Unsupervised cycle-consistent network for removing susceptibility artifacts in single-shot EPI

Weida Xie   876863137@QQ.COM
Wuhan University of Technology

Shi Chen   CHENSHI0205@163.COM
Wuhan University of Technology

Qingjia Bao   BAOQINGJIAHAOBA@GMAIL.COM
Chinese Academy of Sciences

Kewen Liu   LIUKEWEN@126.COM
Wuhan University of Technology

Zhao Li   1051202035@QQ.COM
State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Wuhan Institute of Physics and Math, Innovation Academy for Precision Measurement Science and Technology.

Chongxin Bai   907372720@WHUT.EDU.CN
Wuhan University of Technology

Otikovs Martins   MARTINS.OTIKOVS@WEIZMANN.AC.IL
Weizmann Institute of Science, Rehovot, 76001

Piqiang Li   563040043@QQ.COM
Wuhan University of Technology

Jie Wang   JIE.WANG@WIPM.AC.CN
Chinese Academy of Sciences

Chaoyang Liu   CHYLIU@WIPM.AC.CN
State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Wuhan Institute of Physics and Math, Innovation Academy for Precision Measurement Science and Technology.

Editors: Vineeth N Balasubramanian and Ivor Tsang

Abstract

Single-shot EPI(ssEPI) is one of the most important ultrafast MRI sequences commonly used for diffusion-weighted MRI and functional MRI. However, ssEPI suffers from susceptibility artifacts, especially in the high field or at the tissue boundaries. The widely used blip/down approaches, such as TOPUP, estimate the underlying distortion field from a pair of images with reversed-phase encoding direction. Typically, the iterative methods are used to find a solution to the ill-posed problem of finding the displacement map that maps up/down acquisitions onto each other. Then the geometric and intensity corrections are applied to obtain the undistorted images based on the estimated displacement map. This paper presents a new unsupervised cycle-consistent deep neural network that takes advantage of both the deep neural network and the gradient reversal method. The proposed method consists of three main components: (1) the Resnet50-Unet to map the pair of images with inverted phase encoding to the displacement maps; (2) the geometric and intensity correction module to obtain the undistorted images; (3) the forward model is ap-

Xie Chen Bao Liu Li Bai Martins Li Wang Liu

plied to get the cycled blip up/down images, and the cycle-consistent loss is optimized. In addition, the CNN network will generate two field maps to overcome motion or field drift during the scan. This new network is trained unsupervised on the clinical datasets downloaded from the Human Connection Project website. And we test this method on both preclinical and clinical datasets. The preclinical dataset is collected from 20 mice based on the modified EPI pulse sequence in 7T scanner. Both simulated and experimental results demonstrate that our method outperforms state-of-the-art methods. In conclusion, we proposed an unsupervised cycle-consistent deep neural network for removing susceptibility artifacts. The results on both preclinical and clinical datasets show this new method's acceleration and generalization capabilities.

Keywords: Deep Learning $\bullet MRI \bullet EPI \bullet TOPUP \bullet DTI \bullet Susceptibility Artifacts$.

1. Introduction

Magnetic resonance imaging (MRI) is a widely used imaging method for clinical and preclinical applications Hedouin et al. (2017). Specifically, single-shot Echo Planar Imaging (ssEPI) is one of the most efficient MRI acquisition schemes which can provide relatively high-definition images in 100 ms or less Hu et al. (2019). Furthermore, the quality of ultrafast acquisition makes ssEPI immune to the motion artifacts and desirable for Diffusion Tensor Imaging (DTI) Irfanoglu et al. (2015), functional MRI (fMRI) Yun and Shah (2020), and Dynamic Susceptibility Contrast MRI (DSC-MRI) Holland et al. (2010). However, due to the B0 field inhomogeneity and the low bandwidth along the phase-encoding direction, ssEPI suffers from severe spatial and intensity distortion (susceptibility artifacts), especially at tissue boundaries with different susceptibilities.

There have been many studies for removing the susceptibility artifacts in ssEPI. Among various proposed ssEPI susceptibility correction methods, the two most representative and commonly used methods are the field-mapping methods and the gradient reversal methods. The conventional field-mapping method usually collects two gradient-echo (GRE) images with different TE values to calculate B0 field inhomogeneity Xiong et al. (2019). Then, it calculates each pixel's shift (the displacement map) from the field map and corrects the distorted ssEPI images by coordinating calculation and linear interpolation in the image domain. Although this field-mapping method is easy to implement, its performance is limited by the field map's quality. Inaccurate field maps will lead to the residual artifacts after the distortion correction and reduce this method's efficacy. Note that a full EPI image of the human brain takes only a few seconds while acquiring a GRE-based field-map requires several minutes Liebig et al. (2020). Furthermore, the phase unwrapping procedure in the field-map calculation is vulnerable to various errors, especially near tissue boundaries or regions with high field inhomogeneity. More importantly, it cannot deal with the intensity variation problem Mani et al. (2017).

Instead of acquiring the real field map in the field mapping method, the gradient reversal methods rely on two reversed phase-encoding (PE) images to estimate the displacement map, in which the two images are acquired using identical sequences but with opposite PE directions Ong et al. (2018). The reversed gradient method exploits the fact that when the phase encoding gradients are reversed in polarity, if the voxels of the positive phase image's region are compressed, the corresponding negative phase image's region will be

stretched and vice versa. Chang et al Chang and Fitzpatrick (1992), proposed the gradient reversal method, which estimates the displacement map in each line along the PE direction independently by performing every phase encoding line integral. The main limitation of this method is the use of 1-D unwarping for every phase encoding line without considering the displacement map's smoothness, resulting in streaking or discontinuities in the images. An alternative implementation of the reversed gradient method by Andersson et al Andersson et al. (2003), involves fitting for a smoothly-varying 3-D displacement field using discrete cosine basis functions. This method is also widely used in FSL(FMRIB Software Library) as TOPUP Smith et al. (2004). These conventional approaches' main problem is time-consuming optimization of the objective function, especially for input images with large sizes or severe displacements. And the accuracy of the optimized displacement maps is also limited as the resolution of the field map would be limited by the highest frequency component of the discrete cosine basis functions. Moreover, the optimized objective functions often lead to a poorly conditioned, nonconvex optimization problem and a deficient local minimum.

In recent years, with the development of deep learning in medical image processing, several researchers began to use deep learning to remove susceptibility artifacts in ultrafast MRI correction. Deep learning techniques provide a potential avenue for dramatically reducing computational time and improving the objective function's convergence. These deep learning methods can be divided into two groups: the supervised network and the unsupervised network. The supervised methods need the ground truth data, such as the simulated datasets, to train the model. Liao Liao et al. (2018) proposed to use the convolutional neural network (CNN) for gradient-echo EPI distortion correction. They first used undistorted images to simulate distorted images by the SPROM software. Then the simulated datasets were used to train the network. Zhangxuan Hu et al. (2020) proposed a 2D-Unet based network for removing the distortion in ssEPI, in which they use PSF-EPI images as targets in the training stage. However, the truth displacement maps are tedious to acquire or may not be guaranteed to represent a "true" ground truth due to the phase unwrapping or regularization procedure's errors. Moreover, the displacement maps based on simulations are always different from the experimentally measured maps, especially in the temporal lobes or surroundings of the sinuses.

On the other hand, unsupervised networks do not need the ground truth (the "true" displacement map or undistorted images) to train the network. Thus, these methods are more suitable for MRI studies. Recently, several groups try to combine the deep learning network and the reversed gradient phase encoding method to obtain the displacement map and remove the susceptibility artifacts in ssEPI, which can be much faster than the traditional methods. Soan Duong et al. (2020) proposed an end-to-end deep learning network (S-Net) to correct the reversed-PE EPI image pairs' susceptibility artifacts. Benjamin Zahneisen et al. (2020) explored using a deep convolutional network to estimate the displacement map from a pair of input images. They trained the U-net-based network by minimizing a similarity metric between the two corrected images.

This paper presents a new unsupervised cycle-consistent deep neural network for removing susceptibility artifacts in ssEPI with two reversed-phase encodings, in which the deep neural network is used to generate the displacement map, and the forward model is applied to get the cycled blip up/down images. Based on the cycled blip up/down images, the

network can be trained only with the cycle-consistent loss, without any explicit knowledge of the ground displacement map. The main contribution of this work includes:

1. A cycle architecture is introduced. Compared to the unsupervised networks by minimizing the similarity between the two corrected images (corrected blip-up or blip-down), our cycle architecture can make the network trained with only one final corrected image. As we all know, even with the displacement map, the inversion problem of obtaining the undistorted blip-up or blip-down is rank deficient, and no unique inverse exists. However, the matrix-inverse problem for obtaining the final undistorted is much easier Andersson et al. (2003).

2. Two field maps to overcome motion or field drift during the scan. One of the main assumptions of distortion corrections is that the B0 field (displacement map) is constant between the two acquisitions. Hence the distortions are exactly the opposite of each other, disregarding possible inconsistencies introduced, for example, by subject motion or magnetic field drift. Here, instead of only one displacement map, we will ask the network to generate two codependent displacement maps that are nearly identical but allow flexibility to compensate for any changes such as motion or field drift.

3. The density compensation method without calculating the Jacobian matrix is applied Xiong et al. (2019) Liu et al. (2021).

4. Both clinical and preclinical applications are tested to demonstrate the generalization capabilities of this method. The network can be trained unsupervised with only the clinical datasets, and it can work well not only on the clinical datasets, but also on the preclinical dataset.

2. Method

2.1. Problem Formulation

In general, the susceptibility artifacts in ssEPI can be assumed to be negligible along the frequency-encoding and slice-selection direction due to much higher acquisition or excitation bandwidth, and the acquired signal for two-dimensional ssEPI can be given Lee et al. (2016) Studholme et al. (2000) as:

$$S\left(k_x, k_y\right) \propto \int_x \int_y \rho(x,y) e^{j(k_x x + k_y(y + \frac{\Delta B_0(x,y)t_{esp}}{G_y T_p}))} dx dy \tag{1}$$

Where G_y and T_p enote gradient amplitude and duration, respectively. $\Delta B_0(x, y)$ is the local field inhomogeneity. $\rho(x, y)$ denotes a map of the density of excited protons, the t_{esp} denotes echo spacing. k_x and k_y denote the frequency-encoding and slice-selection direction respectively. From formula 1, we can notice that the signal from the real space location (x, y) will be shifted to $(x, y + \delta y)$, where $\delta y = \frac{\Delta B_0(x,y)t_{esp}}{G_y T_p}$. And the shift in units of pixels(named displacement map) can be written as $u(x, y) = \frac{\delta y}{\Delta y} = \Delta B_0(x, y)t_{esp}N_y$, where

$\Delta y = L_y/N_y$ indicates the resolution of one pixel. L_y, N_y denotes the field of view and phase encoding number, respectively.

The goal of the susceptibility artifact correction is to recover any pixel's real intensity (x, y) from the corresponding position $(x, y + \delta y)$, and it needs both intensity and geometric correction. As the image intensity will be proportional to the spin density, if distortion leads to local compression, the signal will be higher by the accumulation of compression, to the opposite, the signal will be lower for the local stretch. These distortions will be scaled by the Jacobian J of the transformation $y \to y' = y + u(x, y)$:

$$J(x, y) = 1 + \frac{\partial u(x, y)}{\partial y} = 1 + T_{pe} \frac{\partial \Delta B(x, y)}{\partial y} \tag{2}$$

Thus, the susceptibility corrected images can be obtained by $\tilde{I}(x, y) = J(x, y)I(x, y + u)$.

The main idea of the gradient reversal methods is shown in Figure 1(b). If the phase-encoding gradient is reversed, i.e., k-space is traversed in the opposite direction along the phase-encoding axis, the signal emanating from location (x, y) will be shifted to $(x, y - u)$ and $(x, y + u)$. Similarly, we can obtain the correction for the second image $\tilde{I}(x, y) = J_{\text{Up}}(x, y)I_{\text{Up}}(x, y + u) = J_{\text{Down}}(x, y)I_{Down}(x, y - u)$. Thus, the main problem of the susceptibility corrected is to obtain the displacement map $u(x, y)$. As we mentioned before, there are two kinds of methods to obtain $u(x, y)$, the first one is field-map, and the other one is to optimize the following problem:

$$\arg\min_u \sum \left[J_{\text{Up}}(x, y)I_{\text{Up}}(x, y + u) - J_{\text{Down}}(x, y)I_{Down}(x, y - u) \right]^2 \tag{3}$$

However, this optimization function is an ill-posed problem, and it is very time-consuming as the iterative method is always needed. This work proposes a new unsupervised cycle-consistent deep learning model that takes advantage of both the deep neural network and the gradient reversal method, in which the deep neural network is trying to obtain the displacement map $u(x, y)$ instead of the time-consumed iterative method. And based on the displacement map, we can obtain the final corrected image.

The network is trained with the cycle-consistent loss, without any explicit knowledge of the ground displacement map.

$$\arg\min_u \sum \left\{ [I_{CircleUp}(u) - I_{Up}]^2 + [I_{CircleDown}(u) - I_{Down}]^2 \right\} \tag{4}$$

As shown in Fig .1(c), the main idea of this method is to optimize the loss between the $I_{CircleUp}(u)$ which is obtained by the forward model with the displacement map and I_{Up} and I_{Down}. The whole procedure of caculating $I_{CircleUp}(u)$ contains two steps:

1. The corrected image based on the current displacement map $u(x, y)$ Xiong et al. (2019) Liu et al. (2021).

$$I(\boldsymbol{r}) = \frac{2I_{Down}(x, y - u) \cdot I_{Up}(x, y + u)}{I_{Down}(x, y - u) + I_{Up}(x, y + u)} \tag{5}$$

2. After we obtain the corrected image, the forward model is applied to obtain the $I_{CircleUp}(u)$ and $I_{CircleDown}(u)$, according to Equation (1) Lee et al. (2016) Studholme et al. (2000).

2.2. The architecture of the cycle-consistent network

the network is increased by the three residual layers, which is proposed to learn better transformation functions and representations through a deeper perception. And the decoder section is the right part of a four layers U-Net to get two displacement maps. We set the kernel size and stride equal to 3 and 2, respectively for the decoder section. As the field map has both positive and negative values, we do not use the activation module Relu, which output the positive values only. Moreover, our traditional correction module is subdivided into intensity distortion and geometric distortion Jezzard and Balaban (1995). In order to reduce computational complexity, our intensity distortion module use only a convolution layer to map the density map from the displacement map, which can replace the Jacobian matrix in formula 4. For the geometric distortion, we use the density images and two displacement map to get the undistortion image by traditional correction algorithms. In addition, Figure 1(b) has clarified the simplified diagram of the reversed phase encoding image acquisition. The workflow and the architecture of the proposed model are shown in Fig. 1(c). It consists of three main components: (1) the CNN module to map the pair of images($I_{U\mathrm{p}}$) and $I_{D\mathrm{own}}$) with inverted phase encoding to the displacement maps; (2) the geometric and intensity correction module to obtain the undistorted images($I_{U\mathrm{p}}$) and $I_{D\mathrm{own}}$); (3) the forward model is applied to get the cycled blip up/down images($I_{CircleUp}$ and $I_{CircleD\mathrm{own}}$). Finally, the cycle-consistent loss between the cycle images and original images is optimized to train the CNN.

The CNN module is based on the modified Res50-Uet. The input of this CNN is the two uncorrected ssEPI images(I_{Up}) and $I_{D\mathrm{own}}$), and the output is the corresponding displacement maps(I_{Dm}). In order to overcome the motion or the field drift during the two scans, we ask CNN to generate two displacement maps that are nearly identical but allow flexibility. The neural network module contains five encoder-decoder blocks. Each encoder block adopts the residual block structure, which contains two 2D convolutional layers, two ReLU activation functions, and a skip connection. In contrast, the decoder block contains an upsampling layer, a 2D convolutional layer. As the field map contains both positive and negative values, we removed the ReLU activation function of the decoder block to ensure that the output can be positive and negative. Moreover, concatenations were applied between the same scale feature maps from the encoder-decoder blocks, which allow the network to propagate context information to higher resolution layers.

Based on the CNN's displacement map(I_{Dm}), the traditional correction module will do both geometric and intensity correction to get the undistorted image Jezzard and Balaban (1995).Firstly, the geometric correction is applied by shifting the values that map a distorted pixel back to the undistorted grid. As the displacement map contains fractional pixel-shift values, mapping the image intensities to the unwrapped image grid has to be done by interpolation. Secondly, according to the Equation $\tilde{I}(x,y) = J(x,y)I(x,y+u)$, the intensity correction should be applied by multiplying the Jacobian matrix. The traditional way to calculate the Jacobian matrix is based on Numerical discrete differentiation as the field image is a digital image composed of discrete pixels Duong et al. (2020) Zahneisen et al. (2020).To overcome discrete numerical differentiation, the method that can derive the intensity of the undistorted image directly without the Jacobian calculation by combining

the pair of corresponding pixels from the two EPI images with opposite PE directions is used Xiong et al. (2019)., shown in Equation (3).

The third component is to get the cycled blip up/down images by the forward model Studholme et al. (2000). Finally, the cycle-consistent losses between the cycle up/down images and original up/down images are optimized to train the CNN.

2.3. Loss function

To optimize the performance of removing susceptibility artifacts, we propose the hybrid loss function that not only consider the cycle-consistent losses between the cycle up/down images and original up/down images, but also consider the information of smoothness of the field map. The hybrid loss function can be denoted as:

$$L_{All} = Lmse_{Circle} + \lambda_1 Lmse_{Cor} + \lambda2 Lmse_{Map} + \lambda_3 Ltv_{Map} \tag{6}$$

Where $Lmse$ is the mean square error loss, Ltv is the anisotropic total variation norm loss. $\lambda_1, \lambda_2, \lambda_3$ are weighting factors, we manually tried a number of configurations for the three weights in our experiment, and the weights of the losses are empirically set as $\lambda_1 = 1$, $\lambda_2 = 1$ and $\lambda_3 = 1$. This process can be automated through grid search or learn the weights during CNN training. Here, we did not explore this extension because we observed relatively small changes among the configurations we tried. Therefore, we used the above empirical values in this study.

The first term $Lmse_{Circle}$ is the cycle-consistent loss between the circle distortion images($I_{CircleUp}$ and $I_{CircleDown}$).and the distortion inputs (I_{Up}) and I_{Down}):

$$Lmse_{Circle} = \frac{1}{N}(\sum (I_{CircleUp} - I_{Up})^2 + \sum (I_{CircleDown} - I_{Down})^2) \tag{7}$$

The second term is the forword-consistent loss between Corrected Up($\tilde{I}_{Up}$) and Corrected Down($\tilde{I}_{Down}$) images:

$$Lmse_{Cor} = \frac{1}{N}\sum (\tilde{I}_{Up} - \tilde{I}_{Down})^2 \tag{8}$$

the third term is to constrain the two displacement maps(I_{Dm_1} and I_{Dm_2} represent the two different displacement maps for I_{Up}) and I_{Down} respectively) to be similar as the motion or field drift always introduces a smooth phase:

$$Lmse_{Map} = \frac{1}{N}\sum (I_{Dm_1} - I_{Dm_2})^2 \tag{9}$$

Meanwhile, based on the prior information that the displacement map must be smooth, the four term is the Total Variation loss as the smoothing constraint on the displacement maps:

$$Ltv_{Map} = \sum_{i=1,2} \sum_{x,y} |I_{Dm_i}(x + 1, y) - I_{Dm_i}(x, y)| + |I_{Dm_i}(x, y + 1) - I_{Dm_i}(x, y)| \tag{10}$$
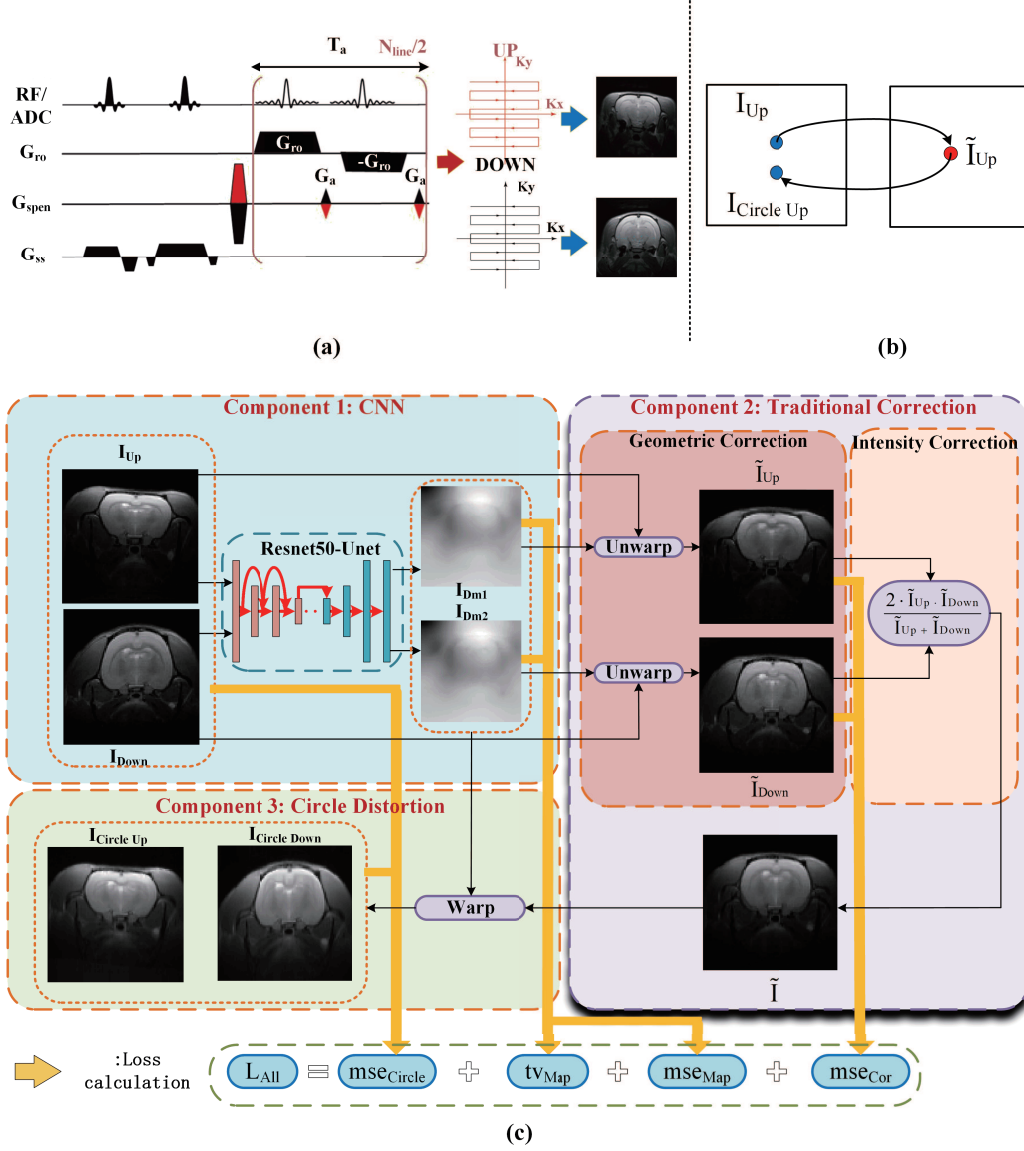
Figure 1: (a) The pulse sequence and k-space acquisition trajectory, (b) our model utilize the cycle-consistency: $I_{Up} \rightarrow \tilde{I}_{Up} \rightarrow I_{CircleUp} \approx I_{Up}$, and (c) The illustration of our proposed learning framework for the EPI correction.

Table 1: The summary of the datasets used in the experiments.

Datasets	No.sub	Image size	Resolution	PE direcions
Preclinical DWI-7T	30	$96 \times 96 \times 20 \times 16$	$0.29 \times 0.29 \times 0.29 mm3$	PA and AP
clinical DWI-3T	37	$144 \times 168 \times 111 \times 97$	$1.25 \times 1.25 \times 1.25 mm3$	LR and RL
Sim.EPI-7T	41	$96 \times 96 \times 18$	$0.29 \times 0.29 \times 0.29 mm3$	PA and AP

3. Experiments and result

3.1. Datasets and training

The clinical datasets DWI-3T are downloaded from the Human Connection project website. It includes 180 healthy subjects, and we randomly selected 37 subjects for our model. The total of Nsub = 37 subjects was split into Ntrain = 30 training case(111×30 slices), Ntest = 5 test case($111 \times 5 \times 97$ slices, 97 means the diffusion directions) and Nval = 2 validation cases(111×2 slices). The preclinical MRI experiments were performed on the mice brain's coronal position using a Bruker Biospec 7.0 T/20cm MRI scanner in our laboratory. A total of Nsub = 31 mice was split into Ntrain = 25 training case(25×20 slices), Ntest = 3 test case($3 \times 20 \times 16$ slices, 16 means the diffusion directions) and Nval = 3 validation cases(3×20 slices). We only use the b0 images to train our model, while all the diffusion-weighted images are used to test for both the clinical and preclinical DWI datasets. Moreover, to compare the performance of all the susceptibility removing methods(TOPUP Andersson et al. (2003) Smith et al. (2004), S-Net Duong et al. (2020), Flow-net Zahneisen et al. (2020)) quantitively, the simulation with reference images is performed based on the Bloch equation. Firstly, the T2 images of RARE (image size is $256 \times 256 \times 18$) data of mice and the corresponding ground displacement map to simulate the distortion bilp up/down images, and the reference image is based on the displacement map is set to zero. A total of Nsub = 41 mice was split into Ntrain = 30 training case(30×18 slices), Ntest = 8 test case(8×18 slices) and Nval = 3 validation cases(3×18 slices).

The deep learning model was implemented on Keras for the Python 3.6 environment on an NVIDIA Geforce GTX 2080Ti with 11GB GPU memory and Intel Core CPU i7-8700 3.7GHz. An Adam optimizer35 with $\beta 1 = 0.9$, $\beta 2 = 0.999$, and $\varepsilon = 10^{-8}$ was used to train the network. The initial learning rate was set to 0.001.

3.2. Results

For Fig. 1(a) is the pulse sequence and k-space acquisition trajectory, where RF/ADC denotes the RF pulse, G_{ro}, G_{spen} and G_{ss} denote frequency, phase and slice coding direction respectively. Fig. 1(b) is a schematic diagram of our cycle consistency, the boxes on the left and right represent the distortion and undistortion domains. And Fig. 1(c) illustrate our proposed learning framework for the EPI correction.

Fig. 2 shows the simulation results with different correction methods(TOPUP Andersson et al. (2003) Smith et al. (2004), S-Net Duong et al. (2020), Flow-net Zahneisen et al. (2020)). In the simulation, the B0 inhomogeneity of the synthetic field map is set to -100 Hz to 150 Hz and -150 Hz to 100 Hz, respectively. This new method can obtain better images than the

traditional TOPUP and the other two state-of-the-art deep learning methods. This can be particularly perceived from the zoomed regions shown in the row 2 and row 4.

Fig. 3 shows the comparison of PSNR(Peak Signal-to-Noise Ratio) and SSIM(Structural Similarity) for the simulation. We can notice that our proposed method is better than TOPUP, Snet, Flownet. Fig. 3 (a) and (b) are based on the correction results of the simulation data with the field map value in the range of -200hz to 200hz. Fig. 3(c) and (d) are the PSNR and SSIM for the correction with the field map values ranging from -150 to 150. The results for all metrics are statistically significant ($p < 0.01$).
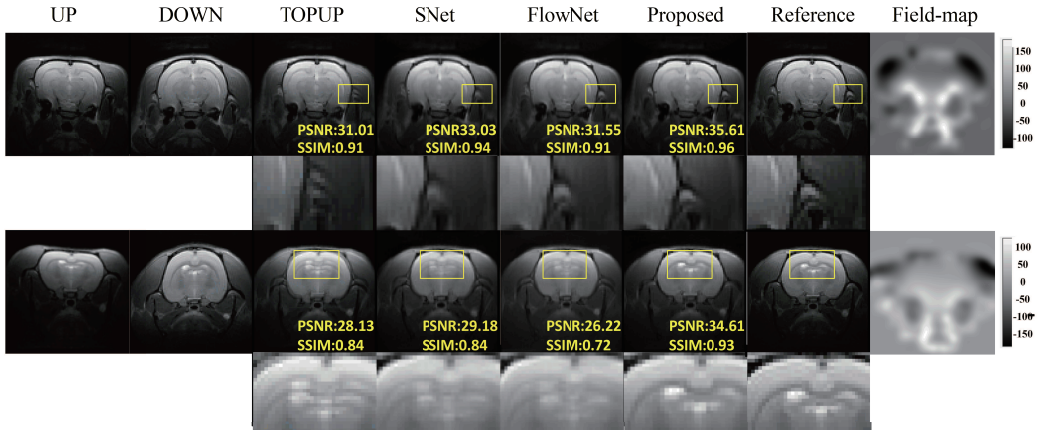
Figure 2: Correction results of various methods on one representative slice under two different degrees distortion of the simulated EPI-7T mice data set. The first two columns show the uncorrected up and down images. Column 3, 4, 5 displays the correction results of TOPUP Andersson et al. (2003) Smith et al. (2004), S-Net Duong et al. (2020), Flow-net Zahneisen et al. (2020). Columns 6 shows the correction results of unsupervised method for generating only one displacement map. And the last two column is the reference and the grand truth of the field map used for our simulation respectively. The color bar on the far right is the range of field map values. At the same time, the PSNR and SSIM values calculated using the true value of the simulation as a reference are also displayed in the lower right corner of the images.

Fig. 4 is shown the comparison of the results between different correction methods for clinical DWI-3T data set. And the results demonstrate that our method can better preserve the details of the image while removing artifacts than other methods. Moreover, it can be seen clearly that the difference map of Proposed method(with two displacement maps) is smaller than that of proposedw/t(with one displacement map). Fig. 5 is the corresponding FA map and the color map of fiber direction calculated by the FSL Smith et al. (2004).

We also applied the method to correct the susceptibility artifacts in the preclinical DWI-7T dataset. Fig. 6 shows the comparison of the results obtained by different correction methods. Fig. 7 is the corresponding FA map and the color map of fiber direction calculated by the FSL Smith et al. (2004).
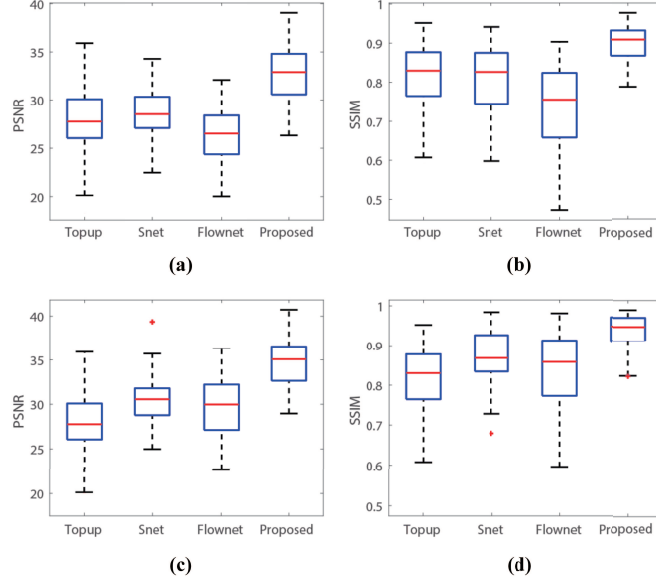
Figure 3: Quantitative comparison between distortion correction methods for two different degrees of distortion. For (a) and (b), the value of field map for simulation ranges from -200Hz to 200Hz, while the value of (c) and (d) is in the range of -150Hz to 150Hz.
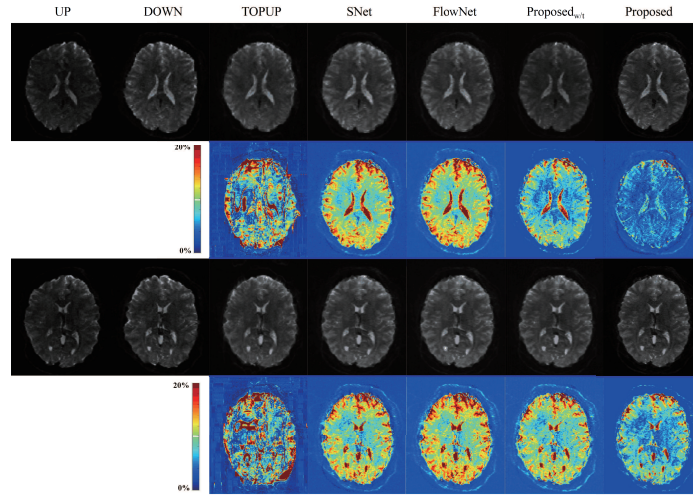

Figure 4: The comparison of the correction results for b5 images of DWI-3T test data. The 1st and 3rd row show uncorrected images and the correspond corrected images for different slices. The 2nd and 4th row display the corresponding absolute error maps between Circle UP and UP.
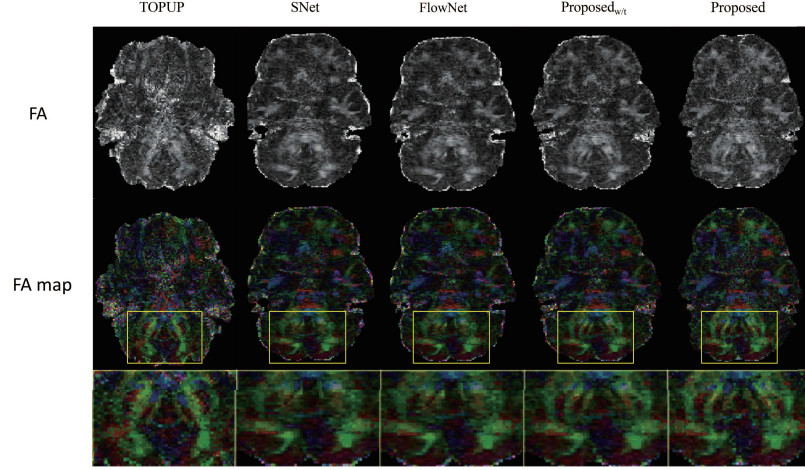
Figure 5: The comparison of the Fractional anisotropy (FA) maps based on the different correction results for HCP DWI-3T.
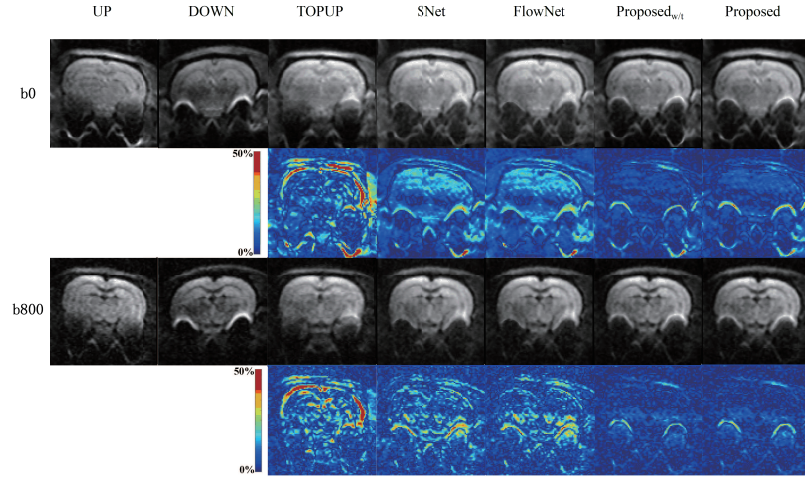

Figure 6: Comparison of the different diffusion-weighted results for different correction methods of our laboratory DWI-7T dataset. The 1st and 3rd row show the correction results under different b values, and The 2nd and 4th row display the corresponding absolute error maps of Circle Up and Up.
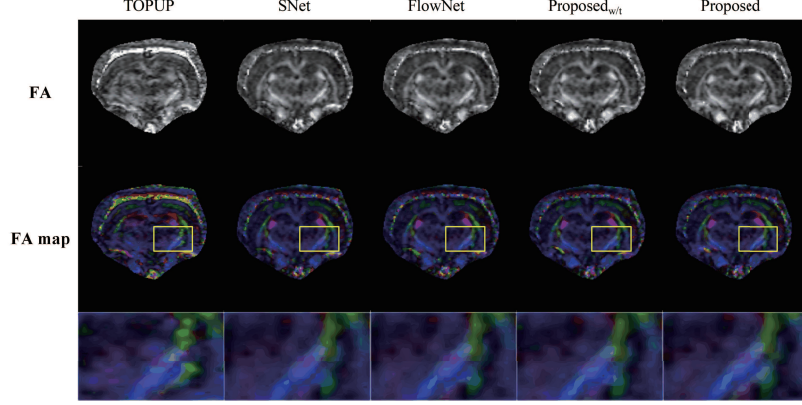
Figure 7: The comparison of the Fractional anisotropy (FA) maps based on the different correction results for our laboratory DWI-7T.

Table 2: The ablation study for our simulated data.

	mse_{Circle}	mse_{Cor}	**PSNR**	**SSIM**
3*Sim.EPI-7T		$\checkmark$	29.7997	0.8984
	$\checkmark$		30.6140	0.9037
	$\checkmark$	$\checkmark$	33.4288	0.9222

To evaluate the loss function in the segmentation network, we train our model with different loss functions, including the cycle-consistent loss(mse_{Circle}) and the forword-consistent loss(mse_{Cor}). Table 1 shows the comparsion results with different loss functions.

4. Conclusion

This paper presents a cycle-consistent deep neural network that combines the deep neural network and the gradient reversal method for removing susceptibility artifacts. The model can be performed with unsupervised training without explicit knowledge of the ground truth field map. The CNN network is trained to generate two field maps to overcome motion or field drift during the scan. The results for both 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

Acknowledgements should go at the end, before appendices and refe rences.

References

Jesper LR Andersson, Stefan Skare, and John Ashburner, editors. *How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging*, volume 20 of *Neuroimage*, 2003. Elsevier. doi: 10.1016/S1053-8119(03)00336-7. URL https://www.sciencedirect.com/science/article/abs/pii/S1053811903003367.

Hsuan Chang and J Michael Fitzpatrick, editors. *A technique for accurate magnetic resonance imaging in the presence of field inhomogeneities*, volume 11 of *IEEE transactions on medical imaging*, 1992. IEEE. doi: 10.1109/42.158935. URL https://ieeexplore.ieee.org/abstract/document/158935.

Soan TM Duong, Son L Phung, Abdesselam Bouzerdoum, and Mark M Schira, editors. *An unsupervised deep learning technique for susceptibility artifact correction in reversed phase-encoding EPI images*, volume 71 of *Magnetic resonance imaging*, 2020. Elsevier. doi: 10.1016/j.mri.2020.04.004. URL https://www.sciencedirect.com/science/article/abs/pii/S0730725X19307325.

Renaud Hedouin, Olivier Commowick, Elise Bannier, Benoit Scherrer, Maxime Taquet, Simon K Warfield, and Christian Barillot, editors. *Block-matching distortion correction of echo-planar images with opposite phase encoding directions*, volume 36 of *IEEE Transactions on Medical Imaging*, 2017. IEEE. doi: 10.1109/TMI.2016.2646920. URL https://ieeexplore.ieee.org/abstract/document/7809089.

Dominic Holland, Joshua M Kuperman, and Anders M Dale, editors. *Efficient correction of inhomogeneous static magnetic field-induced distortion in Echo Planar Imaging*, volume 50 of *Neuroimage*, 2010. Elsevier. doi: 10.1016/j.neuroimage.2009.11.044. URL https://www.sciencedirect.com/science/article/abs/pii/S1053811909012294.

Zhangxuan Hu, Yishi Wang, Zijing Dong, and Hua Guo, editors. *Water/fat separation for distortion-free EPI with point spread function encoding*, volume 82 of *Magnetic resonance in medicine*, 2019. Wiley Online Library. doi: 10.1002/mrm.27717. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.27717.

Zhangxuan Hu, Yishi Wang, Zhe Zhang, Jieying Zhang, Huimao Zhang, Chunjie Guo, Yuejiao Sun, and Hua Guo, editors. *Distortion correction of single-shot EPI enabled by deep-learning*, volume 221 of *NeuroImage*, 2020. Elsevier. doi: 10.1016/j.neuroimage.2020.117170. URL https://www.sciencedirect.com/science/article/pii/S105381192030656X.

M Okan Irfanoglu, Pooja Modi, Amritha Nayak, Elizabeth B Hutchinson, Joelle Sarlls, and Carlo Pierpaoli, editors. *DR-BUDDI (Diffeomorphic Registration for Blip-Up blip-Down Diffusion Imaging) method for correcting echo planar imaging distortions*, volume 106 of *NeuroImage*, 2015. Elsevier. doi: 10.1016/j.neuroimage.2014.11.042. URL https://www.sciencedirect.com/science/article/abs/pii/S1053811914009598.

Peter Jezzard and Robert S Balaban, editors. *Correction for geometric distortion in echo planar images from B0 field variations*, volume 34 of *Magnetic resonance*

in medicine, 1995. Wiley Online Library. doi: 10.1002/mrm.1910340111. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.1910340111.

Juyoung Lee, Kyong Hwan Jin, and Jong Chul Ye, editors. *Reference-free single-pass EPI N yquist ghost correction using annihilating filter-based low rank H ankel matrix (ALOHA)*, volume 76 of *Magnetic resonance in medicine*, 2016. Wiley Online Library. doi: 10.1002/ mrm.26077. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.26077.

Pu Liao, Jun Zhang, Kun Zeng, Yonggui Yang, Shuhui Cai, Gang Guo, and Congbo Cai, editors. *Referenceless distortion correction of gradient- echo echo-planar imaging under inhomogeneous magnetic fields based on a deep convolutional neural network*, volume 100 of *Computers in biology and medicine*, 2018. Elsevier. doi: 10.1016/j.compbiomed.2018.07.010. URL https://www.sciencedirect.com/science/article/abs/pii/S0010482518302014.

Patrick Alexander Liebig, Robin Martin Heidemann, Bernhard Hensel, and David Andrew Porter, editors. *A new approach to accelerate readout seg- mented EPI with compressed sensing*, volume 84 of *Magnetic resonance in medicine*, 2020. Wiley Online Library. doi: 10.1002/mrm.28116. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.28116.

Simin Liu, Yuhui Xiong, Erpeng Dai, Jieying Zhang, and Hua Guo, editors. *Improving dis- tortion correction for isotropic high-resolution 3D diffusion MRI by optimizing Jacobian modulation*, Magnetic Resonance in Medicine, 2021. Wiley Online Library. doi: 10.1002/ mrm.28884. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.28884.

Merry Mani, Mathews Jacob, Douglas Kelley, and Vincent Magnotta, edi- tors. *Multi-shot sensitivity-encoded diffusion data recovery using structured low-rank matrix completion (MUSSELS)*, volume 78 of *Magnetic resonance in medicine*, 2017. Wiley Online Library. doi: 10.1002/mrm.26382. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.26382.

Frank Ong, Joseph Y Cheng, and Michael Lustig, editors. *General phase reg- ularized reconstruction using phase cycling*, volume 80 of *Magnetic resonance in medicine*, 2018. Wiley Online Library. doi: 10.1002/mrm.27011. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.27011.

Stephen M Smith, Mark Jenkinson, Mark W Woolrich, Christian F Beckmann, Timothy EJ Behrens, Heidi Johansen-Berg, Peter R Bannister, Marilena De Lu- ca, Ivana Drobnjak, David E Flitney, et al., editors. *Advances in function- al and structural MR image analysis and implementation as FSL*, volume 23 of *Neuroimage*, 2004. Elsevier. doi: 10.1016/j.neuroimage.2004.07.051. URL https://www.sciencedirect.com/science/article/abs/pii/S1053811904003933.

Colin Studholme, R Todd Constable, and James S Duncan, editors. *Accurate alignment of functional EPI data to anatomical MRI using a physics-based distortion model*, volume 19 of *IEEE transactions on medical imaging*, 2000. IEEE. doi: 10.1109/42.896788. URL https://ieeexplore.ieee.org/abstract/document/896788.

Yuhui Xiong, Guangqi Li, Erpeng Dai, Yishi Wang, Zhe Zhang, and Hua Guo, editors. *Distortion correction for high-resolution single-shot EPI DTI using a modified field-mapping method*, volume 32 of *NMR in Biomedicine*, 2019. Wiley Online Library. doi: 10.1002/nbm.4124. URL https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/abs/10.1002/nbm.4124.

Seong Dae Yun and N Jon Shah, editors. *Analysis of EPI phase correction with low flip-angle excitation to reduce the required minimum TE: Application to whole-brain, submillimeter-resolution fMRI at 3 T*, volume 84 of *Magnetic resonance in medicine*, 2020. Wiley Online Library. doi: 10.1002/mrm.28218. URL https://onlinelibrary.wiley.com/doi/full/10.1002/mrm.28218.

Benjamin Zahneisen, Kathrin Baeumler, Greg Zaharchuk, Dominik Fleischmann, and Michael Zeineh, editors. *Deep flow-net for EPI distortion estimation*, volume 217 of *NeuroImage*, 2020. Elsevier. doi: 10.1016/j.neuroimage.2020.116886. URL https://www.sciencedirect.com/science/article/pii/S1053811920303724.