Retrospective motion correction for Fast Spin Echo based on conditional GAN with entropy loss

Yalei Chen1,2, Pingan Li1,2, Kewen Liu1,2, Qingjia Bao4*, Zhao Li3 and Xiaojun Li1,2

1Information and Communication Engineering, Wuhan University of Technology, Wuhan, Hubei, 430070, China
2Hubei Key Laboratory of Broadband Wireless Communication and Sensor Networks, Wuhan University of Technology, 430070 Wuhan, China
3State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Wuhan Center for Magnetic Resonance, Wuhan Institute of Physics and Mathematics, Innovation Academy for Precision Measurement Science and Technology, 430071Wuhan, China
4Department of Chemical and Biological Physics, Weizmann Institute of Science, Rehovot, Israel

*Corresponding author’s e-mail: qingjia.bao@weizmann.ac.il

Abstract. We proposed a new end-to-end motion correction method based on conditional generative adversarial network (GAN) and minimum entropy of MRI images for Fast Spin Echo (FSE) sequence. The network contains an encoder-decoder generator to generate the motion-corrected images and a PatchGAN discriminator to classify an image as either real (motion-free) or fake (motion-corrected). Moreover, the image’s entropy is set as one loss item in the cGAN’s loss as the entropy increases monotonically with the motion amplitude, indicating that entropy is a good criterion for motion. The results show that the proposed method can effectively reduce the artifacts and obtain high-quality motion-corrected images from the motion-affected images in both pre-clinical and clinical datasets.

1. Introduction

Magnetic resonance imaging (MRI) is one of the important imaging technique for clinical and pre-clinical research since it provides good spatial resolution without the need for harmful radiation[1]. Specifically, fast spin-echo (FSE)[2] sequences are commonly used for MRI because of the ability to prescribe various image in using a large number of shots at different time instances to obtain the high-resolution volumetric image. But the image may be severely degraded due to subject motion between consecutive shots, especially for pediatric or stroke patients in clinical and awake rodents in pre-clinical studies. Any impairment by motion artifacts can reduce the reliability and precision of the diagnosis and a motion-free reacquisition can become time- and cost-intensive. The conventional motion correction methods can be divided into retrospective and prospective motion correction, and most of them need to predict the motion model, or manually outline the artifact area, or need for extra hardware facilities[1][3]. Recently, more and more researchers began to use deep learning as a tool for MRI motion correction[4] because it can provide a potential avenue for dramatically reducing the computation time and improving
the convergence of retrospective motion correction methods. And several groups have proposed the use of GAN for motion correction due to its capability of generating data without the explicit modeling of the probability density function and robustness to over-fitting[5]. In this work, we propose a new end-to-end motion correction method based on conditional GAN (cGAN)[6] and minimum entropy[7] of images for multi-shot FSE sequence, train the network with the simulated random rigid motion data, and apply this network for both clinical and pre-clinical studies. Moreover, we evaluate the new motion correction method on the same weights of the network through the different number of shot for FSE, the different motion pattern such as Markov pattern[8] and periodic pattern[9], and different level motion.

Our contributions
1) We propose a motion correction model that can be used for both pre-clinical (mouse brain) and clinical (human brain) dataset, which can effectively reduce the motion artifacts.

2) We introduced the minimum entropy of MR images into the loss function of our methods.

2. Related Work
The generative adversarial network (GAN)[10] which consists of a generator and a discriminator is a hot research direction in the field of deep learning, in which the generative network generates candidates while the discriminative network evaluates them. Due to its capability of generating data without the explicit modeling of the probability density function and its robustness to over-fitting[5], several groups have proposed the use of GAN for motion correction. Recently, Küstner et al. proposed MedGAN and Cycle-MedGAN with both supervised and unsupervised learning which proved the GAN based retrospective restoration of motion artifacts is feasible resulting in near-realistic motion-free images[11][12]. Jiang et al. proposed a novel model for the abdominal MRI motion correction of FSE and GRE pulse sequence, which incorporating the widespread densely connected UNet with a GAN-guided training schema and perceptual loss function[13]. Usman et al. proposed a novel generative adversarial network (GAN)-based conjugate gradient SENSE (CG-SENSE) reconstruction framework for motion correction in multishot MRI[14].

3. Proposed Method

3.1. The network architecture based on cGAN
The overall architecture of the proposed cGAN-based method which simultaneously train the generator and the discriminator for FSE sequence motion correction is shown in Figure 1(a). The input data of the generator is motion-affected images, and the outputs are the motion-corrected images which contains less artifacts. The input of the discriminator are both the motion-free images and motion-corrected images, and if the input is motion-free, the output of the discriminator is 1, and if the input is motion-corrected, the output is -1.

We designed a residual encoder-decoder generator as shown in Figure 1(c). It contains five encoder blocks, seven residual blocks (ResBlocks), and five decoder blocks. Moreover, concatenations were applied between the same scale feature maps from the encoder and decoder, which allow the network to propagate context information to higher resolution layers. Finally, we introduce the global skip connection, which learns residual artifacts images to ensure train the network faster and model generalizes better. The reason for adopting the residual learning is that, it is easier to optimize the residual mapping than optimizing the direct mapping, and can avoid the gradient vanishing during training when the network is deep.

The architecture of discriminator is identical to PatchGAN[6], used 5 cascaded convolution layer, which is sufficient to restrict the attention to the structure in local image patches. The discriminator divides the input from image to patch, then classifies each patch as either motion-free or motion-corrected, and finally averages out the score of all the image patches as the output.
3.2. Loss functions combined with MR image entropy

Our model training is driven by loss functions, which are composed of adversarial loss and content loss:

\[ L = L_{\text{WGAN}} + \lambda_m L_{\text{mse}} + \lambda_e E_v \]  

(1)

The \( L_{\text{WGAN}} \) is the adversarial loss, and the \( L_{\text{mse}} \) and \( E_v \) equal content loss. We use mean-square error (MSE) loss for the content loss plus the weight of the minimum entropy loss of the MR image, as shown in Equation (1). The \( \lambda_m \) is 0.4 and the \( \lambda_e \) is 0.6.

The minimum entropy loss of MR image is based on the entropy focusing motion correction method proposed by Atkinson[7], as the entropy increases monotonically with the motion amplitude, shown in Figure 1(d). The formula for MR image entropy constraint is as follows:

\[ E = -\sum_{j=1}^{s} \frac{B_j}{B_{\text{max}}} \ln \left( \frac{B_j}{B_{\text{max}}} \right) \]  

(2)

where the \( B_{\text{max}} = \sum_{j=1}^{s} B_j^2 \), \( s \) is the total number of pixels in the motion-corrected image, \( B_j \) is the gray value of the \( j \)-th pixel. The minimum entropy constraint criterion of the image space domain is:

\[ E_v = \min \left( E_v \right) \]  

(3)

The Wasserstein GAN gradient penalty (WGAN-GP)[15] loss was used for the adversarial loss. WGAN-GP adopts Wasserstein distance rather than the JS divergence of the original GAN, which can avoid the gradient vanished problem. The loss function is as follows:

\[ \min_{D} \max_{G} L_{\text{WGAN}} (D, G) = -E_x \left[ D(x) \right] + E_z \left[ D \left( G(z) \right) \right] + \lambda_p E_{\xi} \left[ \left\| \nabla_{\xi} D(\hat{x}) \right\|_2^2 - 1 \right]^2 \]  

(4)

where the \( E(\cdot) \) denotes the expectation operator, the first two indicate the Wasserstein distance estimation, and the last one is the gradient penalty term for network regularization; \( x \) is uniformly sampled along straight lines connecting pairs of generated and real samples; and \( \lambda_p \) is a constant weighting parameter.
3.3. FSE Motion simulation
The FSE sequence images were rigidly transformed based on the six motion parameters (three translations and three rotation parameters) for each time point (phase encode). As the images moving continuously, we extracted 8 or 16 k-space lines at the current time point (phase encode) and appended them to the motion-corrupted k-space. And which k-space lines to be chosen at the current time is determined by the encoding trajectory of the FSE sequence.

We simulated three kinds of motion patterns based on Markov random process motion, periodic motion, and completely random motion. For the Markov motion pattern, the position of the next point of the trajectory is randomly generated based on the previous point velocity and position. The periodic pattern using a simple sine wave with random frequency, phase, and duration was used to simulate the ghosting artifact. And completely random pattern motions with random rotation and with these motion patterns between ±2 mm of translation on the x, y direction, ±1 mm on the z direction, and ± 2° of rotation on all three axes.

3.4. Experimental data acquisition and training
We did both pre-clinical experiments and clinical experiments. The pre-clinical data are rat brain images collected by fast spin-echo (FSE) sequences on Bruker’s 7.0 Tesla scanner in our local laboratory, containing 400 mice brain images. The FOV is 150mm×150mm; the acquisition matrix’s size is 256×256. We randomly choose 210 subjects for the pre-clinical data, 14569 slices as a training set, 10 subjects, 743 slices as the test set. Ten different random motions were simulated for each subject, providing 2100 subjects for training and 100 subjects for testing. We take one slice every ten slices from each subject. For in vivo test to evaluate our model, we did experiments with the awake and uncontrolled rodent.

The clinical dataset is the public dataset HCP Brain dataset. This dataset contains 35 subjects. We randomly choose 15 subjects as the training dataset, 3 subjects as testing. 25 different random motions were simulated for each subject; then, we had 225 subjects for training, 75 subjects for testing. The motion method is as same as the pre-clinical data. Each subject contains 176 slices, and we choose the 70th to 100th slices in the middle of one subject. Our training dataset contains 11250 slices, and the testing dataset contains 2250 slices.

For the training, we train for 100 epochs using ADAM, the initial learning rate is set to 0.001, the first-order momentum to be 0.9, and the second momentum to be 0.999. In each epoch, we train the discriminator every five times and train the generator once. Both the motion-affected image and the motion-free images were normalized, intensity values lie between 0 and 1. The convolutional layers are initialized in the form of glorot-uniform. The number of batchsize is 8. Our deep learning models are implemented on Keras for the Python 3.6 environment on two NVIDIA GeForce GTX 2080Ti with 11GB GPU memory and Intel Core CPU i7-8700 3.7GHz.

4. Results
Figure 2 shows our method’s results of mild and strong motion and corresponding error maps for 2 representative subjects, compared with 2 other deep learning methods (DnCNN, U-net), 1 conventional method (TV Denoiser). Although the network was trained in the range of ±2 mm translation in the x, y direction, ±1 mm translation in the z direction, and ± 2° rotation in all direction, the predictions were performed on a wider motion range (±4 mm in xy, ±2mm in z, and ± 4°) to test the generalization capability of our method to unseen cases. The figure illustrates that our method can obtain better images, which means our method results in the most effective motion artifacts reduction. The detailed texture information of the motion-corrected image corrected by our model is more abundant than the other deep learning models. This can be particularly perceived from the zoomed regions shown in the left-up corner and from the error maps (representing the difference between the reference map and the motion-corrected images obtained by motion correction methods). For the quantitative evaluation, we compared several image evaluation indices (SSIM, PSNR, MSE), which were highest for our method reconstructions, as shown in Table 1.
Figure 2 Mild and strong motion correction results of various methods on multi-shot FSE sequence. Columns of the first is the mild motion result and the third is strong. The columns from left to right show reference images, motion-affected images, motion corrected images using TV Denoiser, DnCNN, UNET, and our method, respectively. In the second and fourth row, the absolute error maps corresponding to the first and third row are presented.

Results of motion artifact reduction of clinical dataset were shown in Figure 3. The simulated motion-affected both contains blurring and ringing, the DnCNN can reduce the ringing artifacts and invalid for the blurring, the results of DnCNN are still blurred. UNET and our method are good at both blurring and ringing. As the zoomed regions of the brain stem show, our method’s result is clearer than the UNET. In terms of PSNR/SSIM/MSE quantitative metrics, our method reached the highest number.

Table 1. Quantitative results

| Method     | SSIM(mild/strong) | PSNR(mild/strong) | MSE(mild/strong) |
|------------|-------------------|-------------------|------------------|
| TV denoiser| 0.6534/0.6019     | 19.34/17.01       | 0.0037/0.0052    |
| DnCNN      | 0.8013/0.7847     | 29.45/25.09       | 0.0018/0.0025    |
| UNET       | 0.8434/0.8403     | 34.63/30.72       | 0.0013/0.0019    |
| Our method | **0.9521/0.9238** | **43.01/36.52**   | **0.0002/0.0008** |
Figure 3 Clinical motion correction results of various methods. Columns of the first and third rows show the different slices of the same subject. From left to right, they show the reference images, motion-affected images, and motion-corrected images using two deep learning models (DnCNN, UNET) and our method, respectively. In the second and fourth row, the absolute error maps corresponding to the first and third row are presented.

5. Conclusion
In this manuscript, we presented a method for restoration of motion-affected multi-shot FSE MR images using cGAN based method, which combined the MR image’s entropy minimization loss. We can achieve the near-realistic restoration of motion-affected MR images and obtain the motion-corrected images that display high similarity to acquired motion-free MR data. Our method outperforms UNET, DnCNN, and TV denoiser in terms of qualitative and quantitative comparisons for the simulation data of pre-clinical and clinical datasets, as shown in the figures above. The reason is probably that our method is powerful and deep enough to decompose all motion features in the motion-affected image, and our method exploits more available information within the MR images (the entropy loss). Compared with the conventional prospective and retrospective method, our method does not need to predict the motion model nor manually outline the artifact area, no need for extra hardware facilities, and can remove the artifact end-to-end.

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 (Grant 2018YFC0115000, 2016YFC1304702), National Natural Science Foundation of China (11575287, 11705274), and the Chinese Academy of Sciences (YZ201677).

Reference
[1] K. Ruppert, D. L. G. Hill, P. G. Batchelor, and M. Holden, “reconstruction Motion correction in MRI of the brain,” doi: 10.1088/0031-9155/61/5/R32.
[2] J. G. Pipe, V. G. Farthing, and K. P. Forbes, “Multishot diffusion-weighted FSE using
PROPELLER MRI,” *Magn. Reson. Med.*, 2002, doi: 10.1002/mrm.10014.

[3] K. A. Salem, “MRI Hot Topics Motion Correction for MR Imaging s medical.”

[4] K. Pawar, Z. Chen, N. J. Shah, and G. F. Egan, “Suppressing motion artefacts in MRI using an Inception- ResNet network with motion simulation augmentation,” no. October, pp. 1–14, 2019, doi: 10.1002/nbm.4225.

[5] J. Luo and J. Huang, “Generative adversarial network: An overview,” *Yi Qi Yi Biao Xue Bao/Chinese Journal of Scientific Instrument*. 2019, doi: 10.19650/j.cnki.cjsi.J1804413.

[6] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-image translation with conditional adversarial networks,” 2017, doi: 10.1109/CVPR.2017.632.

[7] D. Atkinson, D. L. G. Hill, P. N. R. Stoyle, P. E. Summers, and S. F. Keevil, “Automatic correction of motion artifacts in magnetic resonance images using an entropy focus criterion,” *IEEE Trans. Med. Imaging*, 1997, doi: 10.1109/42.650886.

[8] O. Kupyn, V. Budzan, M. Mykhailych, and D. Mishkin, “DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks.”

[9] D. Tamada, M. Kromrey, S. Ichikawa, and H. Onishi, “MAJOR PAPER Motion Artifact Reduction Using a Convolutional Neural Network for Dynamic Contrast Enhanced MR Imaging of the Liver,” pp. 1–13, 2019, doi: 10.2463/mrms.mp.2018-0156.

[10] I. J. Goodfellow et al., “Generative adversarial nets,” 2014, doi: 10.3156/jsoft.29.5_177_2.

[11] K. Armanious, S. Gatidis, K. Nikolaou, B. Yang, and K. Thomas, “RETROSPECTIVE CORRECTION OF RIGID AND NON-RIGID MR MOTION ARTIFACTS USING GANS University of Stuttgart , Institute of Signal Processing and System Theory , Stuttgart , Germany University of T ” ubingen , Department of Radiology , T ” King ’ s College Londo,” no. Isbi, pp. 1550–1554, 2019.

[12] K. Armanious, A. Tanwar, S. Abdulatif, K. Thomas, S. Gatidis, and B. Yang, “UNSUPERVISED ADVERSARIAL CORRECTION OF RIGID MR MOTION ARTIFACTS University of Stuttgart , Institute of Signal Processing and System Theory , Stuttgart , Germany University of T ” ubingen , Department of Radiology , T ” King ’ s College London , Biomedica,” pp. 0–4, 2020.

[13] W. Jiang, Z. Liu, K. Lee, S. Chen, Y. Ng, and Q. Dou, “Respiratory Motion Correction in Abdominal MRI using a Densely Connected U-Net with GAN-guided Training Respiratory Motion Correction in Abdominal MRI using a Densely Connected U-Net with GAN-guided Training Clinical magnetic resonance imaging ( MRI ) oft,” no. November, 2019.

[14] M. Usman, S. Latif, M. Asim, and B. Lee, “Retrospective Motion Correction in Multishot MRI using Generative Adversarial Network,” *Sci. Rep.*, pp. 1–11, 2020, doi: 10.1038/s41598-020-61705-9.

[15] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, “Improved training of wasserstein GANs,” 2017.