Received: 11 November 2022 | Revised: 19 February 2023 | Accepted: 14 March 2023 | Published on: 13 April 2023

DOI: 10.1002/mrm.29653

RESEARCH ARTICLE

Magnetic Resonance in Medicine

Unsupervised cycle-consistent network using restricted subspace field map for removing susceptibility artifacts in EPI

Qingjia Bao[1] | Weida Xie[2] | Martins Otikovs[3] | Liyang Xia[2] | Han Xie[1] | Xinjie Liu[1] | Kewen Liu[2] | Zhi Zhang[1] | Fang Chen[1] | Xin Zhou[1,4,5] | Chaoyang Liu[1,4,5]

[1]Key Laboratory of Magnetic Resonance in Biological Systems, Innovation Academy for Precision Measurement Science and Technology, Wuhan, 430071, People's Republic of China

[2]School of Information Engineering, Wuhan University of Technology, Wuhan, People's Republic of China

[3]Weizmann Institute of Science, Rehovot, 76001, Israel

[4]University of Chinese Academy of Sciences, Beijing, 100049, People's Republic of China

[5]Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology-Optics Valley Laboratory, Hubei, 430074, People's Republic of China

Correspondence
Chaoyang Liu, Key Laboratory of Magnetic Resonance in Biological Systems, Innovation Academy for Precision Measurement Science and Technology, Wuhan, 430071, People's Republic of China.
Email: chyliu@wipm.ac.cn

Funding information
National Key of R&D Program of China, Grant/Award Numbers: 2016YFC1304702, 2018YFC0115000; Chinese Academy of Sciences, Grant/Award Number: YZ201677; National Major Scientific Research Equipment Development Project of China, Grant/Award Number: 81627901; National Natural Science Foundation of China, Grant/Award Numbers: 11705274, 11575287

Purpose: To design an unsupervised deep neural model for correcting susceptibility artifacts in single-shot Echo Planar Imaging (EPI) and evaluate the model for preclinical and clinical applications.

Methods: This work proposes an unsupervised cycle-consistent model based on the restricted subspace field map to take advantage of both the deep learning (DL) and the reverse polarity-gradient (RPG) method for single-shot EPI. The proposed model consists of three main components: (1) DLRPG neural network (DLRPG-net) to obtain field maps based on a pair of images acquired with reversed phase encoding; (2) spin physical model–based modules to obtain the corrected undistorted images based on the learned field map; and (3) cycle-consistency loss between the input images and back-calculated images from each cycle is explored for network training. In addition, the field maps generated by DLRPG-net belong to a restricted subspace, which is a span of predefined cubic splines to ensure the smoothness of the field maps and avoid blurring in the corrected images. This new method is trained and validated on both preclinical and clinical datasets for diffusion MRI.

Results: The proposed network could effectively generate smooth field maps and correct susceptibility artifacts in single-shot EPI. Simulated and in vivo preclinical/clinical experiments demonstrated that our method outperforms the state-of-the-art susceptibility artifact correction methods. Furthermore, the ablation experiments of the cycle-consistent network and the restricted subspace in generating field maps did show the advantages of DLRPG-net.

Conclusion: The proposed method (DLRPG-net) can effectively correct susceptibility artifacts for preclinical and clinical single-shot EPI sequences.

KEYWORDS

cycle-consistent, deep learning, diffusion-weighted images, EPI, subspace field map, susceptibility artifacts

1 | INTRODUCTION

MRI is one of the most widely used imaging modalities for preclinical and clinical applications. Due to its ability to deliver high-definition images in short acquisition time (on the order of 100 ms per slice), single-shot EPI (ssEPI) has become one of the most useful sequences utilized in Magnetic Resonance Imaging (MRI).[1] SsEPI also provides the benefit of avoiding motion artifacts associated with multiple-shot techniques. These factors contribute to its widespread application in Diffusion Tensor Imaging (DTI),[2] functional Magnetic Resonance Imaging (fMRI),[3] and dynamic susceptibility contrast MRI (DSC-MRI).[4] However, the bandwidth of ssEPI along the phase-encoding direction is relatively low, making it sensitive to B_0 field inhomogeneities and leading to pronounced susceptibility artifacts at the boundaries of tissues with differing susceptibilities.[5] The severity of susceptibility artifacts increases with magnetic high field strength.

A variety of ssEPI susceptibility artifact correction methods have been developed, and the most widely used are field mapping[6–8] and gradient reversal methods.[9] The field mapping method corrects distorted images by coordinate calculation and linear interpolation based on a separately acquired B_0 field map,[8] calculated from two gradient-echo (GRE) images acquired using different TE. Whereas it takes only a few seconds for ssEPI to collect images, it can require tens of seconds to a few minutes for GRE-based field map acquisition.[10] Despite its relatively easy implementation, it also has some disadvantages. First, the accuracy of the generated field map is low, which affects the performance of artifact correction, resulting in residual artifacts in the corrected image. Secondly, in the calculation process of field mapping, the existence of various errors will affect the phase unwrapping, making this process more challenging near the tissue boundaries and in the areas with high field inhomogeneity.[11] The gradient reversal methods, instead of using an explicit B_0 field map, infer the distortions using a pair of two EPI scans acquired with the phase-encoding gradients played out with opposite polarities, resulting in two images with equal magnitudes of distortion but of opposite directions. Chang and Fitzpatrick's[12] method corrects artifacts by estimating the displacement of each k-space line in the phase- encoding direction and then using 1D unwrapping to obtain a displacement map. However, due to the limited smoothness of the displacement map, streaking or discontinuities might occur in the corrected image. Andersson et al. proposed a method to fit the displacement field with discrete cosine basis functions, named TOPUP, which improved the smoothness of the field map[13] and formed part of FSL software package (FMRIB Software Library).[14] However, the disadvantage of these traditional methods is that for images with large matrix size, pixel-level iterative solution is time-consuming, and the optimization of nonconvex objective function will face the problem of falling into a local minimum.

Due to its powerful feature extraction ability, deep learning–based methods for correction of susceptibility artifacts in ultrafast MRI images have been implemented by a growing number of researchers, relying on deep learning as a potential way to significantly reduce the computational time and solve the nonconvex optimization problem of field map estimation. The deep learning methods are mainly divided into supervised[15,16] and unsupervised methods.[17,18] The supervised methods need the datasets with ground truth to train the model. Liao et al.[15] corrected distortions present in gradient-echo ssEPI by using convolutional neural network based on the simulated distorted images generated by their homebuilt MRI simulation software. Hu et al.[16] used point-spread-function–encoded EPI images as ground-truth images to generate displacement maps to correct distortions in ssEPI during U-net training. However, due to the errors introduced by phase unwrapping and regularization, the effect of the generated displacement map was not satisfactory. Soan[17] proposed an end-to-end deep learning network by exploring the use of deep convolutional network to estimate the displacement map from a pair of input images. Benjamin[18] trained a deep convolutional U-net architecture that was previously used to estimate optical flow (Flownet[19]) between moving images and to learn distortion map from an input pair of distorted ssEPI images. In our previous work, we proposed a cycle-consistent deep neural network that takes advantage of both the deep neural network and the gradient reversal method for removing susceptibility artifacts in single-shot EPI.[20]

This paper proposes an unsupervised deep learning network that uses a reverse polarity-gradient (RPG) to correct susceptibility artifacts in the reversed phase-encoding ssEPI, named DLRPG-net. Firstly, the deep neural network combined with the restricted subspace technique is applied to generate the displacement map. Specifically, DLRPG-net will try to predict the coefficients of predefined cubic spline vectors,[21,22] and the displacement maps are calculated based on these coefficients. Then, the forward and backward physical model containing both geometric and intensity correction modules is employed to obtain the cycle-consistent loss for the network training. This approach does not require explicit knowledge of the ground truth displacement map. The simulation experiments show that our method outperforms the state-of-the-art traditional and deep learning methods in qualitative and quantitative terms. Experiments on both clinical and preclinical applications demonstrate the generalization capabilities of our method. Furthermore, the ablation experiments show the advantage of the combination of the restricted subspace technique and deep learning

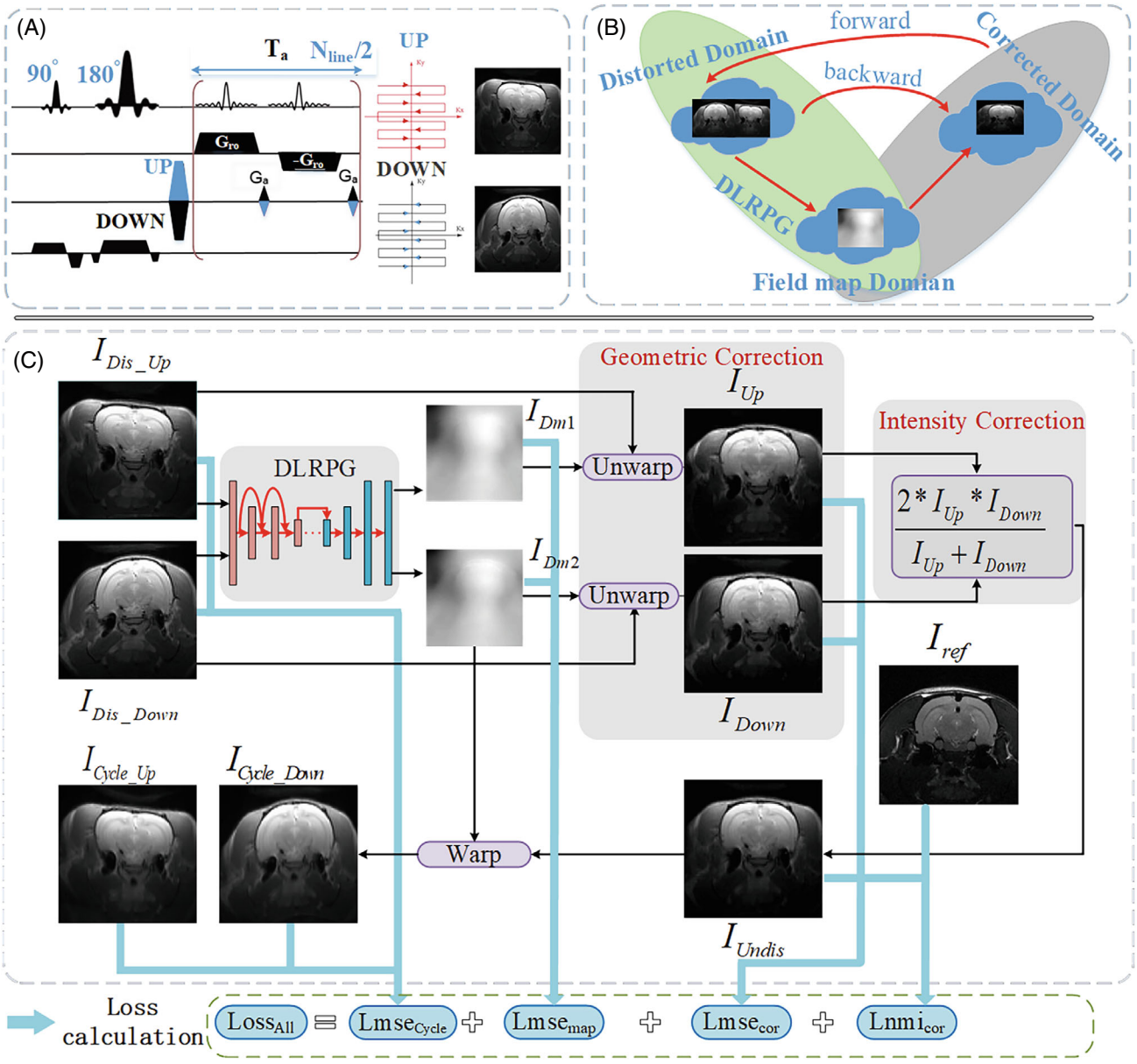

FIGURE 1 (A) The pulse sequence and k-space acquisition trajectory of reversed phase-encoding ssEPI. (B) The cycle-consistency idea of this method. Firstly, the field map is obtained by DLRPG-net based on the pairs of images acquired by the reversed phase-encoding ssEPI. Then, the corrected images can be obtained by the physical backward model (geometric and intensity correction) with the generated field map. Finally, the forward model is applied to obtain the distorted cycle back blip-UP/DOWN images to explore the consistent cycle loss. (C) Detailed procedure of the proposed learning framework for the ssEPI susceptibility artifacts correction. The total loss for training the network is shown at the bottom.

in generating field maps, and the cycle-consistent loss for correcting susceptibility artifacts.

2 | METHODS

2.1 | The overall architecture

The overall architecture of the proposed method is shown in Figure 1. Figure 1(A) shows the pulse sequence of reversed phase-encoding ssEPI. In Figure 1(B), we illustrate the main idea of this method: Firstly, the field map is obtained by the DLRPG-net based on the images acquired with the reversed phase-encoding ssEPI. Then, the corrected images can be obtained by the physical backward model (geometric and intensity correction) using the generated field map. Finally, the forward model is applied to obtain the distorted cycle back blip-UP/DOWN images to calculate the cycle-consistency loss. Figure 1C shows the details of the respective procedures.

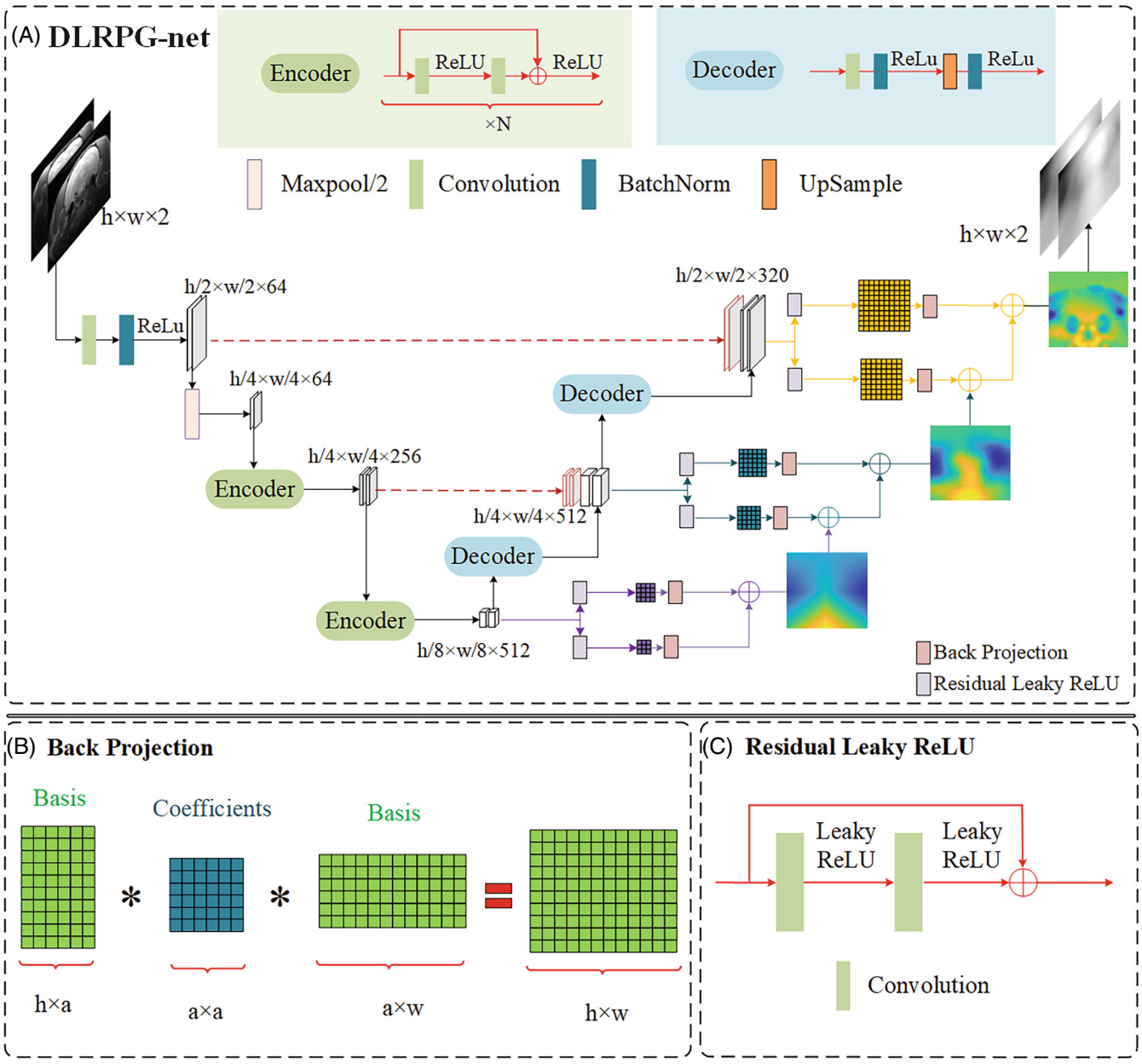

FIGURE 2 (A) The overall architecture of DLRPG-net. (B) The back projection module multiplies the pregenerated bases and the coefficients map generated by the network to obtain the components of the final field map. (C) The residue leaky ReLU is introduced as the activation function to ensure that the output can be positive or negative. ReLU, rectified linear units.

2.2 | DLRPG-net for field map estimation

At the heart of the susceptibility artifact correction lies the accurate estimation of the B_0 field map. However, estimating field maps is not a straightforward problem because the least-squares cost function with respect to field maps is always nonconvex.[22] Neighborhood information is often incorporated into the reconstruction to avoid erroneous field map estimates. Motivated by the work of Tsao and Jiang,[22] we propose a new network (DLRPG-net) that takes both advantages of deep learning and restricts the field map to a subspace. The overall architecture of DLRPG-net is shown in Figure 2. The main idea of the DLRPG-net is to generate the coefficients for the field map in the restricted subspace; namely, these coefficients are expected to be close to the inner product of the field map and the predefined cubic spline vectors.

The architecture of DLRPG-net is similar to the most common network (U-Net[23]). It has a symmetric encoder-decoder architecture, including skip connections, shown in Figure 2. The layers in the encoder part are skip-connected and concatenated with layers in the

decoder part. The skip connections promote U-Net to use fine-grained details learned in the encoder part to construct an image in the decoder part. The encoder part consists of multiple residual blocks and results in a feature map with halved resolution as the input for the next encoder. The residual blocks[24] were used to replace the traditional U-Net feature extraction block. Details of the encoder are shown in the upper left part of Figure 2(A): it contains two 2D convolutional layers, two rectified linear units (ReLU), and a skip connection. The decoder block contains 2D convolutional layers, two batch normalization layers, and an upsampling layer, as shown in Figure 2(A). More details of the DLRPG-net can be found in the Text S1 in Appendix S1. As was mentioned above, with the number of decoding layers increasing, the dimension of the coefficient matrix also increases. The residual leaky ReLU is utilized to convert the feature map to the coefficients, shown in Figure 2C. The leaky ReLU is introduced as the activation function to ensure that the output can be positive or negative. Then, the back projection module is applied to map the deep-learned coefficients to the field map in the image space by multiplying the predefined basis, shown in Figure 2B.

2.3 | Geometric and intensity correction based on DLRPG-net learned field map

In general, the susceptibility artifacts induced by field inhomogeneities in ssEPI are negligible along the slice-selection and frequency-encoding direction as a result of the much higher acquisition or excitation bandwidth. Thus, susceptibility artifacts along phase encoding (PE) in ssEPI can be considered as consisting of two parts: geometric deformations and intensity variations. It is intuitive for the geometric deformation induced by the B_0 inhomogeneities to manifest as a pixel shift along the PE direction. For a pixel in position (x, y), the shift distance along the PE in terms of the number of pixels can be formulated as $u(x,y) = \delta y/\Delta y = \Delta B_0(x,y)t_{\text{esp}}N_y$, where δy is the distortion caused by B_0 inhomogeneity, and Δy indicates the resolution of one pixel. More details on the theoretical background of susceptibility artifacts can be found in Text S2 in Appendix S1. The undistorted spatial images $I_{\text{undis}}(x,y)$ can be obtained based on $u(x,y)$ mapping of the distorted domain to undistorted domain, $I_{\text{undis}}(x,y) \sim I_{\text{dis}}(x, y+u)$. Because the distance of pixel shifts are not necessarily integer multiples of the image resolution, the interpolation of the distorted images is usually needed.[25]

After the geometric deformation correction, the intensity correction should not be neglected. Theoretically, the mean intensity of any subspace should be inversely proportional to its area.[26] Pixel-wise Jacobian modulation has usually been adopted to correct intensity variation in which the contracted areas will experience increased intensity and the dilated areas will experience decreased intensity. More details regarding intensity corrections using Jacobian modulation used in this paper can be found in the Text S3 in Appendix S1. In this work, the optimized Jacobian modulation method is used based on the recent work,[27] and the corrected images can be obtained by the following Equation:

$$I_{\text{undis}}(x,y) = \frac{2I_{\text{dis up}}(x, y+u(x,y)) * I_{\text{dis down}}(x, y-u(x,y))}{I_{\text{dis up}}(x, y+u(x,y)) + I_{\text{dis down}}(x, y-u(x,y))}. \tag{1}$$

If the field map derived by deep learning is exact, the intensities in the image after correction by Jacobian modulation J_{field} and "true" image pixel intensities will be the same. However, it is difficult to obtain an accurate Jacobian matrix. Because the intensity accumulation and corresponding intensity dispersion in the image pair can compensate for each other to a certain degree, the intensity correction by Equation (1) has certain tolerance to field map error propagation and can obtain better images compared to traditional Jacobian modulation.

After obtaining the corrected images, we need to derive the cycle back-calculated images based on the forward model to get the cycle-consistency loss for unsupervised training of the network. In Figure 1C, we can see that the network will try to optimize the cycle-consistency loss between the cycle images ($I_{\text{cycle up}}(u)$ and $I_{\text{cycle down}}(u)$) and uncorrected images ($I_{\text{dis up}}$ and $I_{\text{dis down}}$):

$$\begin{aligned}\text{Lmse}_{\text{cycle}} = &\left[I_{\text{cycle up}}(u) - I_{\text{dis up}}\right]^2 \\ &+ \left[I_{\text{cycle down}}(u) - I_{\text{dis down}}\right]^2.\end{aligned} \tag{2}$$

2.4 | The loss function

To optimize the performance of susceptibility artifact correction, a hybrid loss function that considers several knowledge priors is proposed. The first term of the hybrid loss is the cycle-consistency losses between the cycle back-calculated blip-UP/DOWN images and original UP/-DOWN images. It also consists of the field map loss that enforces similarity between the two field maps, and the structural similarity between the T_2/T_1 weighted image and the corrected EPI images. The hybrid loss function can be defined as:

$$L_{\text{all}} = \text{Lmse}_{\text{cycle}} + \lambda_1 \text{Lmse}_{\text{map}} + \lambda 2 \text{Lmse}_{\text{cor}} + \lambda_3 \text{Lnmi}_{\text{cor}}, \tag{3}$$

Where $\text{Lmse}_{\text{cycle}}$ is the cycle consistency loss representing the difference between the original and the cycle back-calculated images. Lmse_{map} denotes the

loss of the difference between two field maps. [20] $\mathrm{Lmse}_{\mathrm{cor}}$ denotes the loss for the two corrected UP/DOWN images. And $\mathrm{Lnmi}_{\mathrm{cor}}$ means the normalized mutual information (NMI)[28,29] for the structural similarity between the susceptibility artifact-corrected image and the reference image, which is defined as:

$$\mathrm{Lnmi}_{\mathrm{cor}} = 1 - \mathrm{NMI}\left(I_{\mathrm{undis}}, I_{\mathrm{ref}}\right), \tag{4}$$

Where I_{undis} and I_{ref} are the corrected image and the reference image, respectively. The expression of normalized mutual information is as follows:

$$\mathrm{NMI}\left(I_{\mathrm{undis}}, I_{\mathrm{ref}}\right) = \frac{\sum_{i_{\mathrm{undis}}} p_{I_{\mathrm{undis}}}\left(i_{\mathrm{undis}}\right) \log p_{I_{\mathrm{undis}}}\left(i_{\mathrm{undis}}\right) + \sum_{i_{\mathrm{ref}}} p_{I_{\mathrm{ref}}}\left(i_{\mathrm{ref}}\right) \log p_{I_{\mathrm{ref}}}\left(i_{\mathrm{ref}}\right)}{\sum_{i_{\mathrm{undis}}, i_{\mathrm{ref}}} p_{I_{\mathrm{undis}}, I_{\mathrm{ref}}}\left(i_{\mathrm{undis}}, i_{\mathrm{ref}}\right) \log p_{\mathrm{XY}}\left(i_{\mathrm{undis}}, i_{\mathrm{ref}}\right)}, \tag{5}$$

Where $p_I(i)$ represents the probability distribution of I, which means the probability that variable I has value i, $p_{I_{\mathrm{undis}}, I_{\mathrm{ref}}}\left(i_{\mathrm{undis}}, i_{\mathrm{ref}}\right)$ is the joint probability distribution of I_{undis} and I_{ref}. NMI values range from 0 (no mutual information) to 1 (completely relevant). In addition, $\lambda_1, \lambda_2, \lambda_3$ are the weighting factors in formula (2) that are empirically set as $\lambda_1 = 1$, $\lambda_2 = 1$, and $\lambda_3 = 0.2$ after performing an array of optimization. Figure S1 in Appendix S1 depicts the convergence curves of DLRPG-net training under different loss components.

2.5 | Datasets

All animal experiments were conducted following the USA National Institutes of Health (Bethesda, Maryland, USA) animal care guidelines, and the procedures were approved by the Animal Ethics Committee of Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences (APM22022A). The preclinical MRI data were acquired on a Bruker Biospec 7.0 T/20 cm MRI scanner using a homemade modified EPI sequence with an inverse encoding gradient in the coronal view of the rat brain (source code is available at https://github.com/baoqingjia/EPI_correction). A 72-mm-diameter volume coil was used for Radio Frequency (RF) transmission and a quadrature surface coil for signal detection. The rats were anesthetized with isoflurane mixed with oxygen (4.0%–5.0% for induction, 0.5%–1.0% for maintenance). The respiratory rate (30 ~ 50) was continuously monitored during the scan under the anesthesia state scan. The body temperature was maintained with 37°C water circulation. The sequence parameters are as follows: Field of View (FOV) $= 28 \times 28\ \mathrm{mm}^2$, acquisition matrix size $= 80 \times 80$, TE = 30 ms, TR = 4000 ms, slice thick = 1 mm, number of slices for each rat ranged from 20 to 30, number of averages = 1. Thirteen diffusion-weighted scans were distributed equally over 1 shell defined with b-values of 1000s/mm^2, diffusion gradient time $\delta = 3.5$ ms, and gradient separation time $\Delta = 12$ ms. In addition, each rat also has corresponding T_2w images acquired with rapid acquisition with relaxation enhancement (RARE) pulse sequence: FOV $= 28 \times 28\ \mathrm{mm}^2$, acquisition matrix size $= 256 \times 256$, TE = 20 ms, TR = 2500 ms, slice thick = 1 mm, and reference images' region and position are the same as for the EPI sequence.

All clinical data were acquired using multi-shell multi-band EPI sequences with an inverse encoding gradient, downloaded from the Human Connectome Project (HCP) website (Washington University, University of Minnesota, and Oxford University).[30,31] All the images were acquired on the HCP standard 3 T Siemens MRI scanner with 32 channel head coil.[32] Parameters selected for DWI-EPI were: FOV $= 208 \times 180\ \mathrm{mm}^2$; acquisition matrix size $= 104 \times 90$; TE = 89.5 ms; TR = 5520 ms; nominal voxel size of 1.25 mm isotropic; and 270 diffusion-weighted scans distributed equally over 3 shells defined with b-values of 1000, 2000, and 3000 s/mm^2. Each scan was repeated along two reversed phase-encoding directions (Left to Right: L/R and Right to Left: R/L). For the T_1w images, FOV $= 224 \times 224\ \mathrm{mm}^2$, acquisition matrix size $= 320 \times 320$, TE = 2.14 ms, TR = 2400 ms, and nominal voxel size of 0.7 mm isotropic.

2.6 | Implementation

The DLRPG-net was implemented on Pytorch for the Python 3.6 environment on an NVIDIA (Santa Clara, California, USA) Geforce GTX 2080Ti with 11GB GPU memory and Intel (Santa Clara, California, USA) Core CPU i7-8700 3.7 GHz. The network was trained by the hybrid loss function according to section 2.4 and Adam optimizer[33] with $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$. Each dataset was trained for 300 epochs. The initial learning rate was set to 0.001, and it decreased by a factor of 0.95 after five epochs.

Two deep learning methods (Soan[17] and Benjamin[18]) and the well-known traditional method (TOPUP[13,14]) are used for the comparison with DLRPG-net. In terms of the model parameters, the Benjamin model has 2.87 million, DLRPG-net has 3.02 million, and Soan has 2.16 million parameters. All deep learning methods (Soan, Benjamin, and DLRPG-net) running on the PyTorch platform have a correction time of less than 1.8 seconds per subject with GPU. With CPU, the correction times are 2.8, 9.6, and 15.2 s, which are still less than the correction time of the traditional method TOPUP (598 s).

We also performed the Bloch equation–based simulations to quantitatively compare the susceptibility correction results. Simulations were performed based on our preclinical dataset acquired with the Fast Spin Echo (FSE) pulse sequence on a 7.0 T scanner. Moreover, the B_0 inhomogeneity maps were obtained from the field map protocol based on 3D-GRE with two different TEs. In the simulation, Bloch equations were used to consider the precession of spins only as an effect of RF pulses, gradients, and B_0 inhomogeneity effects, ignoring T_1 and T_2 relaxation. To obtain a more accurate description of the impact of B_0 inhomogeneities, each image voxel contained 16 spins, which were summed together after Bloch simulation to obtain the final result. The simulation was performed based on our home-built MatLab 2016a (MathWorks, Natick, MA) code, which can be downloaded from the github. The distorted ssEPI images (UP/DOWN) were simulated with different phase-encoding gradient polarities, and the reference images were simulated with zero B_0 inhomogeneity, in which UP/DOWN images are identical. The quality of the outputs of the network was evaluated by two quantitative metrics: peak Signal Noise Rate (PSNR) and structural similarity index (SSIM).[34] The details of the simulations are described in the Text S4 in Appendix S1.

For all three deep learning–based methods (Soan, Benjamin, and DLRPG-net), we train the network with both the simulated preclinical dataset and HCP dataset. For the preclinical study, we utilize 800 slices of simulated data to train the network and test the network with both simulated and real scan datasets acquired on Bruker 7 T. The 500 slices of simulated data are used to evaluate PSNR/SSIM quantitative metrics. Moreover, a total of 2275 slices (8 mice) of real scan datasets acquired on Bruker 7 T were tested. For susceptibility distortion correction of the clinical dataset, 500/200/1000 slices are used for training, validating, and testing. It is worth mentioning that we only take B_0 images (without diffusion weighting) for training, and both B_0 and diffusion-weighted images are used for testing. Moreover, we also investigate the distribution of the Mean Square Error (MSE) error between correct UP and Down images.

3 | RESULTS

3.1 | Validation for simulated EPI data acquired with reversed gradients

Figure 3 shows the corrected images for simulated EPI data acquired with the reversed gradients and the corresponding quantitative comparison. The columns from left to right show reference images, distorted UP/DOWN images, images corrected by TOPUP, Soan, Benjamin, and DLRPG-net methods. The last column shows the corresponding field maps used for the simulation. Here, we also show the correction results for different B_0 inhomogeneities, one from −150 to 150 Hz and the other from −200 to 200 Hz. As the simulated datasets are used in this test, the B_0 inhomogeneity ranges can be deliberately scaled to two different ranges. Figure 3 illustrates that DLRPG-net can obtain better-corrected images. This can be particularly well perceived by the zoomed regions (the second and fifth row) and by the error maps (the difference between the reference images and the corrected images). This can be appreciated even better from the zoomed regions and the absolute error map. And the last line summarizes the quantitative comparison between the results obtained by TOPUP, Soan, Benjamin, and DLRPG-net methods. Regarding quantitative metrics, PSNR/SSIM for DLRPG-net is 33.6/0.92, which is higher than 28.3/0.82 for TOPUP, 29.8/0.82 for Soan, and 30.1/0.85 for Benjamin methods.

Figure 4 and Figure S2 in Appendix S1 show the comparison of different deep learning–based susceptibility artifact correction methods with traditional TOPUP for simulated data set comprising 500 slices. Every point in Figure 4 represents SSIM for one slice based on TOPUP and deep learning–based method. The first row shows the scatterplot comparison. The points above the red line indicate the results of deep learning–based methods, which are better than TOPUP results. The second row shows the Bland–Altman analysis comparison. We can notice that the Bland–Altman analysis (average mean difference, average + 1.96 SD, average −1.96 SD) in SSIM of DLRPG-net (16.3, 38.8, −6.1) in % also shows the improvement compared to that of Soan (6.1, 26.8, −14.6) and Benjamin (9.2, 33.1, −14.7) methods. The statical distribution of correction results is shown in Figure S2 in Appendix S1.

3.2 | The correction results for clinical data

Figure 5 presents the comparison between correction results for B_0 images of DWI-3 T data with the edges extracted from the corresponding T_2 images using boundary-based registration.[35] The images from left to right correspond to distorted blip-UP/DOWN images, images corrected by TOPUP, Soan, Benjamin, and DLRPG-net methods. The second and fifth rows show the absolute error maps between cycle back-calculated blip-UP and original UP images. The third and sixth rows show the absolute error maps between corrected UP and DOWN images. Figure 5 illustrates that the DLRPG-net can obtain better-corrected images with respect to the other assessed methods. This advantage is especially

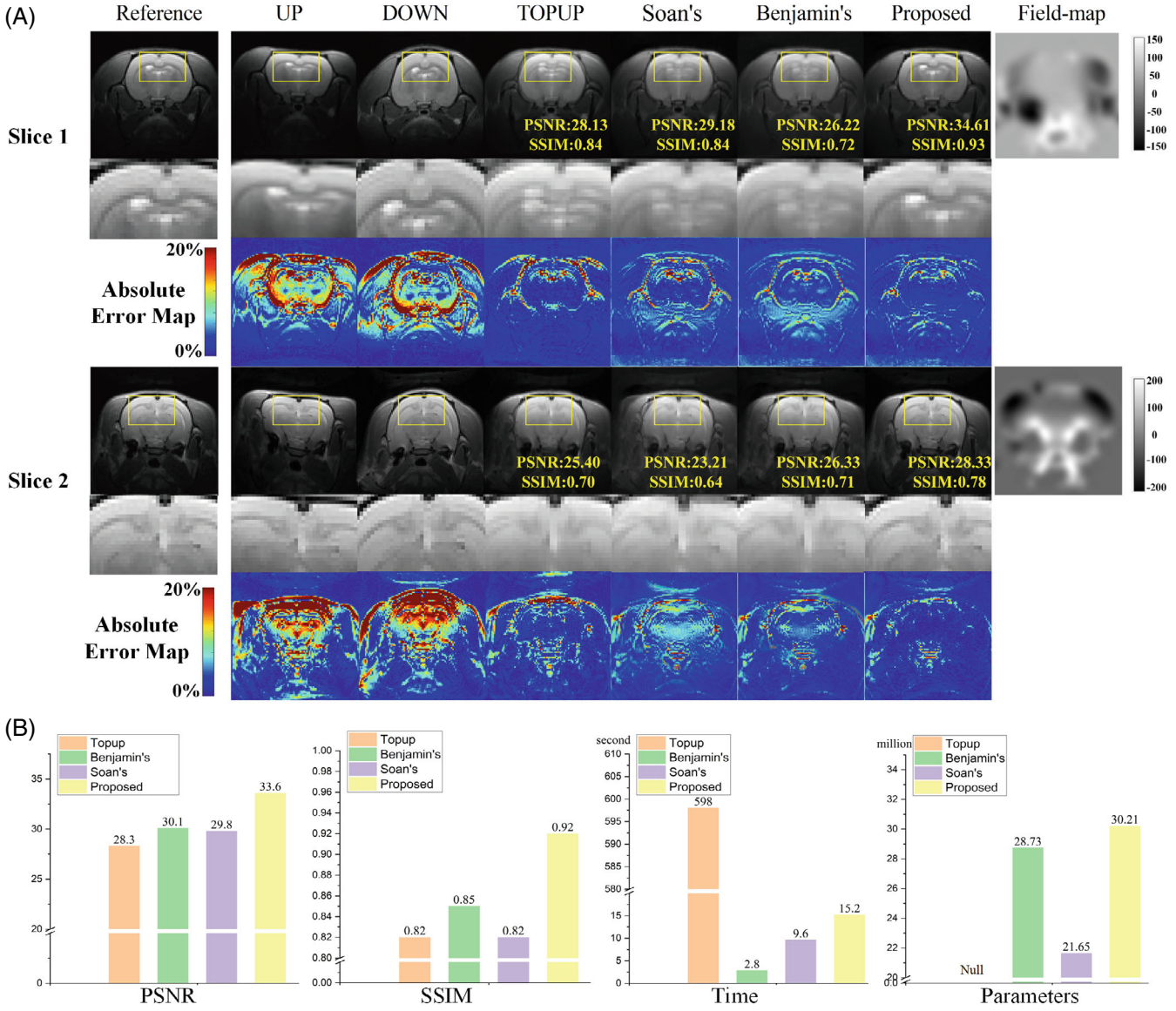

FIGURE 3 Comparison of different susceptibility artifact correction methods for simulated EPI-7 T mice data set and the quantitative results of PSNR, SSIM, correction time, and number of parameters. Columns from left to right correspond to reference images (no B_0 inhomogeneity), distorted blip-UP/DOWN images, and corrected images based on TOPUP, Soan, and Benjamin methods, as well as the herein proposed method (DLRPG-net) along the field map used for simulation.

pronounced at the edges and in the region of the cerebrospinal fluid (CSF). Images corrected with DLRPG-net show the smallest error maps in both cycle back-calculated error maps and error maps calculated between corrected UP and DOWN images.

Because the ground truth image cannot be collected in clinical experiments, the PSNR and SSIM indexes are no longer applicable, so we choose the MSE between the corrected UP and Down images as the evaluating indicator. Figure 6 and Figure S4 in Appendix S1 show the comparison of different deep learning–based susceptibility artifact correction methods with traditional TOPUP for HCP data set of 1000 slices. Every point shows MSE between the corrected UP and Down images for one slice based on TOPUP and deep learning–based methods. The first row shows the scatterplot comparison. The points below the red line indicate the results of deep learning–based methods are better than TOPUP results. The second row shows the Bland–Altman analysis comparison. We can notice that the Bland–Altman analysis (average mean difference, average +1.96 SD, average −1.96 SD) in MSE of DLRPG-net (−0.19, −0.35, −0.02) in % also shows the improvement compared to that of Soan (−0.13, −0.36, −0.1) and Benjamin (−0.15, −0.37, 0.07) methods. The statical distribution of correction results is shown in Figure S4 in Appendix S1.

We also evaluated the corresponding fractional anisotropy (FA) map and diffusion-encoded-color (DEC)

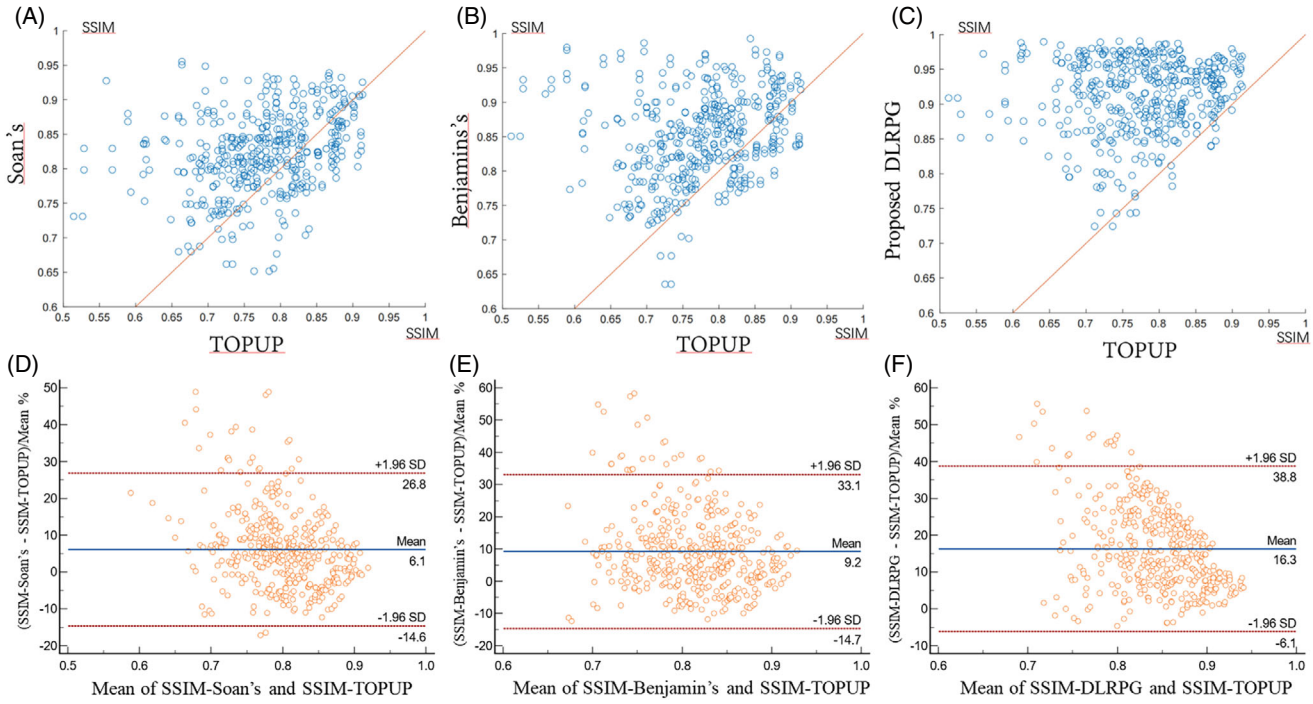

FIGURE 4 Comparison of different deep learning–based susceptibility artifact correction methods with traditional TOPUP for simulated data set comprising 500 slices. Every point represents SSIM for one slice based on TOPUP and deep learning–based methods. The first row shows the scatterplot comparison. The points above the red line indicate the results of deep learning–based methods, which are better than TOPUP results. The second row shows the Bland–Altman analysis comparison. (A, D) present the comparison of Soan and TOPUP methods. (B, E) present the comparison of Benjamin and TOPUP methods. (C, F) present the comparison of DLRPG-net and TOPUP methods. The Bland–Altman analysis (average mean difference, average +1.96 SD, average −1.96 SD) in SSIM of DLRPG-net (16.3, 38.8, −6.1) in % shows the improvement compared to that of Soan (6.1, 26.8, −14.6) and Benjamin (9.2, 33.1, −14.7) methods.

map based on the images obtained after susceptibility artifact correction. Figure S3 in Appendix S1 shows the comparison of different susceptibility artifact correction methods for two slices of the FA and DEC maps. One can notice that the deep learning method can obtain better B_0 images, FA, and DEC maps compared to other evaluated methods.

3.3 | The correction results for preclinical data

We also applied different susceptibility artifact correction methods to correct artifacts on the experimental preclinical DWI-7 T dataset. Figure 7A shows a comparison of the results after applying corrections using different methods, including TOPUP, Soan, Benjamin, and DLRPG-net, which shows that DLRPG-net can obtain better-corrected images compared to the rest of the assessed methods. Figure 7B presents the comparison between the results after applying corrections for FA maps and DEC maps arising from images after applying different correction methods. As can be appreciated from Figure 7, the deep learning method can obtain better *b*0-corrected images, diffusion-weighted images, FA map, and DEC map compared to other evaluated methods.

4 | DISCUSSION

This work has introduced a new method for distortion correction in EPI images that arise from magnetic field inhomogeneities. Quantitative metrics on all examined data sets demonstrate, along with qualitative image assessment, that our proposed method outperforms the widely employed TOPUP method and some of the recently introduced deep learning-based methods. At the same time, the method introduced herein is orders of magnitude faster than TOPUP method while using the same input, namely, so-called UP and DOWN images.

Particular effort has been dedicated to addressing one of the main challenges when deriving field maps: overcoming discontinuities introduced through the fitting process. In TOPUP, this has been addressed by fitting a field map to a set of discrete cosine functions. In line with this idea, we chose spline functions as our basis set. We can notice in Figure 2 that the coefficients of field maps in the different resolution subspaces correspond to different feature layers in the network. For example, the feature maps in DLRPG-net's bottom layer correspond to the coefficients in the restricted low-resolution subspace. In contrast, the feature maps of the upper layers correspond to the coefficients representing the higher-resolution subspace of the

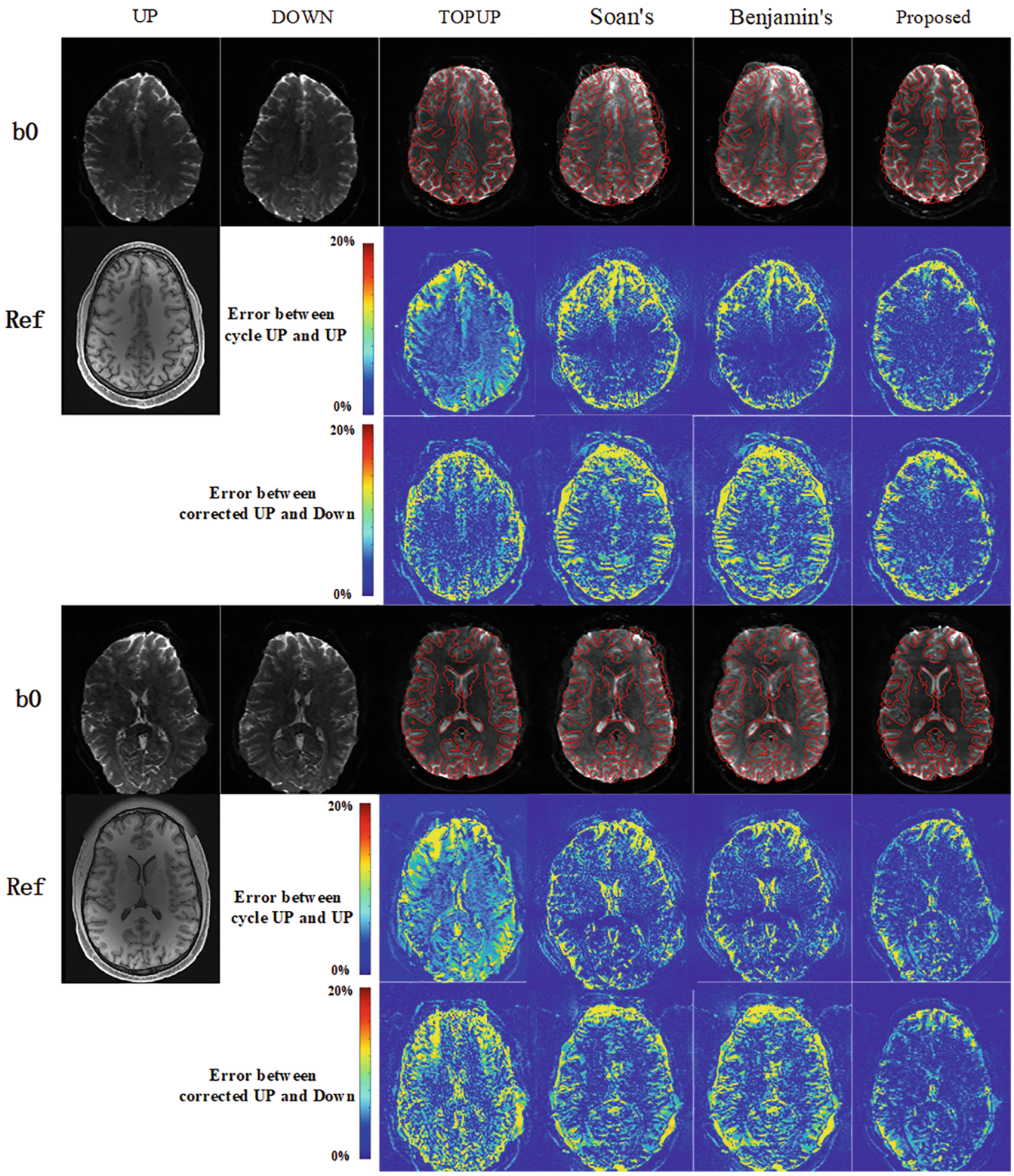

FIGURE 5 The susceptibility artifact correction results were compared for B_0 images of DWI-3 T data. The edges extracted from the corresponding T_1 reference images are overlaid over the corrected images. The second and fifth rows show the corresponding absolute error maps between cycle back blip-UP and original UP. The third and sixth rows show the corresponding absolute error maps between corrected UP and corrected DOWN

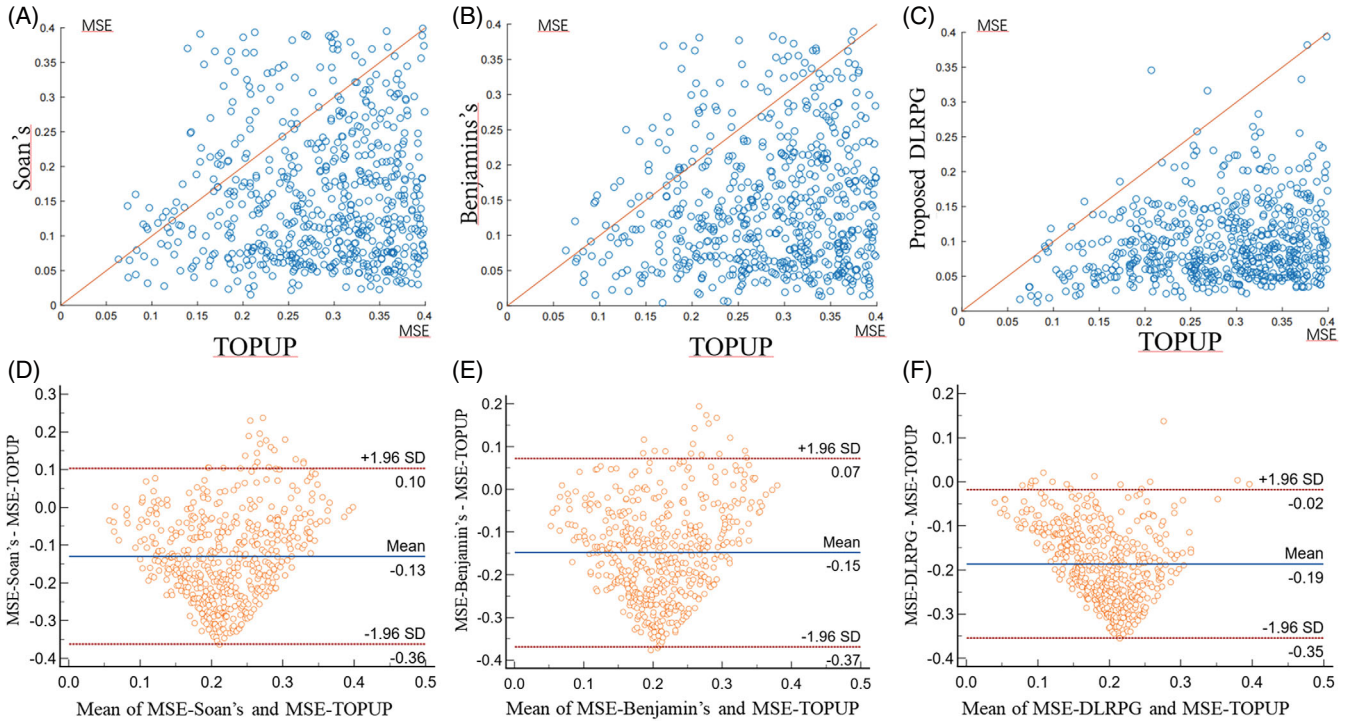

FIGURE 6 Comparison of different deep learning–based susceptibility artifact correction methods with traditional TOPUP for HCP data set of 1000 slices. Every point shows MSE between the corrected UP and Down images for one slice based on TOPUP and deep learning–based methods. The first row shows the scatterplot comparison. The points below the red line indicate the results of deep learning–based methods are better than TOPUP results. The second row shows the Bland–Altman analysis comparison. (A, D) present the comparison of Soan and TOPUP methods. (B, E) present the comparison of Benjamin and TOPUP methods. (C, F) present the comparison of the DLRPG-net and TOPUP methods. The Bland–Altman analysis (average mean difference, average +1.96 SD, average −1.96 SD) in MSE of DLRPG-net (−0.19, −0.35, −0.02) in % shows a big improvement compared to that of Soan (−0.13, −0.36, −0.1) and Benjamin (−0.15, −0.37, 0.07) methods. HCP, Human Connectome Project.

field map. Firstly, this helps to avoid discontinuities in the derived field maps; and secondly, this reduces the space in which the field map is expected to lie by constraining it to a subspace spanned by the basis functions. Whereas some of the methods try to enforce smoothness on the field map by penalizing total variation of the field map, using a set of smooth basis functions ensures this by definition. Impact of the usage of spline subspace on simulated EPI-7 T mice data set and HCP DWI-3 T data dataset is illustrated in Figure S5 in Appendix S1, where SSIM is significantly reduced when directly using extracted feature maps without projecting them to the spline basis set, that is, by ablating the back projection module. This is particularly well exemplified by preclinical images, where a spline basis set helps to reduce the dimensionality of the solution space and improves the convergence, particularly well perceivable in the vicinity of structures with sharp intensity changes, where model without spline basis does not converge to one solution.

After the deep learning network training is completed, it is essential to test the generalization capability of the network. In this work, we train the network with only $b=0$ images and successfully correct distortion in diffusion-weighted images. We also tried to train the network with diffusion-weighted data acquired with diffusion-weighting gradients applied in one orientation and corrected diffusion-weighted data acquired with diffusion-encoding gradients applied in another direction successfully. Moreover, we trained one network on preclinical data and tested it on clinical data and another network on clinical data and tested it on preclinical data. For the network trained only with clinical data, it can also provide satisfactory correction results for some preclinical datasets; however, some preclinical data (~20%) was not corrected to an acceptable level, especially for images displaying large B_0-associated inhomogeneity distortions. It is worth mentioning that the main advantage of this distortion correction method is that it is based on unsupervised training, without reference data (label), which is often challenging to obtain. We calculate the cost function using the physical forward problem of distortion correction during conventional training without a need for reference data. With this physical forward model, we try to correct the images directly from a single dataset ("no pretraining

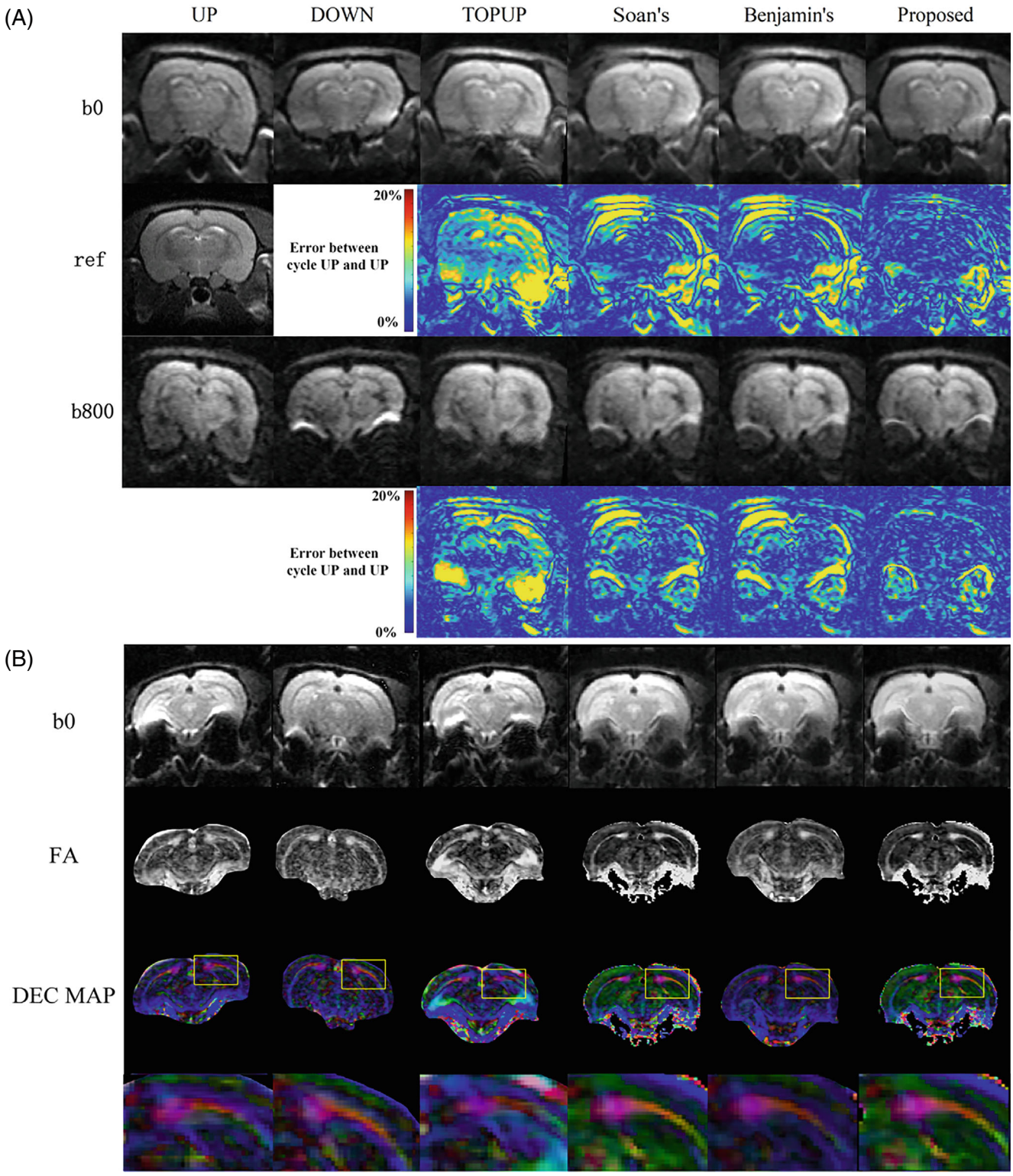

FIGURE 7 (A) Comparison of the correction results of preclinical rats data set obtained by the different methods, including TOPUP, Soan, Benjamin, and DLRPG-net. The second and fourth row shows the absolute error between the cycle UP and UP images and a reference T_2w image. (B) Comparison between the results after applying corrections for FA (the second row) maps and DEC (the third row) maps arising from images by the different methods, including TOPUP, Soan, Benjamin, and DLRPG-net. The fourth row shows the zoomed region of the DEC maps.

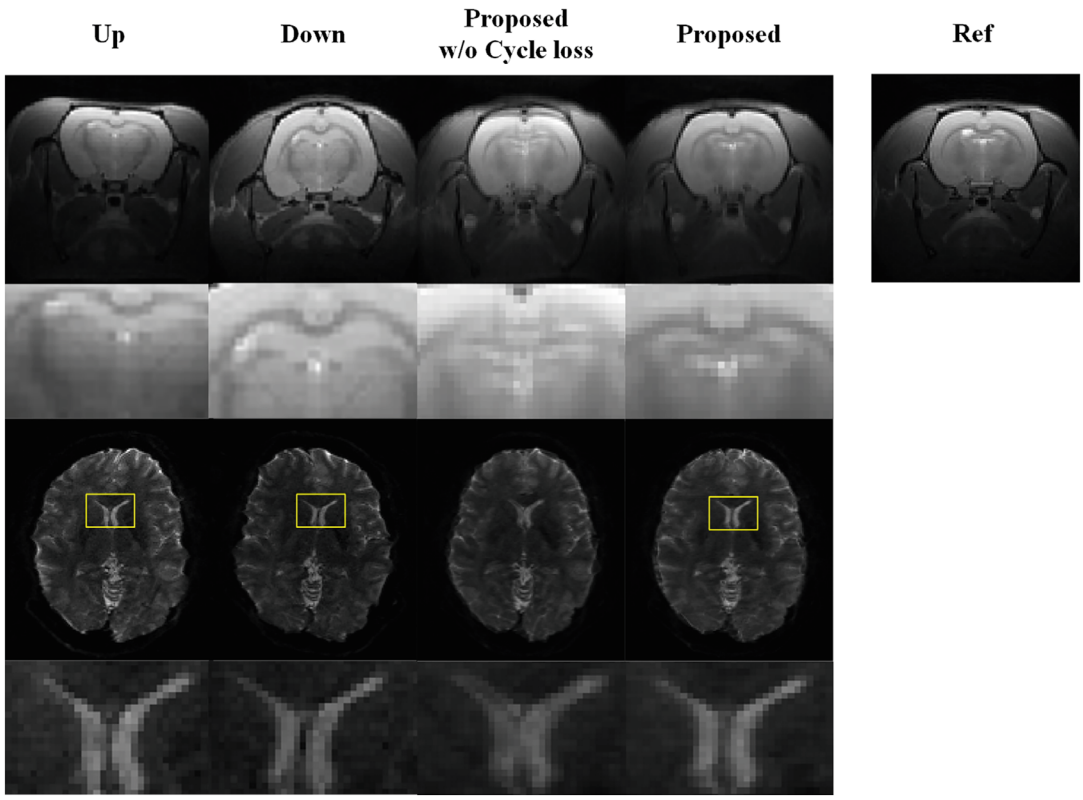

FIGURE 8 Comparison of our proposed method with and without the cycle consistency loss for our laboratory DWI-7 T and HCP DWI data. On the far right are the reference and the corresponding quantitative comparison (SSIM) for a slice of HCP DWI data

network"). Although the generalization will not constitute a challenge for this "no pretraining network," it will take much longer time to correct the images (~30 s for one image). Therefore, in this work two networks were trained for the preclinical and human data, respectively.

The ablation experiments were also performed to evaluate the advantage of the cycle- consistency loss and the NMI loss terms. Firstly, the improvements of the cycle-consistency loss term are shown in Figure 8 both for the simulated EPI data and real in vivo DWI-3 T data. In Figure S6 in Appendix S1, results from the ablation experiment are shown with and without NMI loss. As can be seen, the CSF region and the edge region were better corrected when using NMI loss. Moreover, we demonstrated that our proposed model can perform shift correction with high fidelity and the more challenging task of intensity correction. We assign this asset to the adoption of an optimized method to calculate intensity correction as described by Equation 1. This approach tries to alleviate the limited accuracy of the derived field maps. The benefits of Equation 1 are emphasized in Figure 9, comparing it to the classical Jacobian correction method.

We also explored the appropriate size of training data sets. To this end, we conducted an experiment varying the number of slices used for training from 200 to 1400 in steps of 100 slices. After reaching the slices number of 800, the PSNR/SSIM indices reach a plateau, and the increase in the training data set does not result in obvious performance gain, although it is accompanied by increased memory usage. Therefore, for the current study, we concluded that 800 is the most appropriate dataset size.

One advantage of the TOPUP method is the fast acquisition of input image pairs from which the field map will be derived. This allows for minimizing the impact of motion, which is much more of a concern for alternative distortion correction methods based on time-consuming field mapping. Nonetheless, to define cost function as physically realistic as possible, we accounted for potential impact of motion or field drift and derived two field maps allowing differences in field maps corresponding to "UP" and "DOWN" images, respectively, but applying a slight penalty on these differences to force maximal similarity between them when no or a negligible amount of motion is present. The benefits of this approach in the presence of motion based on simulations are illustrated in Figure S7 in Appendix S1.

The simulated and in vivo preclinical/clinical experiments show the advantages of the susceptibility artifact correction method introduced in this work. However, this method is designed for the single pairs of UP/Down images without taking advantage of the whole diffusion-weighted dataset. Eddy-current-induced distortions caused by the

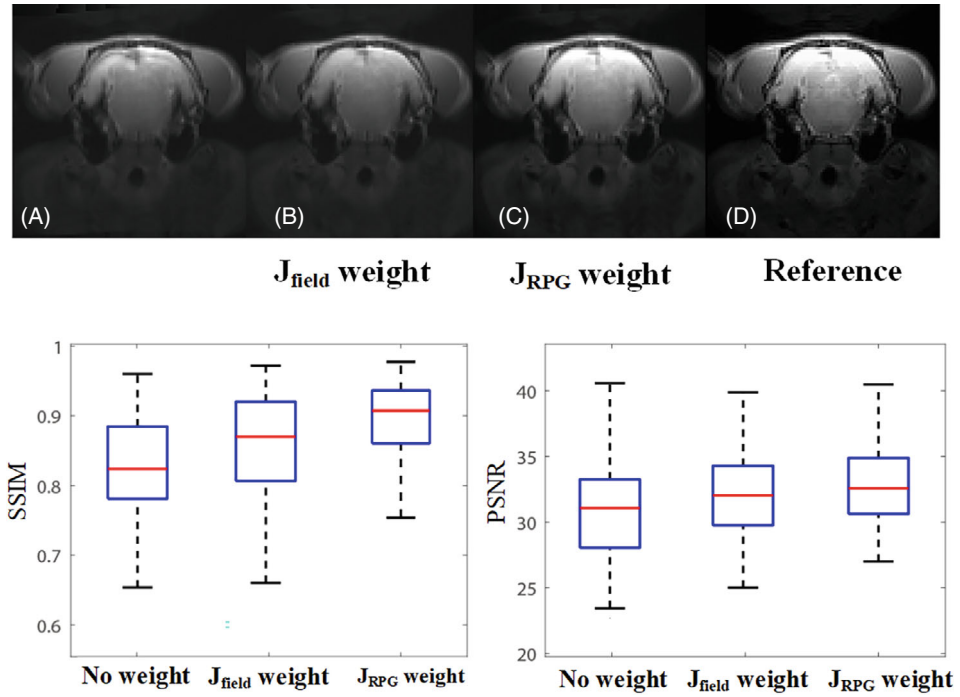

FIGURE 9 Comparison of different density compensation methods are shown in the first row: (A) corrected image without density compensation; in (B, C) are shown corrected images obtained with J_{field} and J_{RPG} density compensation methods, respectively. In addition, in (D) Fast Spin Echo (FSE) image is presented for reference. The second row presents a quantitative comparison (SSIM and PSNR metrics) between no weight, J_{field} weight, and J_{RPG} weight density compensation methods

rapid switching of diffusion gradients are ignored; these distortions may cause misalignment between different DWI volumes, thereby affecting the quality of the FA and DEC MAP. However, the existing eddy-current-distortion correction based on affine transformations is not enough. Our next direction is to further correct the eddy current distortion between different volumes of DWI based on the susceptibility artifact correction.

5 | CONCLUSION

This paper presents a cycle-consistent deep neural network that combines the deep neural network and the gradient reversal method for correcting susceptibility artifacts. The model can be applied with unsupervised training without explicit knowledge of the ground truth field map. Results using clinical and preclinical datasets demonstrate that our method outperforms state-of-the-art methods, and the speed is an order of magnitude faster than the traditional iterative TOPUP method.

ACKNOWLEDGMENTS

We gratefully acknowledge the financial support by National Major Scientific Research Equipment Development Project of China (81627901), the National key of R&D Program of China (grants 2018YFC0115000, 2016YFC1304702), National Natural Science Foundation of China (11575287, 11705274), and the Chinese Academy of Sciences (YZ201677).

FUNDING INFORMATION

National Major Scientific Research Equipment Development Project of China (81627901), the National key of R&D Program of China (Grant 2018YFC0115000, 2016YFC1304702), National Natural Science Foundation of China (11 575 287, 11 705 274), and the Chinese Academy of Sciences (YZ201677).

DATA AVAILABILITY STATEMENT

The source code is avaliable at https://github.com/baoqingjia/EPI_correction.

ORCID

Qingjia Bao https://orcid.org/0000-0003-0940-2535
Han Xie https://orcid.org/0000-0002-2672-3990
Xin Zhou https://orcid.org/0000-0002-9011-0662

REFERENCES

1. Schmitt F, Stehling MK, Turner R. *Echo-Planar Imaging: Theory, Technique and Application*. Springer; 2012.
2. Wu W, Miller KL. Image formation in diffusion MRI: a review of recent technical developments. *J Magn Reson Imaging*. 2017;46:646-662.
3. Bandettini PA. Twenty years of functional MRI: the science and the stories. *Neuroimage*. 2012;62:575-588.
4. Shiroishi MS, Castellazzi G, Boxerman JL, et al. Principles of T2*-weighted dynamic susceptibility contrast MRI technique in brain tumor imaging. *J Magn Reson Imaging*. 2015;41:296-313.
5. Le Bihan D, Poupon C, Amadon A, Lethimonnier F. Artifacts and pitfalls in diffusion MRI. *J Magn Reson Imaging*. 2006;24:478-488.

6. Wang S, Peterson DJ, Gatenby JC, Li W, Grabowski TJ, Madhyastha TM. Evaluation of field map and nonlinear registration methods for correction of susceptibility artifacts in diffusion MRI. *Front Neuroinform*. 2017;11:17.
7. Ruthotto L, Mohammadi S, Heck C, Modersitzki J, Weiskopf N. Hyperelastic susceptibility artifact correction of DTI in SPM. In: Meinzer HP, Deserno T, Handels H, Tolxdorff T, eds. *Bildverarbeitung für die Medizin 2013*. Informatik aktuell. Springer; 2013:344-349.
8. Xiong Y, Li G, Dai E, Wang Y, Zhang Z, Guo H. Distortion correction for high-resolution single-shot EPI DTI using a modified field-mapping method. *NMR Biomed*. 2019;32:124-132.
9. Morgan PS, Bowtell RW, McIntyre DJO, Worthington BS. Correction of spatial distortion in EPI due to inhomogeneous static magnetic fields using the reversed gradient method. *J Magn Reson Imaging*. 2004;19:499-507.
10. Koch KM, Rothman DL, Graaf RAD. Optimization of static magnetic field homogeneity in the human and animal brain in vivo. *Prog Nucl Magn Reson Spectrosc*. 2009;54:69-96.
11. Ong F, Cheng JY, Lustig M. General phase regularized reconstruction using phase cycling. *Magn Reson Med*. 2018;80:112-125.
12. Chang H, Fitzpatrick JM. A technique for accurate magnetic resonance imaging in the presence of field inhomogeneities. *IEEE Trans Med Imaging*. 1992;11:319-329.
13. Andersson J, Skare S, Ashburner J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *Neuroimage*. 2003;20:870-888.
14. Smith SM, Jenkinson M, Woolrich MW, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*. 2004;23:208-219.
15. Liao P, Zhang J, Zeng K, et al. Referenceless distortion correction of gradient-echo echo-planar imaging under inhomogeneous magnetic fields based on a deep convolutional neural network. *Comput Biol Med*. 2018;100:230-238.
16. Hu Z, Wang Y, Zhang Z, Zhang J, Guo H. Distortion correction of single-shot EPI enabled by deep-learning. *Neuroimage*. 2020;221:117-170.
17. Duong STM, Phung SL, Bouzerdoum A, Schira MM. An unsupervised deep learning technique for susceptibility artifact correction in reversed phase-encoding EPI images. *Magn Reson Imaging*. 2020;71:1-10.
18. Zahneisen B, Baeumler K, Zaharchuk G, Fleischmann D, Zeineh M. Deep flow-net for EPI distortion estimation. *Neuroimage*. 2020;217:116886.
19. Dosovitskiy A, Fischer P, Ilg E, et al. Flownet: learning optical flow with convolutional networks. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2015.
20. Weida X, Shi C, Qingjia B, et al. Unsupervised cycle-consistent network for removing susceptibility artifacts in single-shot EPI. *In Proceedings of The 13th Asian Conference on Machine Learning, PMLR*. 2021;157:1723-1738.
21. Sharma SD, Hu HH, Nayak KS. Accelerated water–fat imaging using restricted subspace field map estimation and compressed sensing. *Magn Reson Med*. 2012;67:650-659.
22. Tsao J, Jiang Y. Hierarchical IDEAL: robust water–fat separation at high field by multiresolution field map estimation. In Proceedings of the 16th Annual Meeting of ISMRM, Toronto, Ontario, Canada 2008:653.
23. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, Frangi A, eds.Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015. Lecture Notes in Computer Science. Cham, Switzerlanf: Springer; 2015:234-241.
24. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV. 2016. pp. 770–778.
25. Studholme C, Constable RT. Accurate alignment of functional EPI data to anatomical MRI using a physics-based distortion model. *IEEE Trans Med Imaging*. 2000;19:1115-1127.
26. Holland D, Kuperman JM, Dale AM. Efficient correction of inhomogeneous static magnetic field-induced distortion in Echo planar imaging. *Neuroimage*. 2010;50:175-183.
27. Liu S, Xiong Y, Dai E, Zhang J, Guo H. Improving distortion correction for isotropic high-resolution 3D diffusion MRI by optimizing Jacobian modulation. *Magn Reson Med*. 2021;86:2780-2794.
28. Estévez PA, Tesmer M, Perez CA, Zurada JM. Normalized mutual information feature selection. *IEEE Trans Neural Netw*. 2009;20:189-201.
29. Studholme C, Hill DL, Hawkes DJ. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognit*. 1999;32:71-86.
30. Van Essen DC, Smith SM, Barch DM, et al. The WU-Minn human connectome project: an overview. *Neuroimage*. 2013;80:62-79.
31. Glasser MF, Sotiropoulos SN, Wilson JA, et al. The minimal preprocessing pipelines for the human connectome project. *Neuroimage*. 2013;80:105-124.
32. Van Essen DC, Ugurbil K, Auerbach E, et al. The human connectome project: a data acquisition perspective. *Neuroimage*. 2012;62:2222-2231.
33. Kingma DP, Ba J. Adam: a method for stochastic optimization. 2014; arXiv preprint: arXiv:1412.6980.
34. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process*. 2004;13:600-612.
35. Greve DN, Fischl B. Accurate and robust brain image alignment using boundary-based registration. *Neuroimage*. 2009;48:63-72.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

Appendix S1. Supporting information.

How to cite this article: Bao Q, Xie W, Otikovs M, et al. Unsupervised cycle-consistent network using restricted subspace field map for removing susceptibility artifacts in EPI. *Magn Reson Med*. 2023;90:458-472. doi: 10.1002/mrm.29653