOV, its corresponding sagittal and coronal subnetworks each output an imputed volume by combining every imputed sagittal or coronal slice, respectively. These two imputed volumes are then merged into one volume to improve 3D consistency. The imputation regions of the final merged volume are sampled back to the original subject space and are added to the original DWI volume.

framework to decompose G into two separate generators, G_{sagittal} and G_{coronal} , and learn them independently through small patches of 3D volume in sagittal and coronal views, respectively. Each small patch contains a sequence of neighboring slices of the target slice (n for each side) and is then used to predict a single slice in the sagittal and coronal views. The predictions from the sagittal and coronal views are later merged by voxel averaging to obtain the final output volume. We trained separate models to handle the distribution difference between DWI volumes obtained with a b -value equal to 0 or 1300 s/mm^2 , resulting in four generators in total: $G_{b_0\text{-sagittal}}$, $G_{b_0\text{-coronal}}$, $G_{b_{1300}\text{-sagittal}}$, and $G_{b_{1300}\text{-coronal}}$. The axial view was not included in the model because the axial slices of DWI in the incomplete FOV regions are not available and, therefore, provide no information about the diffusion features for training. We use pix2pix⁴¹ as our generator G for its stable conditional image translation and L_1 loss for preserving the

underlying context of the image,⁴² which is critical for medical image synthesis tasks. The final objective for every G is

$$\mathcal{L}_{\text{GAN}}(G, D) = \mathbb{E}_{y_v} [\log D(y_v)] + \mathbb{E}_{x_v} [\log(1 - D(G(x_{T1}, x_v)))] \quad (1)$$

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x_v, y_v} [\|y_v - G(x_{T1}, x_v)\|_1] \quad (2)$$

$$G^* = \arg \min_G \max_D \mathcal{L}_{\text{GAN}}(G, D) + \lambda \mathcal{L}_{L1}(G) \quad (3)$$

where D is a discriminator to distinguish if the output of generator G looks real.

During training, first, a DWI volume and its corresponding T1w image (registered as in data preprocessing) are randomly selected. The DWI volume is randomly cut off by 0 to 50 mm in the normalized space, which covers the maximum missing distance, as previously shown in Fig. 1, and for model generalizability, from either the top or bottom of the brain. The cutoff DWI is then paired with its T1w image as input. The non-cutoff DWI volume is used as the ground truth for the prediction. Then, small patches of the sagittal and coronal views are created: DWI and T1w patches are concatenated along the plane direction. For example, if sagittal DWI and T1w patches are both $(2n + 1) \times 256 \times 256$, their concatenation will be $((2n + 1) + (2n + 1)) \times 256 \times 256$. Finally, the corresponding G_{sagittal} and G_{coronal} are optimized by stochastic gradient descent using Eq. (3), where the expectation of x_v and y_v is approximated by mini-batches of image slices. In our design, we train the model to predict the whole regions of the brain (both cutoff and non-cutoff regions) instead of cutoff regions only. We reason that this can encourage the model to learn global representations of the image and thus enhance the model's robustness and generalizability for various sizes of incomplete FOVs, including the case in which the input image already has a complete FOV. We adopt the state-of-the-art PyTorch implementations (<https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix>) for training every generator. As suggested in pix2pix, we choose the deterministic G for efficient model training. We used "resnet_9blocks" as the network architecture for G to encourage the model to explore features within both T1w images and DWIs. We set $n = 7$ as the minimum requirement for maintaining 3D consistency. The best model was selected by the imputation performance on the imputed regions only using the validation set.

For testing and application, the model follows the same process to obtain the predicted volume. For the final framework output, we use only the slices in the missing regions of the predicted volume. The imputed regions are sampled back to the original subject space and then combined with the originally acquired regions with an incomplete FOV. A mask m that covers the acquired regions (if $m = 1$: acquired regions, else: missing regions) can be generated from the testing data with any brain-masking methods ("median_otsu" as a simple example), and the final output is therefore $m \odot x_v + (1 - m) \odot \tilde{y}_v$. For all images of the testing subjects, we first cropped them by 30 mm to obtain testing images with an incomplete FOV. We then used the original full FOV images as our ground truth reference images.

The model is implemented using Python 3.11.5 and PyTorch 2.3.0, along with CUDA 11.8. All experiments were run on an Nvidia Quadro RTX 5000 with 16 GB of GPU memory. The batch size is set to 24, and four parallel PyTorch data loading workers are used.

2.4 Analysis

First, we qualitatively and quantitatively evaluate the imputation errors on the WRAP dataset. We report mean squared error (MSE), PSNR, and SSIM for the imputed regions compared with the ground truth reference. The SSIM window is set to 7 for every dimension. Brain masks computed by spatially localized atlas network tiles for intracranial measurements⁴³ are applied to ensure that the metrics are computed for brain areas only. In addition, we study the imputation performance with respect to the missing slice distance and concerning different directions of the diffusion-encoding gradient pulse.

Next, to test our hypothesis that an imputed image with a complete FOV, generated by our approach, can improve whole-brain tractography for corrupted data with an incomplete FOV, we conduct paired t -tests for 72 tracts and specifically investigate 12 of them that are commonly associated with Alzheimer's disease (AD). We present Bland-Altman plots for studying the agreement of bundle shape measurements between the reference and our approach.

Then, to test our hypothesis that T1w images can be helpful for multi-volume dMRI imputation, we conduct an ablation study. This study also serves as a baseline for the proposed model by training a model with the same neural network architecture and settings but without the input of T1w images.

Finally, to evaluate the generalizability of our methods, we report the imputation errors using PSNR and SSIM on an additional NACC dataset. We also conduct the same tractography and bundle analysis on the NACC dataset.

3 Results

3.1 Imputation of Missing Slices

In general, the proposed method is capable of imputing visually similar slices for both the top and bottom of the brain, with similar global contrast and anatomical patterns compared with the ground truth reference. The major differences observed were at the boundaries between the white matter and the gray matter (Fig. 5). The imputation errors increase when the imputed slice is located toward the edges of the brain, i.e., distant from its nearest acquired regions (Figs. 6 and 7). MSE, PSNR, and SSIM for the imputed slices of the testing subjects are recorded in Table 2. In addition, we studied how the imputation performance can vary in relation to the directions of the diffusion-encoding gradient pulse. The apparent diffusion coefficient (ADC) was computed for 40 directions within the testing subjects. The proposed method showed no obvious bias toward specific directions as evidenced by the similar PSNR of ADC observed across all directions (Fig. 8).

3.2 Bundle Analyses

We are interested in how our approach can help repair the bundles and increase the tractography accuracy within both the acquired and imputed regions. To evaluate this, we ran Tractseg⁴⁴ on images with an incomplete FOV, their imputed image generated by our approach, and their ground truth reference images with a complete FOV. In particular, we studied a group of 12 tracts, including Rostrum (CC_1), Genu (CC_2), Isthmus (CC_6), and Splenium (CC_7) of the corpus callosum (CC) as well as left and right cingulum (CG), fornix (FX), Inferior occipito-frontal fascicle (IFO), and superior longitudinal fascicle I (SLF_I). These tracts are commonly associated with Alzheimer's disease (AD)^{45–59} and were examined to explore the potential clinical benefits of the proposed framework.

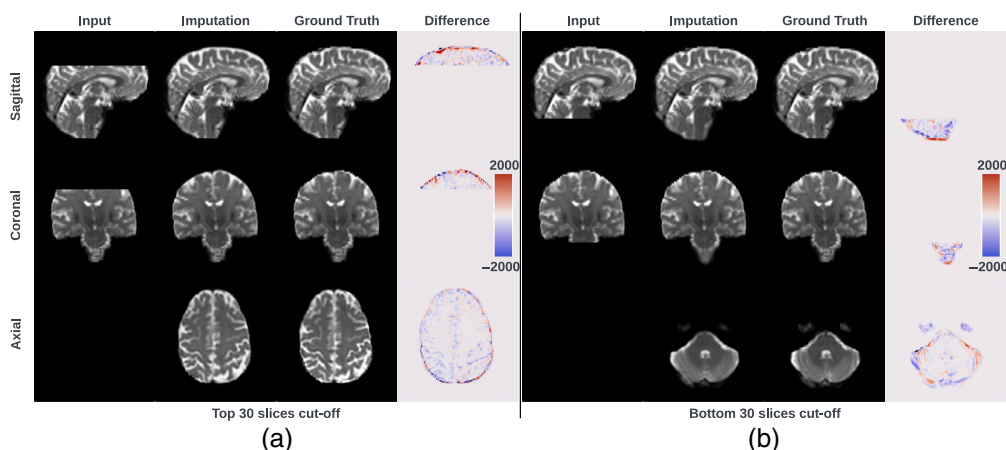


Fig. 5 Imputation for both the top (a) and bottom (a) of the brain. Red and blue indicate that the imputed intensity is larger or smaller than the ground truth, respectively. The imputation achieved similar global contrast and anatomical patterns compared with the ground truth reference. A closer examination of local areas, as indicated by the difference image, reveals large imputation errors at the boundaries between the white matter and the gray matter and at the edges of the brain. In addition, the proposed framework tends to make blurry imputation, thereby losing the high-frequency information that details the brain structure.

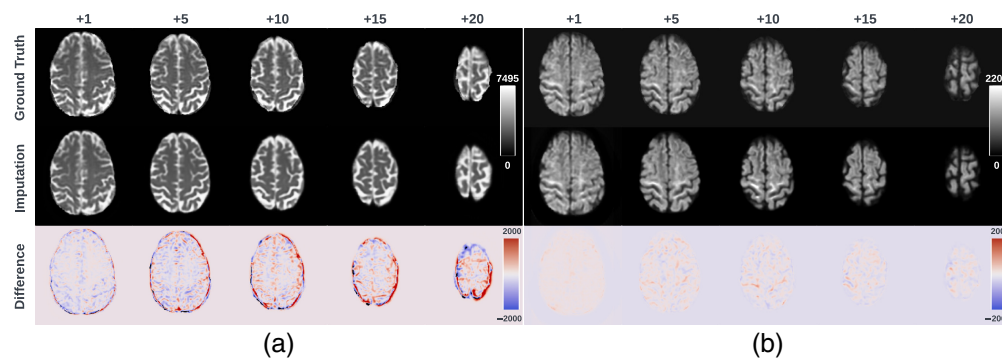


Fig. 6 Axial slice imputations for b_0 images (a) and b_{1300} images (b). The color lookup tables are adjusted with different intensity ranges for a better display of diffusion-weighted volumes. Each column represents the distance to the nearest acquired slice in millimeters (mm). Red and blue indicate that the imputed intensity is larger or smaller than the ground truth reference, respectively. Consistent with Fig. 5, the proposed framework performs imputations that globally align with the ground truth reference, albeit with a blurrier appearance. In addition, increasing imputation errors are observed as the distance of the imputed slices increases, for both b_0 and b_{1300} images.

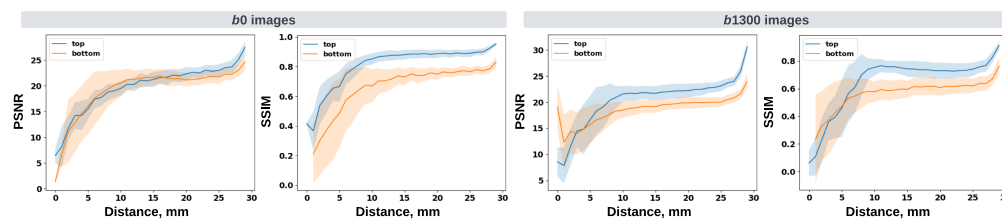


Fig. 7 Imputation performance with respect to the distance from the top or bottom of the brain, assuming a complete brain. The larger the distance is, the closer it is to the acquired region. Both PSNR and SSIM metrics for b_0 and b_{1300} images show an ascending trend, indicating an improving imputation accuracy when approaching the nearest acquired region, and a higher error margin in slices adjacent to the top or bottom of the brain. At a 30 mm distance, which is approximately the closest missing slice to the acquired brain region, the imputation accuracy markedly improves, as evidenced by the rising tail of each plotted line.

Table 2 Average MSE, PSNR, and SSIM (3D) for imputation regions of testing data on the WRAP and NACC datasets.

		WRAP		NACC	
		b_0 images	b_{1300} images	b_0 images	b_{1300} images
Baseline (no T1w)	MSE	45.421 ± 25.106	15.843 ± 3.741	31.552 ± 17.117	8.528 ± 2.071
	PSNR	16.679 ± 0.976	6.458 ± 1.769	16.764 ± 1.358	7.185 ± 1.606
	SSIM	0.691 ± 0.066	0.255 ± 0.079	0.727 ± 0.051	0.279 ± 0.071
Proposed model	MSE	12.483 ± 8.041	0.418 ± 0.192	11.007 ± 5.566	0.320 ± 0.141
	PSNR	22.397 ± 1.573	22.479 ± 1.560	21.304 ± 1.456	21.599 ± 1.299
	SSIM	0.905 ± 0.047	0.893 ± 0.042	0.892 ± 0.040	0.877 ± 0.021

Our method achieved a slightly superior performance on b_0 images than b_{1300} images, as indicated by the SSIM metrics. The ablation of the T1w image inputs significantly decreases the imputation performance, as demonstrated by all three metrics.

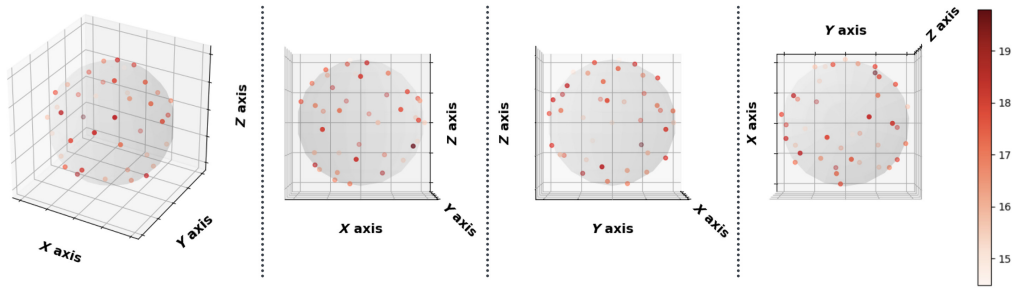


Fig. 8 Imputation performance (PSNR) with respect to 40 directions of diffusion-encoding gradient pulse evaluated by ADC. The average PSNR of ADC is 16.991 ± 1.221 . No obvious visual bias is observed in any direction. p -Value > 0.05 for the Kruskal–Wallis test ($p = 0.999$), which fails to reject the null hypothesis that the medians of each direction’s measurements are the same.

As shown in Fig. 9, the tracts produced in the imputed regions outside of the previously incomplete FOV are visually very similar to their ground truth reference. However, they lack some streamlines around the edge of the brain. In addition, in the acquired regions of the original DWI, our method improves the accuracy and completeness of tracts that are substantially affected by an incomplete FOV. This improvement is particularly evident in tracts that were previously undetected or only partially produced due to the incomplete FOV.

Quantitatively, the Dice similarity coefficient (Dice score) was computed for all 72 tracts generated by Tractseg. For an accurate comparison, we analyze the tracts derived from images with an incomplete FOV alongside those from their corresponding imputed images. Both are matched against the same tract segmentation obtained from the ground truth image with a complete FOV. Subsequently, we calculate two Dice scores: one comparing the reference tracts with those from the incomplete FOV images, and another comparing the reference tracts with those from the imputed images. For ease of reference, we label these scores as “Dice for Incomplete

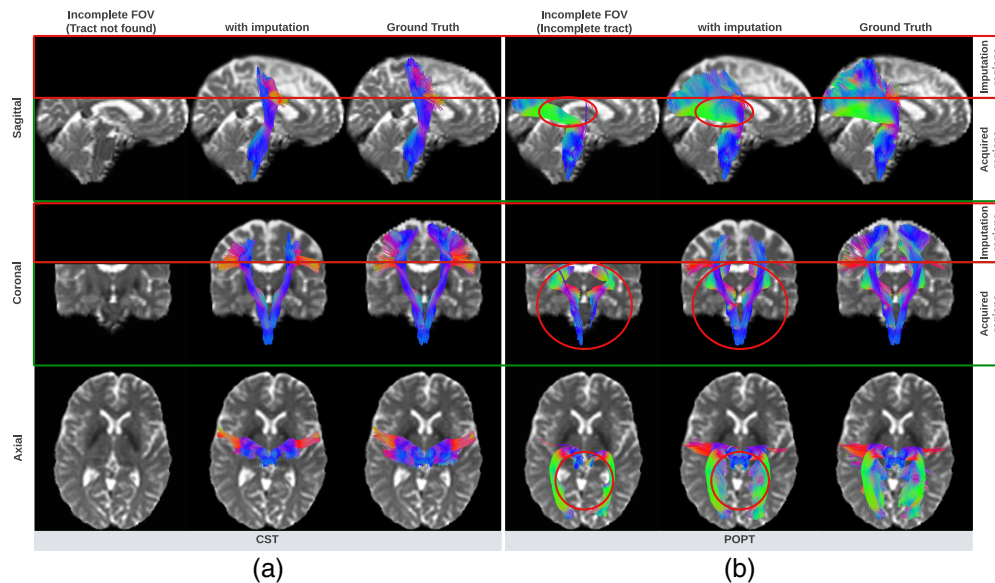


Fig. 9 Tractography results of example tracts for images with an incomplete FOV alongside their imputed counterparts and the ground truth references. The tracts produced through imputed images closely resemble the ground truth reference tracts within the acquired regions but lack some streamlines near the brain’s edge in the imputed regions. Our approach notably enhances the accuracy and completeness of tracts that are significantly compromised with an incomplete FOV. As shown in panel (a), corticospinal tract (CST) is completely not detected for images with an incomplete FOV, but with imputation, CST is produced successfully. In panel (b), the image with an incomplete FOV yields a partial Parieto-occipital pontine (POPT) only, yet the imputed image rectifies and completes the tract’s overall shape and structure within the acquired regions.

Table 3 Average Dice score for 72 tracts produced from an image with an incomplete FOV and with its imputation.

	WRAP		NACC	
	Incomplete FOV	With imputation	Incomplete FOV	With imputation
Acquired regions	0.909 ± 0.026	0.933 ± 0.021	0.884 ± 0.036	0.921 ± 0.022
Imputed regions	N/A	0.646 ± 0.180	N/A	0.643 ± 0.173

The improvement of imputation over the incomplete FOV is statistically significant ($p < 0.001$) from paired t -test on all tracts' results ($p = 2.52 \times 10^{-20}$ for WRAP and $p = 1.25 \times 10^{-24}$ for NACC).

FOV” and “Dice for Imputation,” respectively. Our approach significantly improved ($p < 0.001$) the quality of all 72 tracts on average in the acquired regions while achieving reasonable Dice scores in imputed regions (Table 3). Likewise, the enhancement of the 12 tracts commonly associated with AD in acquired regions was statistically significant ($p < 0.001$), as shown in Table 4. In addition, we analyzed two distinct groups of tracts. One group contains 50 cutoff tracts with ground truth tracts that can be cut off by an incomplete FOV, up to 30 mm from the top of the brain. The other group includes 22 no-cutoff tracts with ground truth tracts that are situated far from the top of the brain and, therefore, are not cut off by an incomplete FOV. For both groups, our approach significantly improved the tractography accuracy (Table 5). For a detailed examination, a comprehensive Dice score comparison of all 72 tracts is presented in Fig. 10. Our approach brought improvements to nearly every tract, particularly for projection pathways heavily impacted by the absence of the top parts of the brain, such as the corticospinal tract (CST). Finally, Bland–Altman plots for examining the bundle shape measurements are presented in Fig. 11. Our approach demonstrates a much more consistent agreement with the reference compared with measurements obtained from images with an incomplete FOV.

Table 4 Average Dice score for 12 tracts that are commonly associated with AD, produced from an image with an incomplete FOV and with its imputation.

	WRAP		NACC	
	Incomplete FOV	With imputation	Incomplete FOV	With imputation
Acquired regions	0.891 ± 0.040	0.920 ± 0.037	0.858 ± 0.059	0.907 ± 0.040
Imputed regions	N/A	0.615 ± 0.281	N/A	0.596 ± 0.256

The improvement of imputation over the incomplete FOV is statistically significant ($p < 0.001$) from paired t -test on AD tracts' results ($p = 0.0006$ for WRAP, $p = 0.00005$ for NACC).

Table 5 Average Dice score for cutoff tracts with ground truth tracts that are cut off by an incomplete FOV and no-cutoff tracts with ground truth tracts that are not cut off by an incomplete FOV in the acquired regions.

	WRAP		NACC	
	Incomplete FOV	With imputation	Incomplete FOV	With imputation
Cutoff tracts	0.910 ± 0.024	0.939 ± 0.010	0.887 ± 0.029	0.926 ± 0.012
No-cutoff tracts	0.905 ± 0.032	0.918 ± 0.0294	0.879 ± 0.048	0.909 ± 0.034

Our approach can improve the tractography accuracy regardless of whether the ground truth tracts are cut off by an incomplete FOV or not. The improvement is statistically significant ($p < 0.001$) for both cutoff tracts ($p = 8.41 \times 10^{-17}$ for WRAP, $p = 6.26 \times 10^{-18}$ for NACC) and for no-cutoff tracts ($p = 5.76 \times 10^{-12}$ for WRAP, $p = 2.06 \times 10^{-8}$ for NACC).

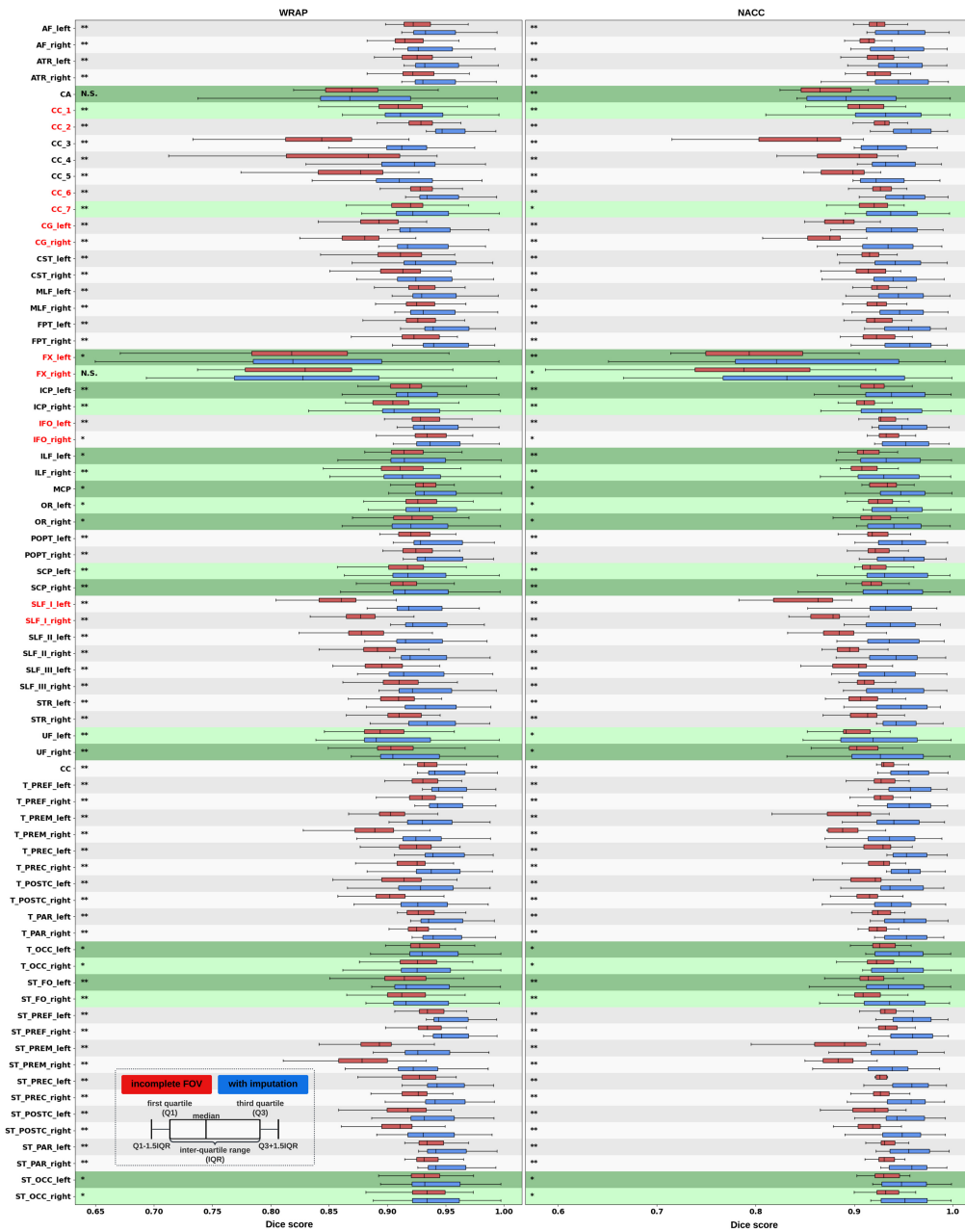


Fig. 10 In both the WRAP and NACC datasets, the proposed framework enhances tractography accuracy through FOV extension (with imputation), as evidenced by the overall higher Dice scores compared with those with an incomplete FOV. Tracts commonly associated with AD have their names in red. Tracts that are not cutoff by an incomplete FOV have green shading in their boxplots. Paired *t*-tests were conducted for each tract, and the statistical significance is denoted by “**” ($p < 0.05$), “***” ($p < 0.01$), and “N.S.” (not significant).

4 Discussion

In the task of imputing missing DWI slices, our framework exhibited a marginally better performance on b_0 images compared with b_{1300} images. This can be attributed to the similarity in patterns between b_0 images and T1-weighted images, which makes their joint distribution simpler for the model to learn. This contrasts with b_{1300} images that require the model to understand additional conditional distributions across various gradient directions. A notable observation was that most imputation errors occurred at the boundary between the white matter and the gray matter. This is likely because our method tends to predict average intensities over the entire

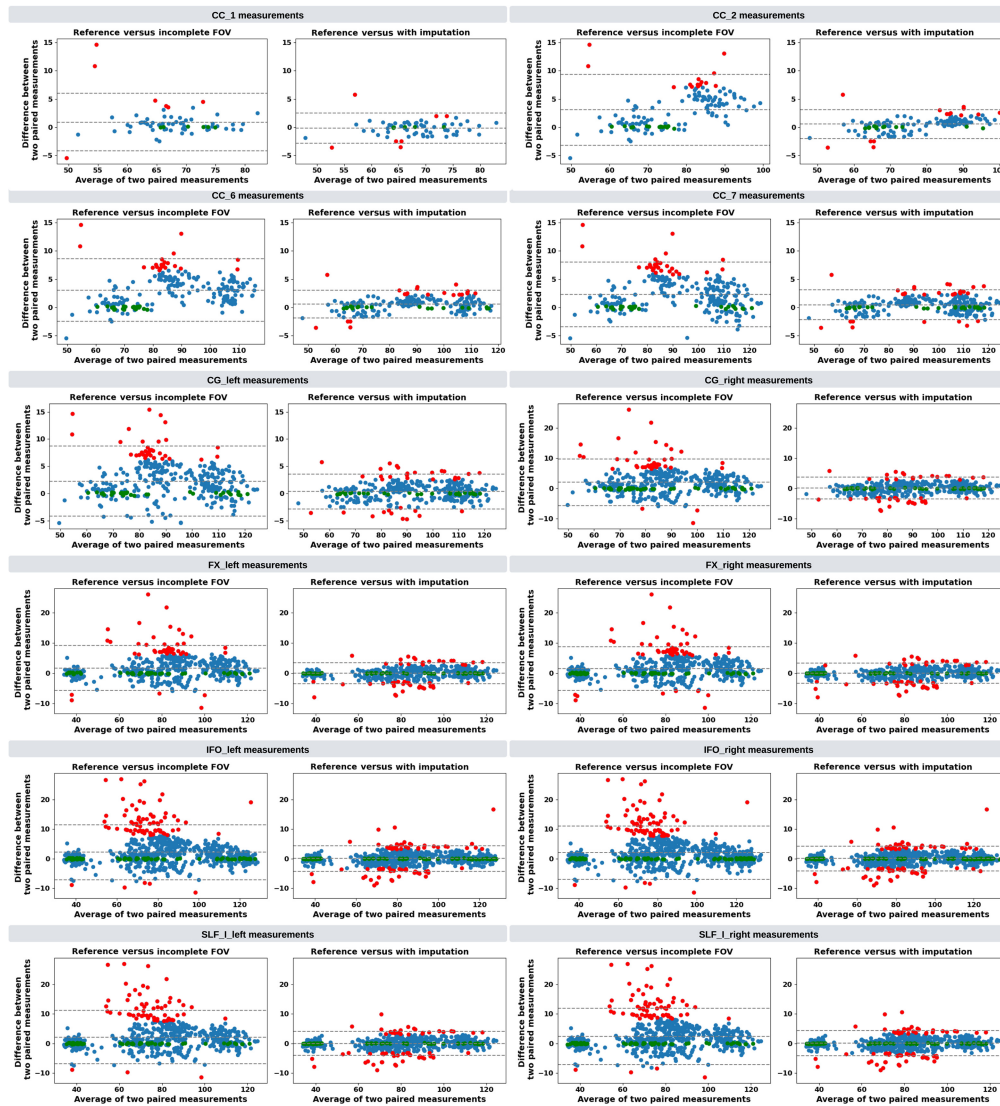


Fig. 11 Bland–Altman plots of the agreement for bundle average length compared with reference. The best 10% measurements with the smallest errors are denoted in green, and the worst 10% measurements with the largest errors are denoted in red. Tracts commonly associated with Alzheimer’s disease (AD) that can be impacted by an incomplete FOV (up to 30 mm from the top of the brain), specifically CC_1, CC_2, CC_6, CC_7, CG, FX, IFO, and SLF_I, are examined. Our approach effectively reduces the significant variations in measurements caused by incomplete FOVs. In the “Reference versus with Imputation” figures, the measurement distribution is tightly clustered and oriented toward the middle dashed line, indicating coherent and consistent agreement with the reference. By contrast, the “Reference versus Incomplete FOV” figures show that the measurements span a large range on the y-axis, suggesting substantial errors and variations. By providing consistent measurements of bundles associated with AD, our method can reduce the uncertainty in AD studies that may contain corrupted data due to an incomplete FOV.

image, which compromises its ability to synthesize sharp intensity contrasts in these areas. In addition, our method faces greater challenges in imputing slices at the brain’s edges. This is evident from the dramatic decrease in PSNR and SSIM when the imputed slice is near the top or bottom of the brain. These imputation challenges therefore affect the tractography results, particularly the difficulties encountered in producing tracts in the same areas.

The comparison of the baseline model with the ablation of T1w image inputs (Table 2) confirms our hypothesis that T1w images are useful for multi-volume DWI imputation. In addition, we noticed that the performance decreases are much larger for the $b1300$ images than the

b_0 images. This finding supports our motivation that the anatomical information contained in T1w images provides a useful reference for imputing DWI across various directions of water diffusion. It further strengthens the contribution of the proposed framework, which learns to integrate features from both T1w images and multi-volume dMRI scans.

It is noteworthy that our approach enhances both the cutoff and no-cutoff tracts. This improvement likely stems from the critical role of whole-brain information in tractography methods. Our findings reinforce the idea that imputing the brain scans in the incomplete part of the FOV can enhance whole-brain tractography and bundle analyses. Consequently, this method holds promise for reducing uncertainty in clinical practice by effectively repairing corrupted data.

5 Conclusion

Completing the missing dMRI data is a crucial task for conducting valuable but time-consuming dMRI scans. In this work, we introduced the first method to solve the FOV extension task for DWI. Our framework successfully imputed missing slices in corrupted DWI with an incomplete FOV, leveraging information from both diffusion-weighted and T1-weighted images. We evaluated the imputation performance qualitatively and quantitatively on both b_0 and b_{1300} DWI volumes on the WRAP and NACC datasets. The results demonstrated that our model not only effectively imputed the missing DWI slices but also improved subsequent tractography tasks. Most notably, the enhanced accuracy and completeness of tractography and bundle analyses, facilitated by our approach in both imputed and observed regions, underscore the substantial potential in effectively repairing corrupted dMRI data. Future research may focus on advancing the generative model to learn features conditioned on the diffusion signal attenuation ratio S/S_0 .

6 Appendix

6.1 Training Graphs

The training graphs for the b_0 model and b_{1300} model are presented in Fig. 12.

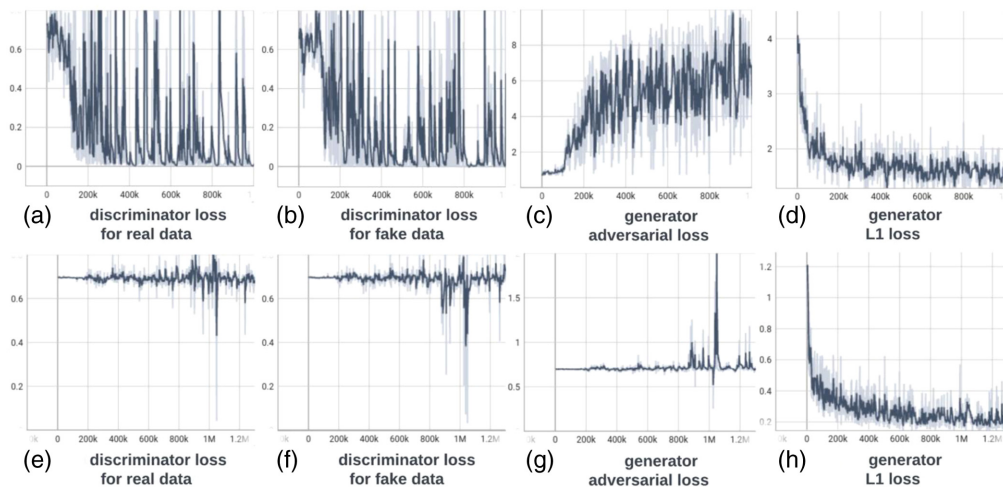


Fig. 12 Training graphs for b_0 model (a)–(d) and b_{1300} model (e)–(h). During training, the losses of the discriminator and generator balance each other, demonstrating stable training for their min-max game. The L1 loss of the generator shows a clear decreasing trend and eventually converges.

Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.

Code and Data Availability

Code can be shared upon request. The data were used under agreement for this study and are therefore not publicly available. More information about the datasets can be found at NACC (<https://www.naccdata.org/>) and WRAP (<https://wrap.wisc.edu/>).

Acknowledgments

This research was supported by NSF CAREER (Grant No. 1452485), National Institutes of Health (Grant No. 1R01EB017230), National Institutes of Health NIDDK (Grant No. K01-EB032898), and NIA (U24AG074855). This study was supported in part using the resources of the Advanced Computing Center for Research and Education (ACCRE) at Vanderbilt University, Nashville, Tennessee, United States (National Institutes of Health S10OD023680). We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Quadro RTX 5000 GPU used for this research. The imaging datasets used for this research were obtained with the support of ImageVU, a research resource supported by the Vanderbilt Institute for Clinical and Translational Research (VICTR), and Vanderbilt University Medical Center institutional funding. The VICTR is funded by the National Center for Advancing Translational Sciences (NCATS) Clinical Translational Science Award (CTSA) Program (Award No. 5UL1TR002243-03). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The ADSP Phenotype Harmonization Consortium (ADSP-PHC) is funded by NIA (Grant Nos. U24 AG074855, U01 AG068057, and R01 AG059716). The harmonized cohorts within the ADSP-PHC included in this paper were the National Alzheimer's Coordinating Center (NACC). The NACC database is funded by NIA/National Institutes of Health (Grant No. U24 AG072122). NACC data are contributed by the NIA-funded ADCs: P30 AG062429 (PI James Brewer, MD, PhD), P30 AG066468 (PI Oscar Lopez, MD), P30 AG062421 (PI Bradley Hyman, MD, PhD), P30 AG066509 (PI Thomas Grabowski, MD), P30 AG066514 (PI Mary Sano, PhD), P30 AG066530 (PI Helena Chui, MD), P30 AG066507 (PI Marilyn Albert, PhD), P30 AG066444 (PI John Morris, MD), P30 AG066518 (PI Jeffrey Kaye, MD), P30 AG066512 (PI Thomas Wisniewski, MD), P30 AG066462 (PI Scott Small, MD), P30 AG072979 (PI David Wolk, MD), P30 AG072972 (PI Charles DeCarli, MD), P30 AG072976 (PI Andrew Saykin, PsyD), P30 AG072975 (PI David Bennett, MD), P30 AG072978 (PI Neil Kowall, MD), P30 AG072977 (PI Robert Vassar, PhD), P30 AG066519 (PI Frank LaFerla, PhD), P30 AG062677 (PI Ronald Petersen, MD, PhD), P30 AG079280 (PI Eric Reiman, MD), P30 AG062422 (PI Gil Rabinovici, MD), P30 AG066511 (PI Allan Levey, MD, PhD), P30 AG072946 (PI Linda Van Eldik, PhD), P30 AG062715 (PI Sanjay Asthana, MD, FRCP), P30 AG072973 (PI Russell Swerdlow, MD), P30 AG066506 (PI Todd Golde, MD, PhD), P30 AG066508 (PI Stephen Strittmatter, MD, PhD), P30 AG066515 (PI Victor Henderson, MD, MS), P30 AG072947 (PI Suzanne Craft, PhD), P30 AG072931 (PI Henry Paulson, MD, PhD), P30 AG066546 (PI Sudha Seshadri, MD), P20 AG068024 (PI Erik Roberson, MD, PhD), P20 AG068053 (PI Justin Miller, PhD), P20 AG068077 (PI Gary Rosenberg, MD), P20 AG068082 (PI Angela Jefferson, PhD), P30 AG072958 (PI Heather Whitson, MD), and P30 AG072959 (PI James Leverenz, MD); National Institute on Aging Alzheimer's Disease Family Based Study (NIA-AD FBS): U24 AG056270; Religious Orders Study (ROS): P30AG10161, R01AG15819, and R01AG42210; Memory and Aging Project (MAP - Rush): R01AG017917 and R01AG42210; Minority Aging Research Study (MARS): R01AG22018 and R01AG42210; Washington Heights/Inwood Columbia Aging Project (WHICAP): RF1 AG054023; and Wisconsin Registry for Alzheimer's Prevention (WRAP): R01AG027161 and R01AG054047. Additional acknowledgments include the National Institute on Aging Genetics of Alzheimer's Disease Data Storage Site (NIAGADS, U24AG041689) at the University of Pennsylvania, funded by NIA.

References

1. P. J. Basser, J. Mattiello, and D. LeBihan, "MR diffusion tensor spectroscopy and imaging," *Biophys. J.* **66**(1), 259–267 (1994).
2. D. C. Alexander et al., "Imaging brain microstructure with diffusion MRI: practicality and applications," *NMR Biomed.* **32**(4), e3841 (2019).
3. P. J. Basser and C. Pierpaoli, "Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI," *J. Magn. Reson.* **213**(2), 560–570 (2011).
4. H. Johansen-Berg and T. E. J. Behrens, *Diffusion MRI: From Quantitative Measurement to In Vivo Neuroanatomy*, Academic Press (2013).
5. C. Pierpaoli et al., "Diffusion tensor MR imaging of the human brain," *Radiology* **201**(3), 637–648 (1996).
6. B. Jeurissen et al., "Diffusion MRI fiber tractography of the brain," *NMR Biomed.* **32**(4), e3785 (2019).

7. J. Y.-M. Yang et al., “Diffusion MRI tractography for neurosurgery: the basics, current state, technical reliability and challenges,” *Phys. Med. Biol.* **66**(15), 15TR01 (2021).
8. E. L. Dennis et al., “Changes in anatomical brain connectivity between ages 12 and 30: a HARDI study of 467 adolescents and adults,” in *9th IEEE Int. Symp. Biomed. Imaging (ISBI)*, pp. 904–907 (2012).
9. N. Nagaraja et al., “Reversible diffusion-weighted imaging lesions in acute ischemic stroke: a systematic review,” *Neurology* **94**(13), 571–587 (2020).
10. S. Cetin-Karayumak et al., “White matter abnormalities across the lifespan of schizophrenia: a harmonized multi-site diffusion MRI study,” *Mol. Psychiatry* **25**(12), 3208–3219 (2020).
11. J. R. Harrison et al., “Imaging Alzheimer’s genetic risk using diffusion MRI: a systematic review,” *NeuroImage Clin.* **27**, 102359 (2020).
12. H. Ni et al., “Effects of number of diffusion gradient directions on derived diffusion tensor imaging indices in human brain,” *Amer. J. Neuroradiol.* **27**(8), 1776–1781 (2006).
13. G. L. Baum et al., “The impact of in-scanner head motion on structural connectivity derived from diffusion MRI,” *NeuroImage* **173**, 275–286 (2018).
14. D. Le Bihan et al., “Artifacts and pitfalls in diffusion MRI,” *J. Magn. Reson. Imaging: Off. J. Int. Soc. Magn. Reson. Med.* **24**(3), 478–488 (2006).
15. D. K. Jones, *Diffusion MRI: Theory, Methods, and Application*, Oxford University Press (2010).
16. C. B. Lauzon et al., “Simultaneous analysis and quality assurance for diffusion tensor imaging,” *PLoS One* **8**(4), e61737 (2013).
17. J. G. Ibrahim and G. Molenberghs, “Missing data methods in longitudinal studies: a review,” *Test* **18**(1), 1–43 (2009).
18. C. Yuan et al., “ReMiND: recovery of missing neuroimaging using diffusion models with application to Alzheimer’s disease,” medRxiv, Cold Spring Harbor Laboratory Preprints (2023).
19. L.-C. Chang, D. K. Jones, and C. Pierpaoli, “RESTORE: robust estimation of tensors by outlier rejection,” *Magn. Reson. Med.: Off. J. Int. Soc. Magn. Reson. Med.* **53**(5), 1088–1095 (2005).
20. Z. Tang et al., “TW-BAG: tensor-wise brain-aware gate network for inpainting disrupted diffusion tensor imaging,” in *Int. Conf. Digit. Image Comput.: Tech. and Appl. (DICTA)*, pp. 1–8 (2022).
21. J. H. Jensen and J. A. Helpert, “MRI quantification of non-Gaussian water diffusion by kurtosis analysis,” *NMR Biomed.* **23**(7), 698–710 (2010).
22. C. M. W. Tax et al., “REKINDLE: robust extraction of kurtosis INDices with linear estimation,” *Magn. Reson. Med.* **73**(2), 794–808 (2015).
23. J. L. R. Andersson et al., “Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images,” *NeuroImage* **141**, 556–572 (2016).
24. A. Koch et al., “SHORE-based detection and imputation of dropout in diffusion MRI,” *Magn. Reson. Med.* **82**(6), 2286–2298 (2019).
25. K. G. Schilling et al., “Synthesized b0 for diffusion distortion correction (Synb0-DisCo),” *Magn. Reson. Imaging* **64**, 62–70 (2019).
26. T. Xiang et al., “DDM²: self-supervised diffusion MRI denoising with generative diffusion models,” in *ICLR* (2023).
27. S. Fadnavis, J. Batson, and E. Garyfallidis, “Patch2Self: denoising diffusion MRI with self-supervised learning,” in *Adv. Neural Inf. Process. Syst.*, Vol. 33, pp. 16293–16303 (2020).
28. Q. Tian et al., “SDnDTI: self-supervised deep learning-based denoising for diffusion tensor MRI,” *NeuroImage* **253**, 119033 (2022).
29. F. Zhang, W. M. Wells, and L. J. O’Donnell, “Deep diffusion MRI registration (DDMReg): a deep learning method for diffusion MRI registration,” *IEEE Trans. Med. Imaging* **41**(6), 1454–1467 (2021).
30. B. Li et al., “Longitudinal diffusion MRI analysis using Segis-Net: a single-step deep-learning framework for simultaneous segmentation and registration,” *NeuroImage* **235**, 118004 (2021).
31. K. Kim et al., “Painting outside as inside: edge guided image inpainting via bidirectional rearrangement with progressive step learning,” in *Proc. IEEE/CVF Winter Conf. Appl. of Comput. Vis.*, pp. 2122–2130 (2021).
32. S. Zhang et al., “Continuous-multiple image inpainting in one-step via positional query and a diffusion-based approach,” in *Int. Conf. Learn. Represent.* (2024).
33. Y.-C. Cheng et al., “InOut: diverse image inpainting via GAN inversion,” in *Proc. IEEE/CVF Conf. Comput. Vis. and Pattern Recognit.*, pp. 11431–11440 (2022).
34. Q. Xiao, G. Li, and Q. Chen, “Image inpainting: hallucinating beyond the image,” *IEEE Access* **8**, 173576–173583 (2020).
35. Z. Tang et al., “High angular diffusion tensor imaging estimation from minimal evenly distributed diffusion gradient directions,” *Front. Radiol.* **3**, 1238566 (2023).
36. L. Y. Cai et al., “Convolutional-recurrent neural networks approximate diffusion tractography from T1-weighted MRI and associated anatomical context,” bioRxiv, Cold Spring Harbor Laboratory Preprints (2023).

37. M. A. Sager, B. Hermann, and A. La Rue, "Middle-aged children of persons with Alzheimer's disease: APOE genotypes and cognitive function in the Wisconsin Registry for Alzheimer's Prevention," *J. Geriatr. Psychiatry Neurol.* **18**(4), 245–249 (2005).
38. S. Weintraub et al., "Version 3 of the Alzheimer Disease Centers' neuropsychological test battery in the Uniform Data Set (UDS)," *Alzheimer Dis. Assoc. Disord.* **32**(1), 10 (2018).
39. L. Y. Cai et al., "PreQual: an automated pipeline for integrated preprocessing and quality assurance of diffusion weighted MRI images," *Magn. Reson. Med.* **86**(1), 456–470 (2021).
40. M. Jenkinson et al., "FSL," *NeuroImage* **62**(2), 782–790 (2012).
41. P. Isola et al., "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. and Pattern Recognit.*, pp. 1125–1134 (2017).
42. D. Pathak et al., "Context encoders: feature learning by inpainting," in *Proc. IEEE Conf. Comput. Vis. and Pattern Recognit.*, pp. 2536–2544 (2016).
43. Y. Liu et al., "Generalizing deep learning brain segmentation for skull removal and intracranial measurements," *Magn. Reson. Imaging* **88**, 44–52 (2022).
44. J. Wasserthal, P. Neher, and K. H. Maier-Hein, "TractSeg-Fast and accurate white matter tract segmentation," *NeuroImage* **183**, 239–253 (2018).
45. O. A. Williams et al., "Vascular burden and APOE 4 are associated with white matter microstructural decline in cognitively normal older adults," *NeuroImage* **188**, 572–583 (2019).
46. P. Neher, D. Hirjak, and K. Maier-Hein, "Radiomic tractometry reveals tract-specific imaging biomarkers in white matter," *Nat. Commun.* **15**(1), 303 (2023).
47. Y. Yang et al., "White matter microstructural metrics are sensitively associated with clinical staging in Alzheimer's disease," *Alzheimer's Dement.: Diagn. Assess. Dis. Monit.* **15**(2), e12425 (2023).
48. M. Bozzali et al., "Damage to the cingulum contributes to Alzheimer's disease pathophysiology by deafferentation mechanism," *Hum. Brain Mapp.* **33**(6), 1295–1308 (2012).
49. C. E. Sexton et al., "A meta-analysis of diffusion tensor imaging in mild cognitive impairment and Alzheimer's disease," *Neurobiol. Aging* **32**(12), 2322.e5–2322.e18 (2011).
50. D. B. Archer et al., "Development of a transcallosal tractography template and its application to dementia," *NeuroImage* **200**, 302–312 (2019).
51. T. M. Nir et al., "Effectiveness of regional DTI measures in distinguishing Alzheimer's disease, MCI, and normal aging," *NeuroImage Clin.* **3**, 180–195 (2013).
52. J. L. da Rocha et al., "Fractional anisotropy changes in parahippocampal cingulum due to Alzheimer's disease," *Sci. Rep.* **10**(1), 2660 (2020).
53. T. M. Schouten et al., "Individual classification of Alzheimer's disease with diffusion magnetic resonance imaging," *NeuroImage* **152**, 476–481 (2017).
54. M. Dumont et al., "Free water in white matter differentiates MCI and AD from control subjects," *Front. Aging Neurosci.* **11**, 270 (2019).
55. C. Metzler-Baddeley et al., "CSF contamination contributes to apparent microstructural alterations in mild cognitive impairment," *NeuroImage* **92**, 27–35 (2014).
56. N. H. Stricker et al., "Decreased white matter integrity in late-myelinating fiber pathways in Alzheimer's disease supports retrogenesis," *NeuroImage* **45**(1), 10–16 (2009).
57. M. Bergamino, R. R. Walsh, and A. M. Stokes, "Free-water diffusion tensor imaging improves the accuracy and sensitivity of white matter analysis in Alzheimer's disease," *Sci. Rep.* **11**(1), 6990 (2021).
58. D. B. Archer et al., "The relationship between white matter microstructure and self-perceived cognitive decline," *NeuroImage Clin.* **32**, 102794 (2021).
59. D. B. Archer et al., "Free-water metrics in medial temporal lobe white matter tract projections relate to longitudinal cognitive decline," *Neurobiol. Aging* **94**, 15–23 (2020).

Biographies of the authors are not available.