Joint Multi-Modal Learning for Susceptibility-Induced Distortion Correction in Diffusion MRI

Jianhui Feng[1], Xiang Li[1,2], and Yuchuan Qiao[1(✉)]

[1] Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China
YuchuanQiao@fudan.edu.cn
[2] College of Computer Science and Artificial Intelligence

Abstract. Susceptibility-induced distortions in diffusion MRI (dMRI) have significantly hindered the study of human brain pathways. Despite the availability of various B0-image-based and FOD-based distortion correction methods, significant distortions persist in certain brain regions, such as the brainstem. In this paper, we present a novel unsupervised learning-based distortion correction framework, named Multi-Modal Distortion Correction Network (MMDC-Net). To aggregate the complementary information in B0 and FOD images, we model distortion correction as a multi-modal optimization problem. We further establish a multi-modal learning paradigm for distortion correction, including modality-aware attention, intensity correction and gradient optimization. First, a Modality-Aware Attention Module (MAM) is designed to extract correction-related features from each modality. We introduce a differentiable intensity correction unit to correct the intensity bias in B0 images. Finally, PCGrad is integrated to mitigate gradient conflicts when optimizing similarity losses across the two modalities. Our method was evaluated on the HCP dataset with RL-LR phase encoding and the HCP-Aging and HCLV datasets with AP-PA phase encoding. The results demonstrated that MMDC-Net achieved superior and robust distortion correction performance compared to other state-of-the-art methods.

Keywords: Susceptibility-induced distortion correction · Diffusion MRI · Unsupervised deep learning

1 Introduction

Diffusion magnetic resonance imaging (dMRI), a non-invasive technique, captures water molecule diffusion patterns in biological tissues and is widely used for human brain white matter fiber tracking [10]. However, the echo-planar imaging (EPI) sequence, commonly used in dMRI, is prone to severe geometric and intensity distortions along the phase encoding (PE) direction due to field inhomogeneities and low bandwidth [9,4]. These susceptibility-induced distortions typically occur at tissue-cavity boundaries, such as the brainstem and orbitofrontal regions [12,11].

A common way to correct susceptibility-induced distortion is to estimate an undistorted image by acquiring two images with opposite PE directions, known as the "blip-up blip-down" method. Among these methods, the `topup` method in the FSL toolbox [1] uses a discrete cosine basis function to estimate a smooth displacement field from two B0 images. In recent years, with advancements in deep learning, several learning-based methods have been proposed to estimate a displacement field using B0 images [17,6,3,18].

By leveraging the structural information in B0 images, B0-based methods can effectively reduce susceptibility-induced distortions in dMRI. However, their correction performance remains limited in certain regions due to insufficient local information. To address this issue, Qiao et al. [13] proposed `FODReg`, a method that uses fiber orientation distribution (FOD) images to estimate residual distortions, and subsequently developed a learning-based approach, `DrC-Net` [12]. While these FOD-based methods improve regional correction, both rely on a cumbersome two-step pipeline using `topup` for preprocessing. This not only leads to a complex correction workflow but also fails to exploit the complementary information between B0 and FOD images.

To further improve the efficiency and accuracy of distortion correction, a natural solution is to aggregate information from both B0 and FOD images into a unified, one-step correction pipeline. In this paper, we propose a novel unsupervised learning-based framework named the Multi-Modal Distortion Correction Network (`MMDC-Net`). We formulate the distortion correction problem as a multi-modal optimization problem, aiming to minimize the discrepancy between the opposing phase-encoding (PE) directions of B0 and FOD images using a single, shared deformation field. Furthermore, we develop a novel multi-modal learning paradigm for distortion correction, which incorporates modality-aware attention, intensity correction, and gradient optimization. Specifically, we first introduce a Modality-Aware Attention Module (MAM) to adaptively emphasize the most informative multi-modal features. Subsequently, a differentiable intensity correction unit is employed to correct the intensity bias in B0 images, ensuring consistent optimization directions across modalities. Finally, PCGrad is introduced to mitigate gradient conflicts arising from multi-modal similarity losses. Experimental results demonstrate that `MMDC-Net` achieves superior and more robust distortion correction performance compared to state-of-the-art methods.

2 Methods

2.1 Susceptibility-Induced Distortion Modeling

In the following, the right-to-left (R/L) or posterior-to-anterior (P/A) PE B0 and FOD images are referred to as positive PE images I^+, and the opposite PE as negative PE images I^-. The fundamental assumption underlying this framework is that the distortion field of each PE image has the same magnitude but opposite direction [1]. Therefore, the distortion correction problem can be formulated by minimizing the difference between the two deformed images:

$$\hat{\phi} = \arg\min_{\phi} \left\{ \mathcal{L}_{sim}(\boldsymbol{I}^{+} \circ \phi, \boldsymbol{I}^{-} \circ \phi^{-1}) \right\}, \quad (1)$$

where $\hat{\phi}$ is the optimal deformation field, $\boldsymbol{I}^{+} \circ \phi$ and $\boldsymbol{I}^{-} \circ \phi^{-1}$ are the deformed images, $\mathcal{L}_{sim}(\cdot, \cdot)$ is the similarity measurement. Existing distortion correction methods only use inputs from one modality for optimization, such as B0 or FOD images. However, a single-modal optimization objective could learn an incorrect deformation in regions of missing contrast, resulting in the optimal deformation field $\hat{\phi}$ failing to represent the true distortion field. To fully utilize the B0 and FOD images, we model the distortion correction problem as minimizing the difference between positive PE and negative PE of B0 and FOD images with the same deformation field. Therefore, the multi-modal optimization objective can be expressed as:

$$\hat{\phi} = \arg\min_{\phi} \left\{ \mathcal{L}_{sim}(\boldsymbol{I}^{+}_{B0} \circ \phi, \boldsymbol{I}^{-}_{B0} \circ \phi^{-1}) + \mathcal{L}_{sim}(\boldsymbol{I}^{+}_{FOD} \circ \phi, \boldsymbol{I}^{-}_{FOD} \circ \phi^{-1}) \right\}, \quad (2)$$

where $\hat{\phi}$ is the optimal distortion field learned from both B0 and FOD images, $\boldsymbol{I}^{+}_{B0} \circ \phi$ and $\boldsymbol{I}^{-}_{B0} \circ \phi^{-1}$ are the deformed B0 images, $\boldsymbol{I}^{+}_{FOD} \circ \phi$ and $\boldsymbol{I}^{-}_{FOD} \circ \phi^{-1}$ are the deformed FOD images.

2.2 Network Architecture

The details of our framework are illustrated in Fig. 1(a). First, we acquire dMRI data in both PEs and perform eddy current correction using the FSL toolbox [2]. Subsequently, we estimate the FOD image from dMRI using the method based on a spherical deconvolution framework [14]. The number of FOD components is set to 6, which has been proven to be adequate [13,12].

Given the significant contrast differences in the FOD and B0 images, our encoder is designed to separately extract features from each modality. First, B0 and FOD images from positive and negative PE are concatenated into 2-channel and 12-channel inputs and fed into two encoder branches. As shown in Fig. 1(a), each convolution block in the encoder consists of a convolution layer with a kernel size of $3 \times 3 \times 3$ and a stride of 2, as well as a LeakyReLU activation layer with a negative slope of 0.2.

A rigid transformation head is designed to correct the linear misalignment between two PEs. Specifically, a MLP layer maps the learned feature from the B0 and FOD images to a set of parameters: $\boldsymbol{\mu} = (\theta_x, \theta_y, \theta_z, t_x, t_y, t_z)$, where θ and t represent the rotation and translation along each respective axis. Then, the rigid transformation matrix $\mathcal{M} = \mathcal{R}_x \circ \mathcal{R}_y \circ \mathcal{R}_z \circ \mathcal{T}$ can be calculated from μ, and converted into a deformation field $\boldsymbol{\phi}_r$. Finally, through the Spatial Transformer Network (STN) [8], the deformation field $\boldsymbol{\phi}_r$ is applied to the negative PE images to achieve rigid alignment of the subject in both PEs.

The B0 and FOD images in opposite PEs contain complementary features for correction. However, according to the prior knowledge, the contribution of

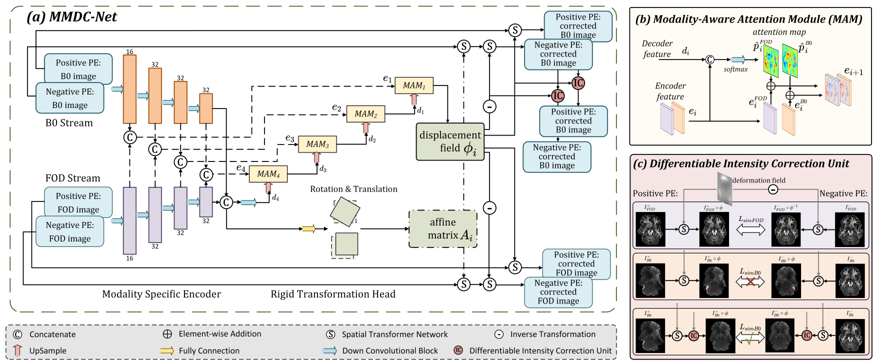

Fig. 1. The proposed pipeline to correct susceptibility-induced distortion in dMRI data. (a) Detail of the proposed `MMDC-Net`. (b) Detail of the Modality-Aware Attention Module. (c) Illustration of the Differentiable Intensity Correction Unit.

different modalities to distortion correction performance varies across brain regions. For example, in the brainstem regions, FOD images provide more relevant information, while B0 images contribute limited information due to the lack of anatomical details. Therefore, it is essential to disentangle the multi-modal features from the encoder into correction-related and correction-agnostic features, which allows the model to focus on the most informative features for correction. Through the layer-wise integration of the Modality-Aware Attention Module (MAM), our decoder adaptively emphasizes correction-relevant features across modalities.

As shown in Fig. 1(b), our MAM module first captures features of the encoder and decoder and learns a 2-channel attention map. Specifically, we reduce the channels of the concatenated features to 2 using a convolutional layer. Then, they are normalized by applying a softmax layer over channel dimensions. The 2-channel attention map $\hat{p}_i$ reflects the contribution between modality in each region to distortion correction, highlighting the importance of correction-related characteristics and decreasing the attention to correction-agnostic characteristics. The output feature e_{i+1}^m can be obtained by multiplying the features with the attention map. In the final stage of the decoder, a fine convolution layer is used to generate a 3D distortion field. Since susceptibility-induced distortion occurs along the PE direction, the distortion field is constrained to be one-dimensional along the PE direction. Finally, the learned distortion field $\boldsymbol{\phi}_d$ is applied to both PE directions.

2.3 Differentiable Intensity Correction Unit

As shown in Fig. 1(c), the B0 images show significant intensity bias due to susceptibility-induced distortion, which becomes more severe after transformation. This leads to intensity inconsistencies between the two PEs and thus impacts the similarity calculation. Conversely, FOD images show no intensity bias.

Therefore, to ensure consistent optimization for both modalities, it is essential to correct the intensity distortion in B0 images.

Inspired by Jacobian modulation in conventional methods [1], we propose a differentiable Intensity Correction (IC) Unit that calculates compensation factors based on distortion field shifts. It reduces the intensity in signal build-up regions and increases it in signal loss regions. Using the RL-LR direction distortion as an example, we compute the coordinate matrices u_i^+ and u_i^- in PE direction:

$$\begin{aligned} u_i^+(\phi) &= \phi(x_i+1, y_i, z_i) - \phi(x_i, y_i, z_i), \\ u_i^-(\phi) &= \phi(x_i, y_i, z_i) - \phi(x_i-1, y_i, z_i), \end{aligned} \tag{3}$$

where ϕ is the distortion field. The compensation factor J_i is computed as:

$$J_i(\phi) = 1 + \frac{u_i^+(\phi) - u_i^-(\phi)}{2}, \tag{4}$$

where $i \in \Omega$ and Ω is the whole domain of the image. The corrected image is obtained by applying the compensation factor J to the B0 image after transformation: $\boldsymbol{I}_{B0}^+ = J(\phi_d) \cdot (\boldsymbol{I}_{B0}^+ \circ \phi_d)$ and $\boldsymbol{I}_{B0}^- = J(\phi_d^{-1}) \cdot (\boldsymbol{I}_{B0}^- \circ \phi_d^{-1})$.

2.4 Loss Function and Optimization

As mentioned above, the optimization objective of our proposed method includes the similarity of two corrected B0 images $\mathcal{L}_{simB0}$ and two corrected FOD images $\mathcal{L}_{simFOD}$. Additionally, the regularization term $\mathcal{L}_{reg}$ is used to promote smoothness of the distortion field. The mean squared difference (MSD) is selected as the similarity measure, while the regularization term is defined as the spatial gradient magnitude: $\mathcal{L}_{reg} = \sum_{p \in \Omega} \parallel \nabla \boldsymbol{\phi}(p) \parallel^2$, where p is image voxel position and Ω is the whole domain of the image.

However, due to the contrast and structural differences between different modality-specific image pairs, the gradients of the two similarity loss functions may conflict. To address this, PCGrad [16,19] is used to reduce gradient conflicts between the similarity losses $\mathcal{L}_{simFOD}$ and $\mathcal{L}_{simB0}$. Specifically, we compute the gradient of the two modal losses separately and project one onto the orthogonal plane of the other to eliminate conflicts, ensuring stable optimization.

3 Experimental Setup and Results

3.1 Experimental Setup

Our method was evaluated on three datasets: HCP, HCP-Aging, and HCLV. The Human Connectome Project (HCP) [15] dataset contains raw 3T dMRI scans from RL-LR PEs with an isotropic spatial resolution of 1.25 mm. We randomly split the HCP dataset into 550 scans for training and 50 scans for testing. The Lifespan Human Connectome Project Aging (HCP-Aging) dataset [5] comprises 3T dMRI data (AP-PA PEs) with 1.5 mm isotropic spatial resolution, split into 544 training and 50 testing scans. The Human Connectome for

Low Vision (HCLV) dataset includes 3T dMRI data (AP-PA PEs) with 1.5 *mm* isotropic spatial resolution; all 58 subjects served as test data for generalization. We compared our method with two state-of-the-art conventional methods `topup` [1], `FODReg` [13], and four learning-based methods `S-Net` [6], `flow-net` [17], `FD-Net` [18], `DrC-Net` [12]. For all methods, we used their recommended configurations for training and inference.

Our proposed `MMDC-Net` was trained using the PCGrad optimizer with a learning rate of $1e^{-4}$, a weight parameter λ_{sim} of $1e^{-2}$, and a regularization parameter λ_{reg} of $5e^{-4}$. We trained our model for 500 epochs with 100 steps per epoch. The implementation and testing of our method were carried out on an NVIDIA Tesla V100 SXM2 GPU using the PyTorch backend.

We evaluated the distortion correction performance of our method using two metrics: mean squared difference (MSD) of fractional anisotropy (FA) from the tensor model and minimum angular difference (MAD) between the major fiber directions of opposite PEs. The angular difference is defined as the minimum angle between the main fiber of the positive PE FOD and the top three fibers of the negative PE FOD. The whole white matter (wm) region was used as an ROI to measure the overall correction performance. Additionally, we particularly focused on brain regions with significant residual distortion after correction, including the brainstem, orbitofrontal cortex (ofc), and corpus callosum (cc). We segmented the ROIs from the T1 image using FreeSurfer 6.0 [7], and then linearly registered them to dMRI space for analysis.

3.2 Results

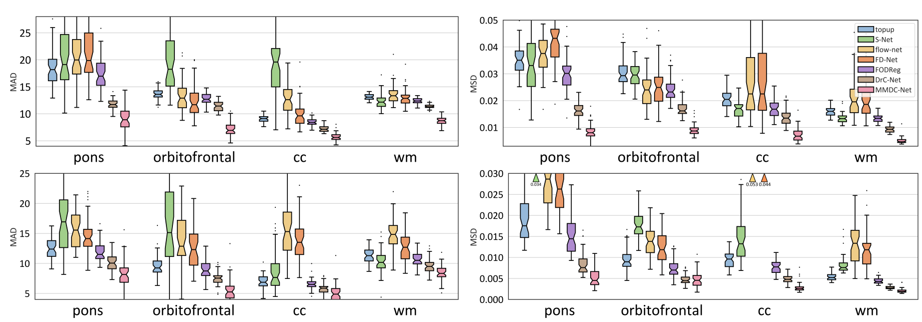

Fig. 2. The quantitative results evaluated on four ROIs. The top row is the result of HCP data and the bottom row is the result of HCP-Aging data. The left box plots show the results on MAD, while the right box plots are MSD results.

Fig. 2 shows the correction results comparing our proposed `MMDC-Net` with other SOTA methods. Besides the MSD in the cc region in the HCP-Aging dataset for `DrC-Net`, the results from `MMDC-Net` were statistically different from those of the other compared methods in terms of a Wilcoxon test ($p < 0.05$). This indicates that our method not only performed superiorly in severely distorted brain regions but also provided excellent correction effects across the entire image. Furthermore, the results demonstrated the ability of our method to correct dMRI data with different phase-encoding directions and different resolutions.

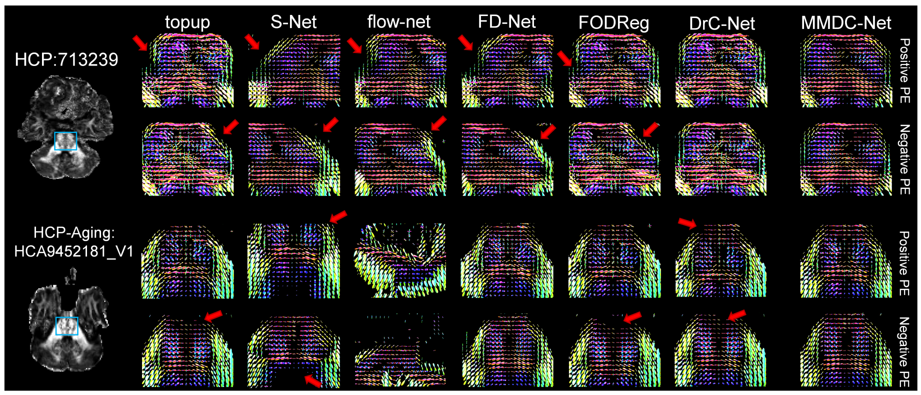

Fig. 3. Distortion correction effects are shown in the FOD images in the brainstem region. The blue box indicates the location of the ROI in the first component of the FOD image.

We visualize the corrected FOD of both PEs in Fig. 3. It can be seen that even after applying existing correction methods, the fiber bundle trajectories generated from the two PEs remain internally irregular, which could lead to incorrect analysis. Specifically, for the HCP example above, the left-right direction offset between the two PEs still exists after correction by existing methods, especially `topup`, `S-Net`, `flow-net`, and `FODReg`. Furthermore, although `DrC-Net` corrects the positional offset, it introduces significant irregularities and noise within the fiber bundles. For the following HCP-Aging example, some methods also suffer from insufficient deformation, especially in the negative PE. In contrast, `MMDC-Net` corrected the distortion of the two PEs completely, resulting in clear and regular fiber bundle trajectories in the brainstem region.

3.3 Generalizability Analysis

To evaluate the generalization of our `MMDC-Net`, we trained all learning-based methods on HCP-Aging data and tested on HCLV data. As shown in Table 1, our method outperformed other learning-based and conventional methods with results comparable to `DrC-Net`, which demonstrates the generalization and robustness of our method.

Table 1. Generalizability analysis for distortion correction methods in HCLV data. For learning-based methods `S-Net`, `flow-net`, `FD-Net`, `DrC-Net` and our `MMDC-Net`, we trained on HCP-Aging data and tested on HCLV data. For conventional methods `topup` and `FODReg`, we used the default parameter settings. **Bold** values indicate the best results.

	MAD ↓				MSD ($\times 10^{-2}$) ↓			
	pons	*ofc*	*cc*	*wm*	*pons*	*ofc*	*cc*	*wm*
`topup`	12.75±1.92	9.67±1.22	7.06±1.06	10.98±0.44	2.18±0.58	1.08±0.21	0.98±0.21	0.67±0.09
`FODReg`	11.82±2.14	9.10±1.23	6.24±0.97	7.20±0.39	1.84±0.53	0.85±0.16	0.85±0.17	0.55±0.06
`S-Net`	21.61±11.04	18.73±6.31	8.50±4.07	10.72±0.72	5.20±1.38	3.94±0.51	6.03±1.05	3.53±0.16
`flow-net`	16.36±3.89	13.19±4.11	9.73±3.55	14.05±1.65	2.72±0.70	1.49±0.51	6.45±2.44	1.32±0.41
`FD-Net`	15.87±3.34	11.98±3.37	8.89±2.97	12.74±1.72	2.65±0.68	1.41±0.44	5.57±2.31	1.28±0.37
`DrC-Net`	9.88±1.20	7.21±0.79	6.67±0.99	**6.40**±0.37	0.87±0.21	**0.49**±0.08	0.67±0.16	**0.29**±0.04
`MMDC-Net`	**8.85**±1.66	**6.91**±1.28	**6.36**±2.34	6.47±0.37	**0.67**±0.18	0.54±0.12	**0.51**±0.10	0.36±0.03

Table 2. Ablation study of input modality and components of the framework in HCP dataset. The MAD and MSD metrics in different ROIs are reported. **Bold** values indicate the best results. * denotes the baseline model. Components are added cumulatively to the baseline.

	MAD ↓				MSD ($\times 10^{-2}$) ↓			
	pons	*ofc*	*cc*	*wm*	*pons*	*ofc*	*cc*	*wm*
B0*	15.06±3.82	9.94±6.15	8.66±1.95	11.77±0.77	2.6±0.7	2.0±1.0	1.2±0.4	1.3±0.2
FOD*	9.19±2.58	11.17±2.67	6.83±1.80	9.08±0.84	1.6±0.6	3.1±0.5	1.3±0.4	0.9±0.2
(B0+FOD)*	9.87±2.68	10.49±2.15	6.43±1.40	9.11±0.78	1.6±0.6	2.8±0.5	1.1±0.3	0.9±0.2
+MAM	**9.01**±2.18	7.64±1.40	5.86±1.08	8.99±0.94	1.3±0.4	1.4±0.3	0.9±0.4	0.7±0.2
+IC Unit	9.11±2.64	7.54±1.37	5.65±0.90	8.90±0.67	1.2±0.4	1.1±0.3	0.9±0.4	0.7±0.2
+PCGrad	9.04±2.28	**7.11**±1.29	**5.64**±0.80	**8.69**±0.79	**0.9**±0.3	**0.9**±0.2	**0.7**±0.3	**0.5**±0.1

3.4 Ablation study

As shown in Table 2, we investigated the impact of input modalities on distortion correction. For a fair comparison, a baseline was established to accommodate multi-modal inputs, retaining the rigid transformation head and replacing MAM with a series of convolutional blocks. The results show that using only the B0 image partially corrected distortions in the white matter and orbitofrontal regions, while using only the FOD image improved correction in specific regions such as the pons in the brainstem. However, simply concatenating B0 and FOD images did not yield better results, as it did not effectively utilize the complementary information between the two modalities. Our paradigm gradually improved overall correction performance. However, introducing the IC unit increased MAD in the pons region, which was attributed to low B0 image intensity and severe distortion in this region.

4 Discussion and Conclusion

We propose an unsupervised learning-based method to correct susceptibility-induced distortion in dMRI data. Unlike prior works that use only B0 or FOD images, we establish a multi-modal correction paradigm to fully extract complementary information across modalities and ensure stable convergence. Our proposed MAM efficiently extracts the correction-related features and explicitly reveals the contribution of B0 and FOD images for distortion correction. Moreover, IC Unit and PCGrad resolve the multi-modal conflicts during optimization.

Although our proposed method achieves superior correction accuracy, it incurs higher computational overhead compared to B0-only methods due to the required eddy current correction and FOD reconstruction steps. Notably, several fast deep learning-based tools for these preprocessing steps have been developed in recent years, which have significantly reduced this computational burden and made our pipeline feasible for large-scale neuroimaging studies.

We hope that this unified multi-modal framework will offer valuable insights and pave new directions for susceptibility distortion correction in dMRI.

Acknowledgments. Jianhui Feng, Xiang Li and Yuchuan Qiao were supported by the National Natural Science Foundation of China (Grant No. 82102002), the Science and Technology Program of Hunan Province (Grant No. 2026QK3020), Open Funding Project of the Nankai University Optometry and Vision Science Institute (Grant No. NKSGY202412).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Andersson, J.L., Skare, S., Ashburner, J.: How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. Neuroimage **20**(2), 870–888 (2003)
2. Andersson, J.L., Sotiropoulos, S.N.: An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. Neuroimage **125**, 1063–1078 (2016)
3. Bao, Q., Xie, W., Otikovs, M., et al.: Unsupervised cycle-consistent network using restricted subspace field map for removing susceptibility artifacts in EPI. Magnetic Resonance in Medicine (2023)
4. Begnoche, J.P., Schilling, K.G., Boyd, B.D., et al.: EPI susceptibility correction introduces significant differences far from local areas of high distortion. Magnetic Resonance Imaging **92**, 1–9 (2022)
5. Bookheimer, S.Y., Salat, D.H., Terpstra, M., et al.: The lifespan human connectome project in aging: An overview. NeuroImage **185**, 335–348 (January 2019)
6. Duong, S.T., Phung, S.L., Bouzerdoum, A., et al.: An unsupervised deep learning technique for susceptibility artifact correction in reversed phase-encoding EPI images. Magnetic Resonance Imaging **71**, 1–10 (2020)
7. Iglesias, J.E., Van Leemput, K., Bhatt, P., et al.: Bayesian segmentation of brainstem structures in MRI. Neuroimage **113**, 184–195 (2015)
8. Jaderberg, M., Simonyan, K., Zisserman, A., et al.: Spatial transformer networks. Advances in Neural Information Processing Systems **28** (2015)
9. Jezzard, P., Balaban, R.S.: Correction for geometric distortion in echo planar images from B0 field variations. Magnetic Resonance in Medicine **34**(1), 65–73 (1995)

10. Le Bihan, D., Breton, E., Lallemand, D., et al.: MR imaging of intravoxel incoherent motions: Application to diffusion and perfusion in neurologic disorders. Radiology **161**(2), 401–407 (1986)
11. Li, Y., Hou, Y., Li, Q., et al.: Optimized tractography mapping and quantitative evaluation of pyramidal tracts for surgical resection of insular gliomas: A correlative study with diffusion tensor imaging–derived metrics and patient motor strength. Journal of Digital Imaging **35**(2), 356–364 (2022)
12. Qiao, Y., Shi, Y.: Unsupervised deep learning for FOD-based susceptibility distortion correction in diffusion MRI. IEEE Transactions on Medical Imaging **41**(5), 1165–1175 (2022)
13. Qiao, Y., Sun, W., Shi, Y.: FOD-based registration for susceptibility distortion correction in brainstem connectome imaging. Neuroimage **202**, 116164 (2019)
14. Tran, G., Shi, Y.: Fiber orientation and compartment parameter estimation from multi-shell diffusion imaging. IEEE Transactions on Medical Imaging **34**(11), 2320–2332 (2015)
15. Van Essen, D., Ugurbil, K., Auerbach, E., et al.: The human connectome project: A data acquisition perspective. Neuroimage **62**(4), 2222–2231 (2012)
16. Yu, T., Kumar, S., Gupta, A., et al.: Gradient surgery for multi-task learning. Advances in Neural Information Processing Systems **33**, 5824–5836 (2020)
17. Zahneisen, B., Baeumler, K., Zaharchuk, G., et al.: Deep flow-net for EPI distortion estimation. Neuroimage **217**, 116886 (2020)
18. Zaid Alkilani, A., Çukur, T., Saritas, E.U.: FD-Net: An unsupervised deep forward-distortion model for susceptibility artifact correction in EPI. Magnetic Resonance in Medicine **91**(1), 280–296 (2024)
19. Zheng, Y., Bai, Y., Qiao, Y.: GSSD: A self-distillation paradigm with gradient surgery for end-to-end deformable image registration. In: International Conference on Information Processing in Medical Imaging. pp. 64–78. Springer (2025)