Deep Learning Based Self-Navigated Diffusion Weighted Multi-Shot EPI with Supervised Denoising

Yiming Dong[1], Kirsten Koolstra[2], Laurens Beljaards[3], Marius Staring[3], Matthias J.P. van Osch[4], and Peter Börnert[4,5]
[1]LUMC, Leiden, Netherlands, [2]Philips, Best, Netherlands, [3]Division of Image Processing, Department of Radiology, LUMC, Leiden, Netherlands, [4]C.J. Gorter MRI Center, Department of Radiology, LUMC, Leiden, Netherlands, [5]Philips Research, Hamburg, Germany

Synopsis

Advanced diffusion weighted self-navigated multi-shot MRI can run at high scan efficiencies resulting in good image quality. However, the model-based image reconstruction is rather time consuming. Deep learning-based reconstruction approaches could function as a faster alternative. Tailored network architectures with appropriately set physical model constraints can help to shorten reconstruction times, resulting in good image quality with reduced noise propagation.

Introduction

Single-shot EPI is one of the standard protocols for diffusion-weighted imaging (DWI) in clinical practice[1]. However, the relatively long readout time of single-shot acquisitions may result in signal loss, image blur, and large geometric distortions. Consequently, multi-shot EPI (msh-EPI) has become an increasingly preferred method for DWI to produce high-resolution images with less geometric distortion[2,3]. However, an inherent challenge for msh-EPI-based DWI is the occurrence of shot-to-shot phase errors induced by physiological motion in presence of the strong diffusion-sensitizing gradients. Many studies have proposed to use either additionally measured navigators[2] or self-navigation[4,5] corrections to deal with such phase errors. However, these methods may either prolong scan-times or reconstruction times. In this work, a neural network with two U-net architectures was trained to reconstruct a shot-specific phase map for each shot and one joint magnitude image. Those were merged, including physical model constraints, in the joint loss calculation of the two U-nets[6,7]. This approach not only realizes fast image reconstruction, but also benefits from image denoising due to the choice of training data and the nature of convolutional neural networks (CNNs). In this work, the training pairs for DWI were simulated using T2w leg images with extra noise added to mimic the lower SNR in diffusion measurements.

Methods

Inspired by many self-navigation approaches in DWI, in this work a joint image constraint was set up by assuming all different shots share one underlying magnitude representation in combination with different, shot-specific phases. Two separate U-net-like CNNs for magnitude and phase components were trained together using a combined loss function (see Fig.1). The outputs of the two individual CNNs were concatenated to calculate a joint loss. The signal model of the msh-EPI-based DWI can be expressed as:

$$S_{i,l} = K_i F C_l e^{j\phi_i} I$$

where $S_{i,j}$ is the k-space data of the i-th shot and l-th coil, K_i is the sampling operator, F is the Fourier transform operator, C_l is the coil sensitivity, and ϕ_i and I are the shot-specific phase maps and the joint magnitude that need to be estimated. The task of the network is to minimize:

$$\text{loss}(I, \phi_i) = \sum_i \sum_l \left\| f_m(X) \cdot e^{j f_p(P)_i} \cdot C_l - M_{i,l} \right\|_2^2$$

where $f_m(X) = \mathrm{I}$ is the joint magnitude predicted by the first "magnitude" U-net, and $f_p(P)_i = \phi_i$ is the *i*-th shot's phase map predicted by the second "phase" U-net The aliased magnitude images X and phase images P are the separate inputs of the two U-nets, and are calculated by taking the magnitude and phase from the inverse Fourier transform of each shot i and coil l of zero-filled k-space data $S_{i,l}$. $M_{i,l}$ denotes the ground truth image for the loss calculation, which can be calculated via $M_{i,l} = C_l e^{j\phi_i} I$. The architecture of the network is shown in Fig. 1.

The training data were generated from b=0 s/mm^2 leg images (from 18 volunteers) measured in the knee/lower leg (3T, Philips, Best, The Netherlands). The scan parameters can be found in Table 1. In this work, 4-shot and 8-coils diffusion data were simulated for training. To generate training inputs, each corresponding series of the 4-shot images (and 8 coils) were undersampled with R=4 according to the 4-shot msh-EPI trajectory. Data augmentation was performed by randomly applying in-plane rotation, flipping, and resizing. Varying shot-specific phases were simulated based on a second-order random Gaussian profile with spatial variation between ± π. To simulate low SNR at higher b-values, random complex Gaussian noise was added to the training input for SNR levels ranging between 8 to 26. Two different models were trained with/without synthetic noise added to the target sets, to obtain two networks with/without strong denoising ability. In total, 1120 different 4-shots DW synthetic training pairs were generated, which are divided into 32 aliased magnitude/phase images as inputs for each U-net, and 32 complex fully k-space data as target sets. The loss function calculation is illustrated in Fig. 2. The training was run for 400 epochs with a batch size of 5, using the Adam optimizer with a learning rate of 0.0001.

Test data were measured using a fat-suppressed DW 4-shot msh-EPI sequence of 2 volunteers' leg on the 3T scanner. The deep learning results were compared to model-based results using extra-navigated[2] and two self-navigated methods (POCS-ICE[4], MUSSELS[5]).

Results

Fig. 3 shows the reconstructions and corresponding ADC maps of three b-values, qualitatively comparing no-navigation, extra-navigated, and deep learning reconstruction with denoising training on/off. Fig. 4 shows comparisons between different methods (2D-navigated[2],MUSSELS[4],POCS-ICE[5]). The network outperformed all the other methods in terms of reconstruction speed (see the caption of Fig.4.). The reduced noise floor is appreciated as well.

Discussion and conclusion

The fully deep learning-based multi-shot EPI reconstruction method proposed in this work demonstrates good preliminary results in the leg region based on simulated training pairs. This helps to accelerate the reconstruction of a joint full-k-space magnitude image compared to traditional model-based solutions. Specifically, the well-known challenge of correcting shot-to-shot phase variations in multi-shot acquisitions was addressed by using the "double" U-net architecture. In addition, due to the inherent properties of CNNs, denoising capability was also trained in the network. The next step will be to train on other anatomies (e.g. the brain), to test the robustness of the method and to prevent oversmoothing. More quantitative analysis is necessary and should be subject of future work.

Acknowledgements

The authors would like to acknowledge NWO-TTW (HTSM-17104).

References

1. Le Bihan D, Mangin J- F, Poupon C, et al. Diffusion tensor im- aging: concepts and applications. J Magn Reson Imaging. 2001;13:534- 546.

2. Jeong H-K, Gore JC, Anderson AW. High-resolution human diffusion tensor imaging using 2-D navigated multishot SENSE EPI at 7 T. MRM. 2013;69(3):793-802.

3. Butts, K., Pauly, J., De Crespigny, A. and Moseley, M. (1997), Isotropic diffusion-weighted and spiral-navigated interleaved EPI for routine imaging of acute stroke. Magn Reson Med., 38: 741-749.

4. Guo H, Ma X, Zhang Z, Zhang B, Yuan C, Huang F. POCS-enhanced inherent correction of motion-induced phase errors (POCS-ICE) for high-resolution multi-shot diffusion MRI. Magn Reson Med. 2016;75(1):169-180.

5. Mani M, Jacob M, Kelley D, Magnotta V. Multi-shot sensitivity-encoded diffusion data recovery using structured low-rank matrix completion (MUSSELS). Magn Reson Med. 2017;78(2):494-507.

6. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention. Berlin: Springer; 2015. p 234–241.

7. Jafari R, Spincemaille P, Zhang J, et al. Deep neural network for water/fat separation: Supervised training, unsupervised training, and no training. Magn Reson Med. 2021;85(4):2263-2277.

8. Dong Y, Koolstra K, Riedel M, van Osch MJP, Börnert P. Regularized joint water–fat separation with B0 map estimation in image space for 2D-navigated interleaved EPI based diffusion MRI. Magn Reson Med. 2021; 00: 1– 18.

Figures

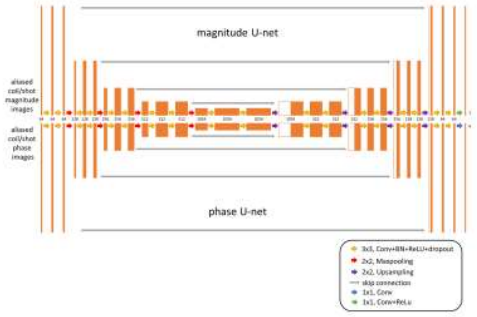

Figure 1. DL architecture. The two U-nets have almost the same structure. The only difference is a ReLu activation applied for the output layer of the magnitude U-net to ensure a positive-valued output of the magnitude image, while 4 phase maps as the output of the phase network. The dropout ratio for each layer is set as 0.05. The inputs of the two U-nets are aliased magnitude/phase images of each shot-coil (4 shots and 8 coils), so 32 input channels of each U-net respectively.

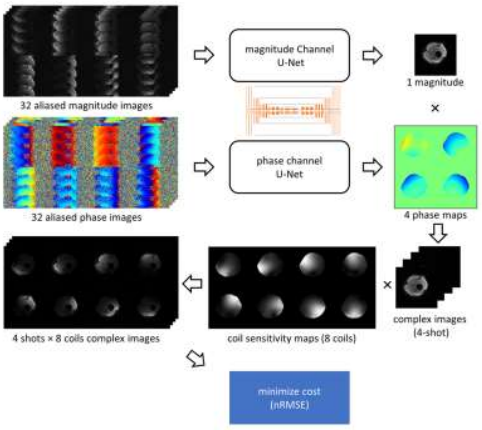

Figure 2. Illustration of how to calculate the combined loss. The individual full-k-space reconstructed images were multiplied with the simulated 4-phase maps and measured 8-element coil sensitivities to generate 32 complex images as ground truth in the loss calculation. The output magnitude image and four phase maps were combined to 4 complex images, then multiplied with 8 elements coil sensitivity maps to generate 32 complex images and minimize the nRMSE loss with the ground truth.

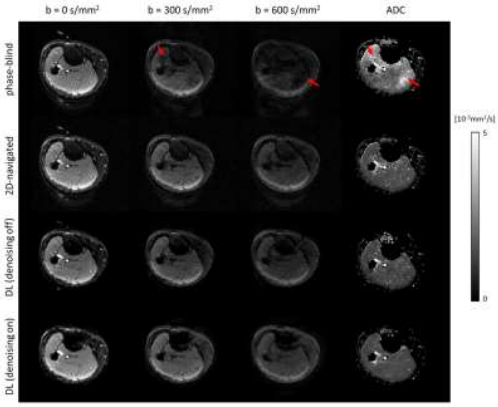

Figure 3. Results of leg DW images with corresponding ADC quantification. Different reconstructions are compared, using no-navigation (phase-blind), 2D-navigator, and deep learning with denoising property on/off. The shot-to-shot phase variations create signal loss in the phase-blind image, resulting in abnormal high ADC values in such regions (red arrow). These can be corrected by using 2D-navigated or deep learning-based reconstructions, while the images can be further denoised by using deep learning and generating smoother DW images/ADC map.

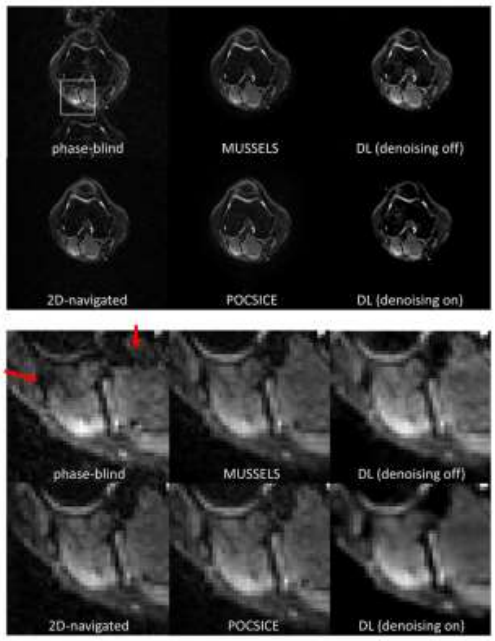

Figure 4. Results for a knee DW images (b = 300 s/mm2) in a healthy volunteer. Different reconstructions are compared, using no-navigation (phase-blind), 2D-navigator, MUSSELS, POCS-ICE, and deep learning (DL) with denoising on/off. Some artefacts show up in the chosen ROI (white square) of the phase-blind reconstruction, which can be avoided by using other options. MUSSELS and POCS-ICE both spent roughly 20 s to reconstruct such slice, while DL only needs 0.4 s. all reconstruction were done with an Intel Core i7 CPU (3.0 GHz, 8 cores).

	sequence	number of volunteers	resolution (mm^3)	shots	TE/TR (ms)	slice	fat suppression
training	chemical-shift encoded spin-echo msh-EPI	9	1.5 x 1.5 x 4	6	62/2000	4	Dixon
		4	1.2 x 1.2 x 4	6	64/2000	4	Dixon
		5	1.5 x 1.5 x 4	4	69/2000	4	Dixon
testing	DW msh-EPI	2	1.5 x 1.5 x 4	4	69/2000	8	SPAIR

Table 1. Scan parameters. The reconstruction for the training sets were performed using a water/fat separation algorithm for Dixon EPI to generate fat-free fully sampled T2w images. Then these images were segmented in k-space to simulate the 4-shot msh-EPI synthetic data for training.