Magnetic Resonance Imaging 71 (2020) 93–102

Contents lists available at ScienceDirect

Magnetic Resonance Imaging

journal homepage: www.elsevier.com/locate/mri

Original Contribution

Correction of out-of-FOV motion artifacts using convolutional neural network

Chengyan Wang[a,b], Yucheng Liang[c], Yuan Wu[b], Siwei Zhao[b], Yiping P. Du[b,*]

[a] Human Phenome Institute, Fudan University, Shanghai, China
[b] Institute for Medical Imaging Technology (IMIT), School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China
[c] Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

ARTICLE INFO

Keywords:
Motion correction
Machine learning
Convolutional neural networks
Out-of-FOV motion
Prior image

ABSTRACT

Purpose: Subject motion during MRI scan can result in severe degradation of image quality. Existing motion correction algorithms rely on the assumption that no information is missing during motions. However, this assumption does not hold when out-of-FOV motion happens. Currently available algorithms are not able to correct for image artifacts introduced by out-of-FOV motion. The purpose of this study is to demonstrate the feasibility of incorporating convolutional neural network (CNN) derived prior image into solving the out-of-FOV motion problem.

Methods and materials: A modified U-net network was proposed to correct out-of-FOV motion artifacts by incorporating motion parameters into the loss function. A motion model based data fidelity term was applied in combination with the CNN prediction to further improve the motion correction performance. We trained the CNN on 1113 MPRAGE images with simulated oscillating and sudden motion trajectories, and compared our algorithm to a gradient-based autofocusing (AF) algorithm in both 2D and 3D images. Additional experiment was performed to demonstrate the feasibility of transferring the networks to different dataset. We also evaluated the robustness of this algorithm by adding Gaussian noise to the motion parameters. The motion correction performance was evaluated using mean square error (NMSE), peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM).

Results: The proposed algorithm outperformed AF-based algorithm for both 2D (NMSE: 0.0066 ± 0.0009 vs 0.0141 ± 0.008, $P < .01$; PSNR: 29.60 ± 0.74 vs 21.71 ± 0.27, $P < .01$; SSIM: 0.89 ± 0.014 vs 0.73 ± 0.004, $P < .01$) and 3D imaging (NMSE: 0.0067 ± 0.0008 vs 0.070 ± 0.021, $P < .01$; PSNR: 32.40 ± 1.63 vs 22.32 ± 2.378, $P < .01$; SSIM: 0.89 ± 0.01 vs 0.62 ± 0.03, $P < .01$). Robust reconstruction was achieved with 20% data missed due to the out-of-FOV motion.

Conclusion: In conclusion, the proposed CNN-based motion correction algorithm can significantly reduce out-of-FOV motion artifacts and achieve better image quality compared to AF-based algorithm.

1. Introduction

Voluntary and involuntary subject motion during MRI scan can result in severe degradation of image quality, especially when high-resolution multi-shot imaging is performed [1]. Inconsistent *k*-space data can produce variety of artifacts such as blurring, ghosting and signal loss, which may lead to errors for subsequent image analysis and diagnosis. For some cases, rescan would be necessary if the motion artifacts are severe, which results in prolonged scanning time.

Due to the prevalence of motion in MRI, many techniques have been proposed to correct or compensate for the motion artifacts [2–19]. They can be broadly categorized as prospective [2–6] and retrospective [7–19] approaches. The prospective motion correction (PMC) methods are capable of compensating through-plane motion artifacts and avoiding spin history effects during scanning. However, PMC often requires additional hardware (e.g., optical motion tracking device) and calibrations between the external-device coordinates and the scanner coordinates, which is a concern for routine applications [20,21]. MR

Abbreviations: PMC, prospective motion correction; AF, autofocusing; FOV, field-of-view; CNN, convolutional neural network; TV, total variation; CG, conjugate gradient; HCP, human connectome project; MPRAGE, magnetization-prepared rapid-acquisition with gradient-echo; ReLu, rectified linear unit; NMSE, normalized mean square error; PSNR, peak signal-to-noise ratio; SSIM, structural similarity index; SNR, signal-to-noise ratio.

* Corresponding author at: Institute for Medical Imaging Technology (IMIT), School of Biomedical Engineering, Shanghai Jiao Tong University, China.
E-mail addresses: wangcybit@126.com (C. Wang), yipingdu@sjtu.edu.cn (Y.P. Du).

https://doi.org/10.1016/j.mri.2020.05.004
Received 18 February 2020; Accepted 14 May 2020

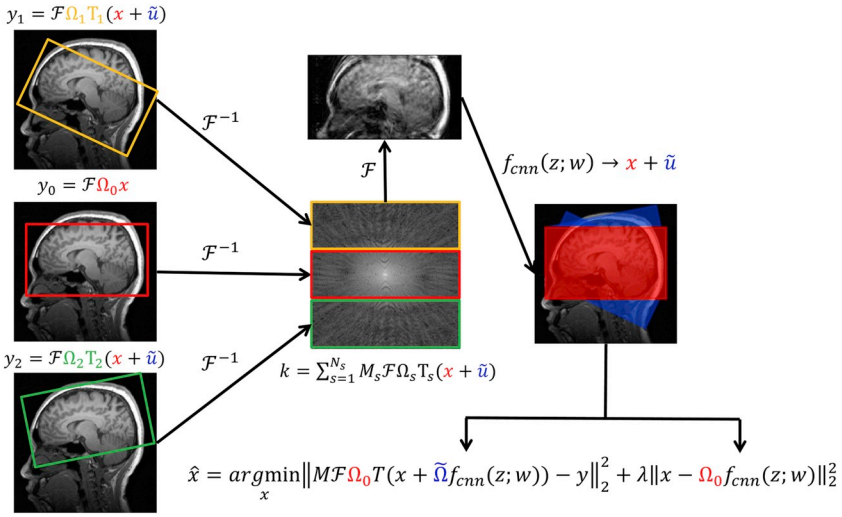

Fig. 1. Illustration of the CNN-based motion correction algorithm.

navigator-based PMC techniques, on the other hand, do not have the described limitations and are thus more convenient to be used. However, navigator methods [3,4] require additional time within the sequence for the calculation of motion parameters, which makes the estimated motion parameters less reliable for motion compensation. Sometimes the requirement for additional delay time even makes it incompatible with some fast sequences. Besides, gradient non-linearity appears to be another limiting factor for PMC techniques, in particular when dealing with large-scale motion.

Retrospective techniques, on the other hand, attempt to remove motion artifacts after scans. This can be done by introducing a priori motion information (obtained from navigator signal or external sensors) into the degradation model [7–11], or applying an autofocusing (AF) algorithm with respect to certain quality metrics, e.g., mutual information, gradient entropy, total variation, etc. [12–15]. Retrospective motion correction techniques can be incorporated into a wide range of MRI protocols and have the advantages of additional hardware/sequence independence and have minimal impact on the study setup. Motions during multi-shot imaging produce inconsistencies for different *k*-space lines due to the change of excitation volume/plane and its orientation. Theoretically, both rotation and translation have their counterparts in *k*-space according to the Fourier theory [22,23]. However, this relationship does not hold in some practical situations when the B_0/B_1 variations and spin history effects exist. In this sense, inversion of the *k*-space data according to the motion parameters is not a good solution for motion correction. An alternative approach is to reconstruct the motion-free image by solving an optimization problem with consideration of a forward motion model. In this way, both data fidelity and a priori knowledge of the motion parameters can be utilized efficiently. Other studies applied the sparse representation theory to estimate motion parameters jointly with image reconstruction by taking advantage of the redundancy in *k*-space data [24]. Parallel-imaging can also be applied to correct motion artifacts using the coil sensitivity information as well as the multi-channel data [25].

Most of the approaches mentioned above rely on the assumption that no out-of-FOV happened during motions. However, this assumption does not hold when out-of-FOV motion (e.g., large-angle rotation or through-plane translation) happens. For partial brain imaging, object motion can cause varying RF excitation volumes, which becomes problematic for the correction of any motion-corrupted data in *k*-space. This problem is more severe in 2D imaging compared to 3D imaging because the acquired signal no longer comes from intended imaging slices when out-of-FOV motion happens. An alternative approach is to apply optimization algorithms to the motion model and reconstruct an image with larger FOV containing all the sub-FOVs during motion. Unlike whole brain imaging, the out-of-FOV motion problem is challenging due to a lack of image information in the out-of-FOV regions. Even in the cases when motion parameters are known, the problem of motion correction is still ill-conditioned. In general, currently available retrospective algorithms are not able to correct for image artifacts introduced by out-of-FOV motion.

Recently, deep learning has been introduced for MRI reconstruction and has achieved good performances [26–33]. By using the convolutional neural network (CNN), a priori information extracted from the previous database can be applied to improve the performance of image reconstruction. Moreover, a few groups have attempted to use CNN to solve the in-plane motion problems in MRI [34–36]. Nevertheless, out-of-FOV motion problems have not been properly addressed yet.

In this study, we propose a CNN-based motion correction algorithm to solve the out-of-FOV motion problem. This approach is designed to train a neural network with the capability of learning the relationship between motion-corrupted image and motion-free image. Unlike other studies, we developed a neural network with the introduction of motion parameters into the loss function to reduce the complexity of the problem. Instead of using CNN output as a final result, we fed the data fidelity reconstruction with the prior generated by the pre-trained CNN. The primary purpose of this study was to demonstrate the feasibility of using CNN to compensate for out-of-FOV motion artifacts using the acquired motion parameters from tracking cameras [5,7], navigators [8–11], or directly generated from images using AF methods [16–19].

2. Theory

An overview of the CNN-based motion correction algorithm is presented in Fig. 1. The following subsections describe details of the problem formulation and implementation.

To simplify the problem formulation, we assume that all *k*-space data are densely sampled in Cartesian coordinates and noise are prewhitened. Taking 3D multi-shot imaging as an example, a rigid-body motion correction problem can be generically formulated as:

$$\hat{x} = \arg\min_x \ \|M\mathscr{F}Tx - y\|_2^2, \tag{1}$$

where $y \in \mathbb{C}^N$ denotes the measured *k*-space data, $x \in \mathbb{C}^N$ the image to be reconstructed, $N = n_x \cdot n_y \cdot n_z$, in which n_x, n_y, n_z are voxel sizes along *x*, *y* and *z* axis); $T \in \mathbb{R}^{N\times N}$ the rigid body motion transformation matrix characterized by three translational motion parameters $(x,y,z) \in \mathbb{R}^{3N}$ and three rotational motion parameters $(\rho_x,\rho_y,\rho_z) \in [-\pi\sim\pi]^{3N}$; $\mathscr{F} \in \mathbb{C}^{N\times N}$ the discrete Fourier transform (DFT) matrix, and $M \in [0,1]^N$ the diagonal masking matrix in *k*-space.

Commonly, we partitioned full k-space into N_s segments with each including L acquired readouts. We assume that the subject motion is piecewise constant within each segment. For more general line-by-line settings, N_s could be equal to the phase encoding steps ($n_{ky} \cdot n_{kz}$). In this study, we suppose the segment number to be $N_s = 12 \times 8$.

This motion model is widely used to estimate motion-free image x with given T and M [37]. However, it only models the subject motion as coordination changes. This assumption does not apply for some situation when outside spin signals move into the FOV and contribute to the acquired data. In this case, the optimization problem should be re-written as:

$$\hat{x} = \arg\min_x \ \|M\mathcal{F}\Omega_0 T(x + \tilde{u}) - y\|_2^2, \tag{2}$$

where $\Omega_0 \in [0,1]^{N\times N}$ denotes the imaging FOV, $\tilde{u}$ the outside image moving into the FOV due to the motion. Suppose we have N_s k-space segments, $\tilde{u}$ should be a union set of $\tilde{u}_s$, in which $\tilde{u}_s$ represents the s-th out-of-FOV images. Note that when $\tilde{u}$ is noise, problem [2] is equals to [1] since the additional signal coming from the outside regions makes no contribution to imaging. This is exactly the case when whole brain imaging is performed. However, in partial brain imaging, signals out of the FOV should be taken into consideration.

Because of the randomness of motion in each segment, $\tilde{u}_s$ is always partially sampled in k-space. Directly solving x and $\tilde{u}$ is not feasible due to the ill-condition nature of this problem. A straightforward idea to improve the condition number of this problem is to introduce some regularization terms. One commonly used regularization term is total variation (TV), which is defined as a finite differences approximation of image gradient. Another class of regularizers is the transform domain sparsity (e.g., wavelet transform). However, these terms only improve the condition number of the problem by adding constraints to x, but can not provide useful information for $\tilde{u}$.

As mentioned above, conventional motion correction algorithms are not able to handle the problems of out-of-FOV motion due to a lack of information about $\tilde{u}$. Deep learning techniques, on the other hand, have the potential to capture the missing information by extracting a priori knowledge from a large dataset. Inspired by the successful applications of learning-based approaches in solving image reconstruction and blind deconvolution problems, we propose a neural network with the capability to remove rigid-body motion artifacts with given motion parameters.

Next we will introduce our idea of applying a priori knowledge from CNN into the optimization problem [2]. Let us define $\Omega_s = \Omega_0 \cdot \mathrm{T}_s - \Omega_0$, which represents the real FOV apart from the original FOV during the s-th k-space segment acquisition. We define $\widetilde{\Omega}$ as:

$$\widetilde{\Omega} = \bigcup_{s=0}^{N_s} \Omega_s, \tag{3}$$

The loss function of CNN is defined as an ℓ_2-norm of the output and ground truth.

$$Loss = \frac{1}{N_T}\sum_{i=1}^{N_T} \|\widetilde{\Omega} f_{cnn}(z_i; w) - \tilde{x}_i\|_2^2 + \lambda_w \ \|w\|_1, \tag{4}$$

where $\tilde{x}_i$ denotes the i-th ground truth image with a union of FOV $\widetilde{\Omega}$, $f_{cnn}(z;w)$ the CNN output, z_i a zero-padding of $\mathcal{F}y_i$ to match the size of $\tilde{x}_i$, w the neural network parameters, N_s the number of k-space segments, N_T the amount of training samples, and $\|w\|_1$ a regularization term to constrain the sparsity of network parameters.

Once we have learned the hidden parameters (w), we can directly apply the CNN prediction to the optimization problem [2] and replace $\tilde{u}$ with $\widetilde{\Omega} f_{cnn}(z;w)$:

$$\hat{x} = \arg\min_x \ \|M\mathcal{F}\Omega_0 T(x + \widetilde{\Omega} f_{cnn}(z;w)) - y\|_2^2, \tag{5}$$

By exploiting the CNN output into the optimization problem, we now have excellent a priori knowledge about $\tilde{u}$, thus makes the problem much easier to be solved. To further alleviate the ill-conditionness of the problem, we introduce another regularization term into [5]:

$$\hat{x} = \arg\min_x \ \|M\mathcal{F}\Omega_0 T(x + \widetilde{\Omega} f_{cnn}(z;w)) - y\|_2^2 + \lambda \ \|x - \Omega_0 f_{cnn}(z;w)\|_2^2, \tag{6}$$

The first term ensures data fidelity while the second term serves as a regularizer. λ is used to balance the trade-off between data fidelity and the regularization term. Other regularization terms can also be added into this formulation to further improve the performance of optimization, e.g., wavelet transform sparsity, TV, etc.

If we denote $M\mathcal{F}\Omega T$ as A, then the optimization problem [6] admits a closed form solution as:

$$\hat{x} = (A^H A + \lambda I)^{-1}(A^H y - A^H A\widetilde{\Omega} f_{cnn}(z;w) + \lambda \bullet \Omega_0 f_{cnn}(z;w)) \tag{7}$$

which can be addressed in an iterative fashion using conjugate gradient (CG) algorithm. In addition, the CNN prediction $\Omega_0 f_{cnn}(z;w)$ can serve as an initial guess for the iterations.

3. Methods

3.1. Data acquisition

Evaluation of the proposed algorithm was performed on an open database consisting of 1113 healthy subjects from the WU-Minn Human Connectome Project (HCP) [38,39]. 3D whole-brain magnetization-prepared rapid-acquisition with gradient-echo (MPRAGE) data were acquired on a 3.0 Tesla MRI scanner (MAGNETOM Skyra, Siemens, Erlangen, Germany) with TR/TE = 2400/2.14 ms, FOV = 224 × 224 × 192 mm^3, voxel size = 0.7 × 0.7 × 0.7 mm^3, and flip angle = 8°. Motion-corrupted images were generated from the motion-free images for training and validation (see details in the Motion Simulation section).

In vivo brain images were acquired on four subjects as approved by the local ethics board. One of the subject was repeatedly scanned in three different positions in order to include some major physical effects during practical motion corrections, e.g., variations in the B_0 and B_1 fields. The subject was asked to perform two types of head motions: [1] «side-to-side» motion and [2] «nodding» motion. All the scans were performed on a 3.0 Tesla MR system (Achieva, Philips Healthcare, Best, The Netherlands), with a maximum amplitude of 40 mT/m and a slew rate of 150 mT/m/ms. 3D MPRAGE imaging was performed with identical scan parameters as the HCP data.

3.2. Image normalization

In order to reduce the spatial variation among subjects, we applied geometrical normalization to all the training/validation data using a diffeomorphic registration scheme [40]. Generally speaking, this deformable model first calculated a “flow field” map between a template and individual images according to their mutual information, and then applied the “flow field” map to each individual images. By using this model, geometrical variation between images can be significantly reduced, thus improving both training and validation performances. We implemented geometrical normalization process using the SPM DARTEL toolbox.

3.3. Motion simulation

We generated motion-corrupted images according to the forward motion model $\hat{x} = M\mathcal{F}\Omega T x$. We assumed the phase encoding order to be center-out. For 2D imaging, phase encoding was simulated along the anterior-posterior direction (x-axis), and frequency encoding was selected along the superior-inferior direction (z-axis). The full k-space was partitioned into 8 segments along the x-axis. For 3D imaging, an extra phase encoding was performed along the left-right direction (y-axis). Correspondingly, the full k-space was partitioned into 12 segments

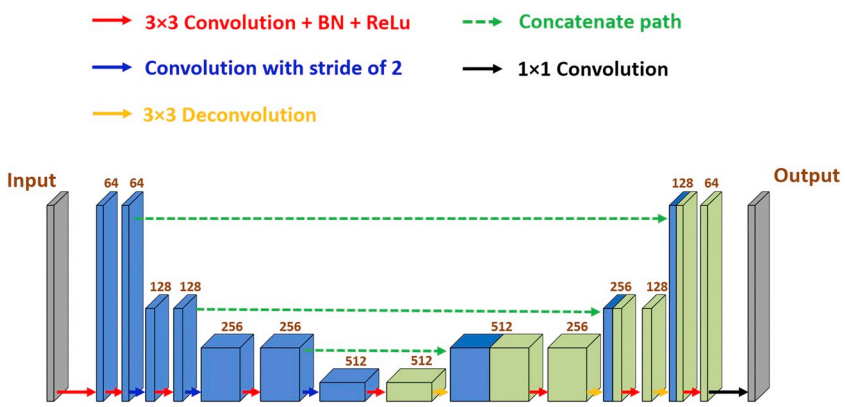

Fig. 2. Architecture of 2D U-net. The structure is modified by replacing max pooling layer with a convolution layer with a stride of 2, and adding batch normalization layer after each convolution for faster convergence.

along the y-axis and 8 segments along the x-axis. Motion parameters were simulated in each *k*-space segments relative to the center *k*-space line.

Two types of motion trajectories (oscillating and sudden motion) were simulated for both 2D and 3D images (Supporting Fig. S1-S4). Spline interpolation was performed after rigid body motion transformation. To evaluate the motion correction performances on different motion patterns, we generated simulated motion-corrupted data with translation amplitudes in a range of 1 to 10 pixels (0.07–7.0 mm) and rotation angles in a range of 2° to 20°. We calculated the percentage of the out-of-FOV image relative to the brain regions within FOV (blue regions in Fig. 1).

3.4. CNN architecture

A U-net architecture [41] was used in this study, which was initially designed for image segmentation but has recently been applied to image reconstructions. Both 2D and 3D U-net were implemented using the TensorLayer library. Fig. 2 illustrates the modified 2D U-net architecture. We modified the U-net by replacing max pooling with a convolution layer with a stride of 2. Batch normalization (BN) layer was used after each convolution for faster convergence [42]. Each layer contains a 3 × 3 convolution kernel and a rectified linear unit (ReLu) activation function. A 3 × 3 convolution layer with a stride of 2 was connected immediately after the first convolution layer. In the synthesis path, each layer consists of a 3 × 3 upconvolution kernel with a stride of 2, followed by another 3 × 3 convolution layer with ReLu activation function. Concatenation connections were applied to provide essential high-resolution features to the synthesis path and to improve the gradient behavior for back-propagation. The similar architecture was carried out for 3D U-net with an additional dimension of convolution kernel with a size of 3.

Simulated motion-corrupted images and their corresponding motion-free images were used as input and output of the network. For each group of simulation, data were randomly grouped into two subsets, i.e., training set (70%) and validation set (30%). Loss function was defined as the ℓ_2-norm of ground truth and CNN output. Adam algorithm [43] was used for gradient descent with the learning rate of 10^{-4} and β of 0.9.

Training was performed on NVIDIA Tesla P100 × 4 GPU (each with 16GB memory). To speed up the training process, we implemented a multi-GPU parallel computational algorithm by dividing the model's inputs into multiple sub-batches and applying a sub-model in each sub-batch. The results were concatenated into a combined batch after each training iteration. For 2D U-net, batch size was selected to be 100. A total of 1000 iterations was applied for training. For 3D U-net, a batch size of 8 and total iterations of 100 were applied. It takes approximately 40 min for 2D training, and 5 h for 3D training.

3.5. Image analysis

After CNN training, CG algorithm was applied to solve Eq. [7] with the CNN output image as a priori knowledge. λ was selected from 0.1 to 3.0 with an increment of 0.1. Maximal iteration of 20 and tolerance error of 10^{-6} was used in the CG optimizer. On average, it takes approximately 30 s for 2D image reconstruction and 5 min for 3D image reconstruction on CPU (Intel Xeon Gold 6130 @ 2.10GHz).

The performance of the proposed motion correction algorithm was assessed in terms of normalized mean square error (NMSE), peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) [44]. We compared the CNN prediction results with and without geometrical normalization. In addition, the robustness of our proposed algorithm was evaluated by adding different levels of Gaussian noise to the given motion parameters. We also compared the performance of the proposed approach with regard to a gradient-based AF algorithm [17].

Paired student *t*-test was conducted to compare the NMSE, PSNR and SSIM measurements between different approaches. A *P* value of less than 0.01 was considered statistically significant.

4. Results

As shown in Supporting Video S1, the spatial variations between subjects were obviously reduced by applying geometrical normalization. Both training and validation losses converged faster when using normalized data (Fig. 3). Taking oscillating motion as an example, the ℓ_2-errors were significantly lower when the network was trained with normalized images (2D: training loss 9.1×10^{-4}, validation loss 1.2×10^{-3}; 3D: training loss 8.1×10^{-4}, validation loss 1.1×10^{-3}) compared to that with original images (2D: training loss: 1.2×10^{-3}, validation loss: 3.2×10^{-3}; 3D: training loss 1.2×10^{-3}, validation loss 1.8×10^{-3}).

Intermediate training results during different epochs are shown in Fig. 4 and Supporting Video S2. The motion-corrected images corresponding to different motion parameters were reconstructed by CNN during iterations. The pre-trained neural network performed well in removing the motion artifacts. However, the reconstructed images seem to be blurring compared to the ground truth (Fig. 4). With the application of data fidelity term o the CNN output, image quality was improved in both 2D and 3D images (Fig. 5-7). As shown in Fig. 8, a combination of CNN prediction and data fidelity performs the best in terms of PSNR with $\lambda = 1.5$ for 2D image reconstruction (NMSE = 7.2×10^{-3}, PSNR = 29.8, and SSIM = 0.91), and $\lambda = 2.1$ for 3D reconstruction (NMSE = 6.8×10^{-3}, PSNR = 35.2, and SSIM = 0.92).

Comparison of motion correction performances of the CNN-based algorithm and AF-based algorithm is shown in Fig. 9. The proposed algorithm outperformed AF-based algorithm for both 2D (NMSE:

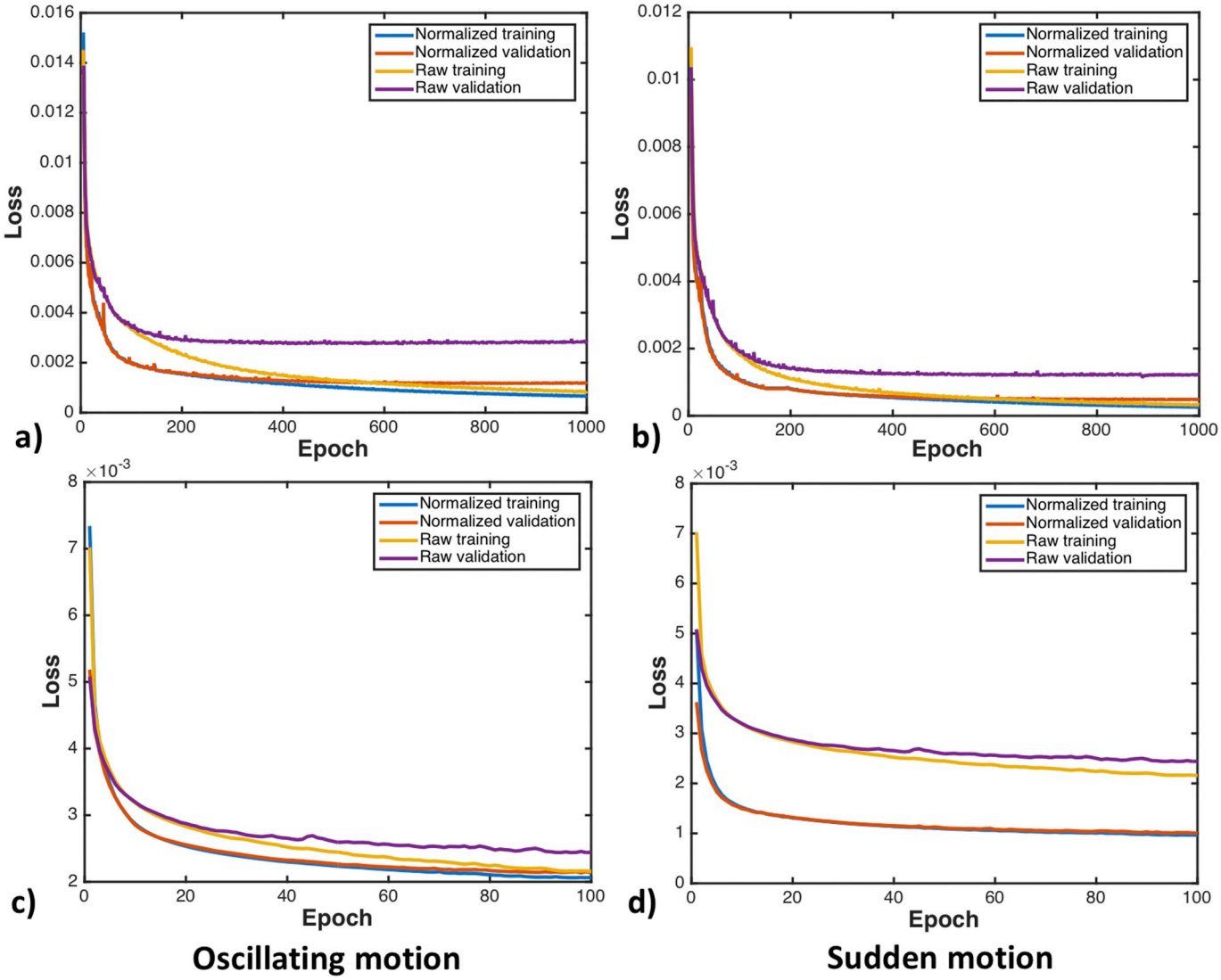

Fig. 3. CNN training and validation loss curves. a,b) 2D U-net training with oscillating motion and sudden motion corrupted images; c,d) 3D U-net training with oscillating motion and sudden motion corrupted images.

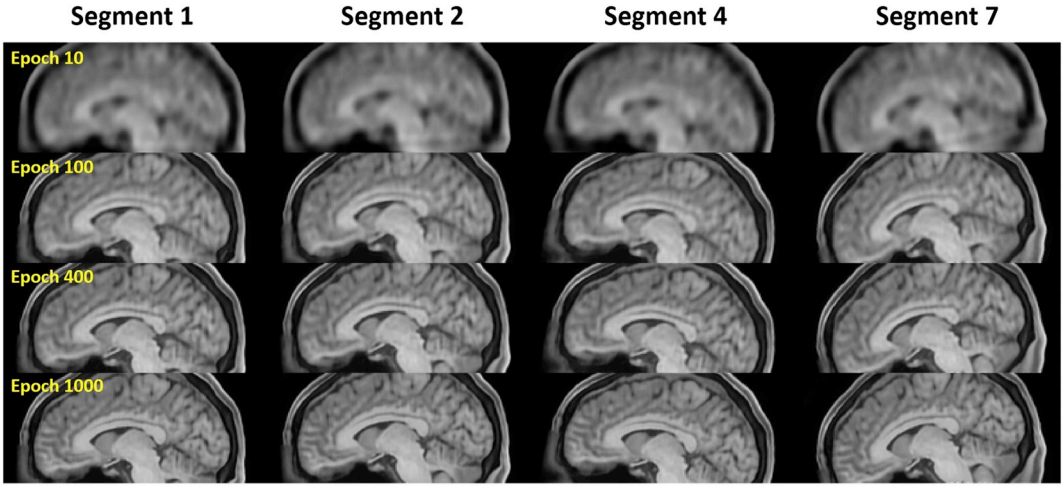

Fig. 4. Intermediate results during CNN training. FOV is moved during data acquisition for each *k*-space segment due to motion.

0.0066 ± 0.0009 vs 0.041 ± 0.008, $P < .01$; PSNR: 29.60 ± 0.74 vs 21.71 ± 0.27, $P < .01$; SSIM: 0.89 ± 0.014 vs 0.73 ± 0.004, $P < .01$) and 3D imaging (NMSE: 0.0067 ± 0.0008 vs 0.070 ± 0.021, $P < .01$; PSNR: 32.40 ± 1.63 vs 22.32 ± 2.78, $P < .01$; SSIM: 0.89 ± 0.01 vs 0.62 ± 0.03, $P < .01$). Two examples of 3D motion correction with all slices are shown in **Supporting Video**

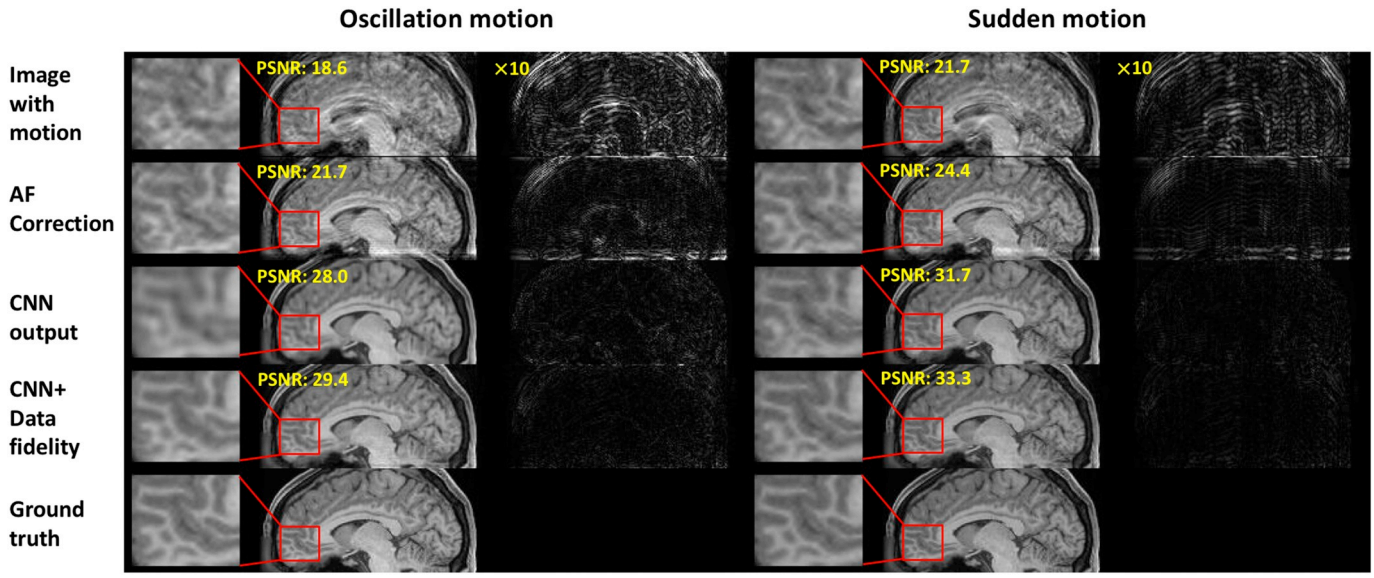

Fig. 5. Comparison of motion correction performances for 2D imaging with oscillating and sudden motion trajectories.

4–5. As shown in Supporting Fig. S5-S6, the proposed algorithm performs well for the motion corrupted data simulated from interleaved *k*-space data with repeated scans in different orientations. The image qualities of the motion-corrected images were acceptable with PSNR of 30.3 in 2D imaging and 40.4 for 3D imaging. Although the improvement was less significant in 3D imaging compared to 2D imaging according to the image quality metrics, data fidelity operation is still necessary to make the reconstruction results less biased by the training dataset.

The proposed algorithm has better tolerance to motion estimation errors compared to the AF-based algorithm. Although image quality reduced as the estimation errors increase, it was still acceptable when the SNR decreased to 1 dB (Fig. 10). Although direct comparison of the motion correction performances between the 2D and 3D image is not feasible due to different acquisition schemes, the 3D images seem to be more robust to the motion parameter errors according to all the image quality metrics.

On average, when more than 5% brain images were missing, the CNN-based algorithm performed better than conventional algorithms. More accurate motion correction performance was achieved when motion occurs at the peripheral *k*-space lines. When severe motion happens near the center of *k*-space, the performance of conventional correction algorithm becomes worse, while the CNN-based algorithm is more robust with 20% data missed for both 2D and 3D imaging (NMSE = 0.0087 for 2D imaging and NMSE = 0.0083 for 3D imaging).

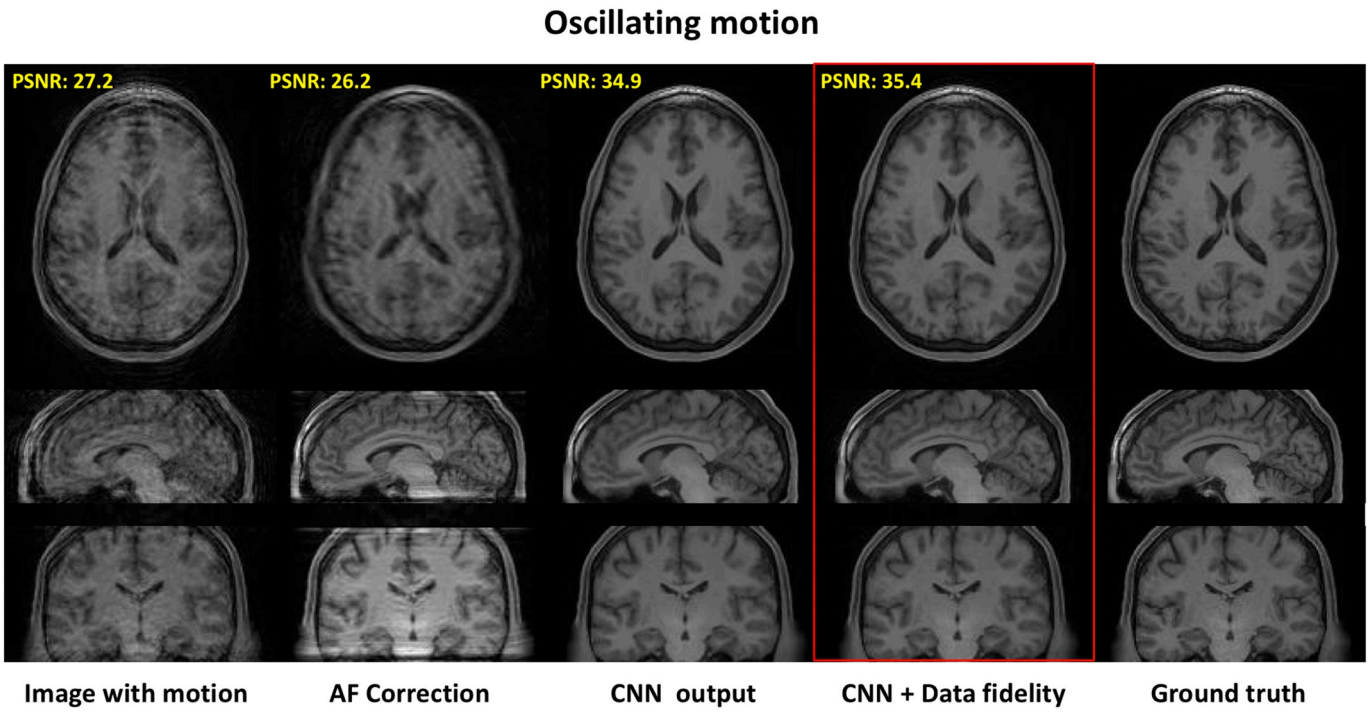

Fig. 6. Comparison of motion correction performances for 3D imaging with oscillating motion trajectory.

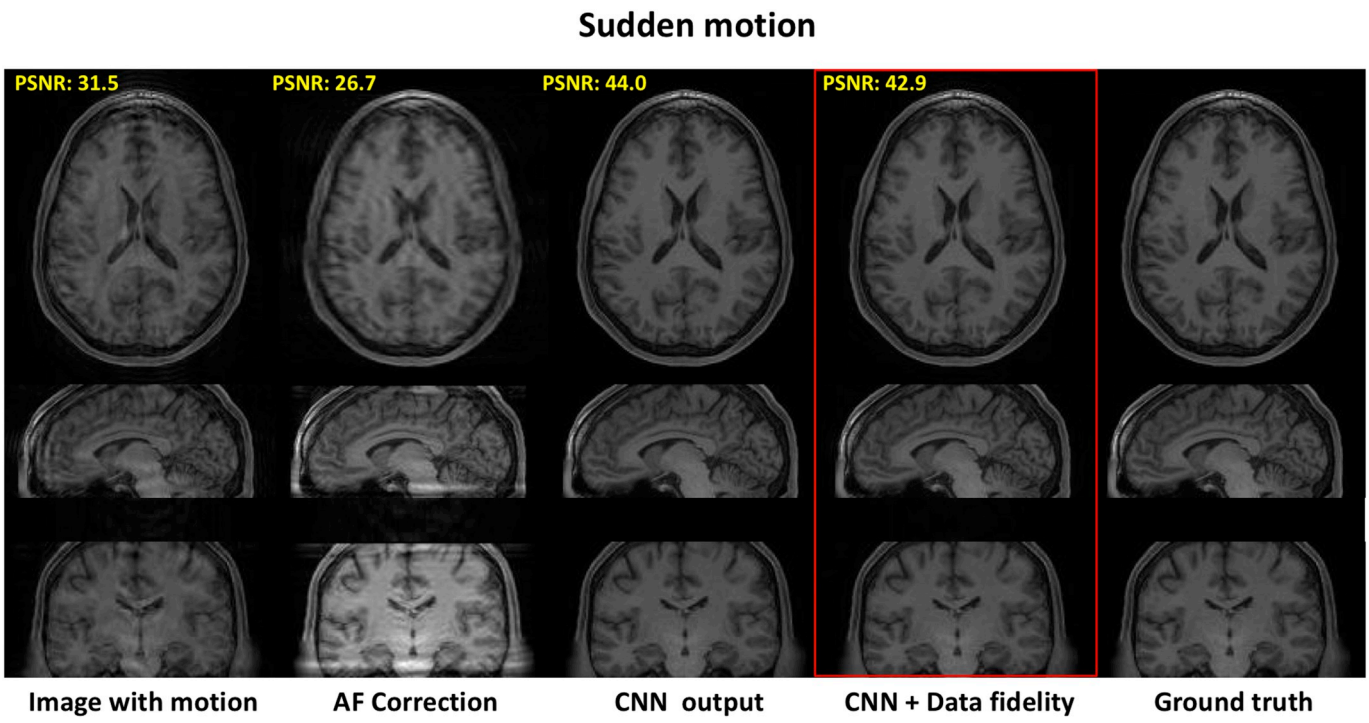

Fig. 7. Comparison of motion correction performances for 3D imaging with sudden motion trajectory.

5. Discussion

This study presented a new motion correction algorithm with the incorporation of deep learning technique for both 2D and 3D imaging. An essential feature of this algorithm is the application of CNN-based prior image for solving the out-of-FOV motion problem. We trained the network with given motion parameters, which significantly reduced the complexity of the problem and improved the training performance. Results show that our algorithm outperformed the conventional AF-based algorithm with significantly lower NMSE and higher SSIM.

Unlike other studies [1–20], our work mainly focused on the problem of out-of-FOV motion during multi-shot imaging. Most existing techniques were only able to recover motion artifacts composed of translation and small angle rotation (commonly < 5°), we characterized

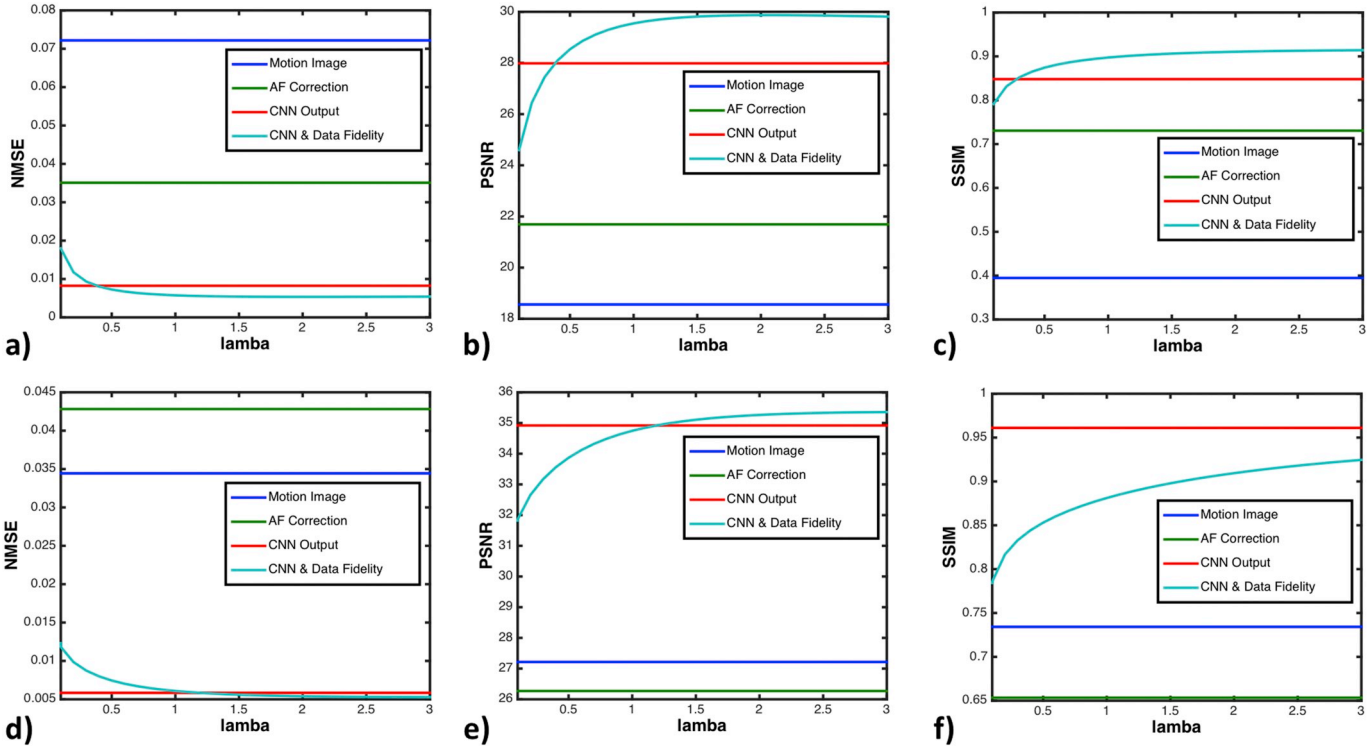

Fig. 8. Image reconstruction performances with the selection of different λ for the proposed algorithm. a) NMSE, b) PSNR and c) SSIM of the 2D motion correction; d) NMSE, e) PSNR and f) SSIM of 3D motion correction.

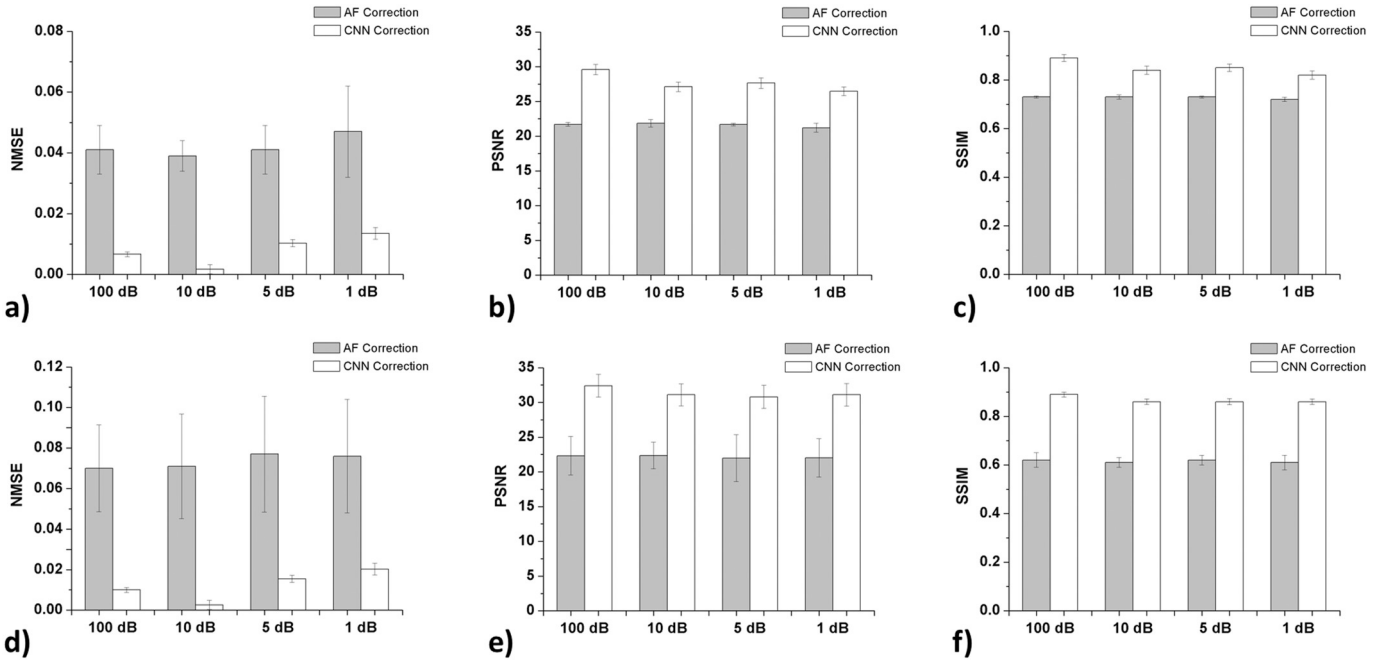

Fig. 9. Comparison of motion correction performances of the proposed CNN-based algorithm with AF-based algorithm in testing dataset with metrics of a) NMSE, b) PSNR and c) SSIM for the 2D imaging, and d) NMSE, e) PSNR and f) SSIM for 3D imaging.

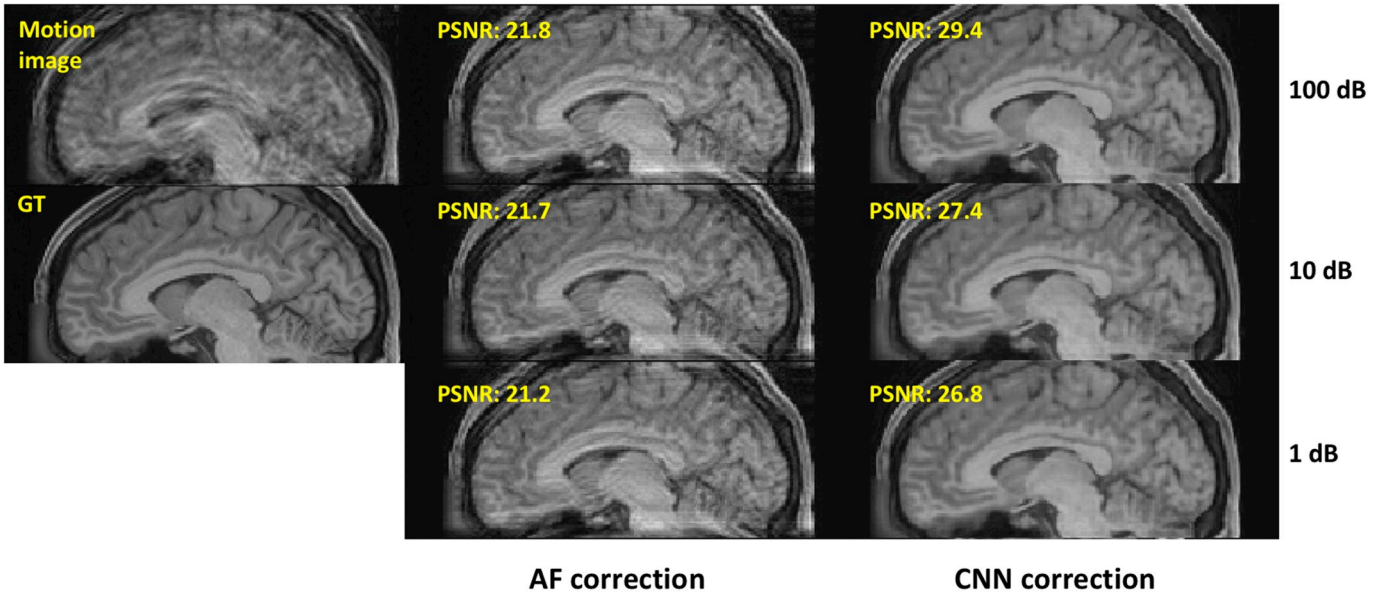

Fig. 10. Motion parameter errors on the performance of CNN-based algorithm and AF-based algorithm.

the great potential of incorporating a priori knowledge into the motion problem and solved cases with even 20% data missed during scans. The reason that our approach performed better than conventional algorithms is quite reasonable since missing data makes the motion problem ill-conditioned. This study suggests that CNN is capable of providing good image priors with motion artifacts removed and thus make the motion model well-conditioned. In addition, the proposed algorithm does not rely on other regularization terms such as transform domain sparsity or TV. According to our experimental result from the interleaved motion-corrupted data, the quality of the motion-corrected image is acceptable with the existence of several major physical effects, i.e., B_0/B_1 inhomogeneity.

We tested the robustness the proposed algorithm by adding random noise to the motion parameters. According to all image quality metrics, the proposed algorithm has good tolerance to motion estimation errors when SNR decreased to 1 dB. Based on our results, 2D and 3D motion correction algorithms show comparable performances. Although direct comparison is not feasible because even the same motion parameters will result in different motion artifacts due to the different *k*-space data acquisition schemes, 3D CNN appears to be more robust to estimation errors in motion parameters.

One should also note that the neural network only provides us the most probable prediction of the image based on the prior datasets. It is still necessary to incorporate data fidelity term into the model to further minimize the reconstruction errors. Results show that with a combination of CNN output and data fidelity term, our algorithm performs much better in terms of visual observation as well as all image quality metrics.

The proposed CNN model is based on the U-net architecture, which has gained great popularity in image reconstruction due to its strong representation power and minimal training parameters required [41]. Previous studies have shown that the multi-scale network can learn global features more effectively, and requires much fewer parameters compared to single-scale networks [45]. To better preserve the image resolution, max pooling layers are removed in the modified network. Instead, additional convolution layers with a stride of two are used to compress the feature maps. Besides, batch normalization layers are applied to accelerate the training convergence.

As a part of pre-processing, geometrical normalization helps in reducing the spatial variation of training dataset, which decreases the complexity of the problem. The results have shown that the data normalization helps our neural network perform better in both simulated data and scanned data. Generally speaking, normalization reduces the CNN training burden and leads to faster convergence. Since it is not practical to conduct normalization directly on the motion-corrupted image, a reference image (e.g., T_1w or T_2w image) can be used instead to calculate the deformable field map. And apply the normalization process to all the training data using the estimated deformable field map. In this way, all the training images can be normalized to the scanned data, which makes the pre-trained network directly applicable for practical usage.

Unlike many other applications of neural networks, our network was trained without data augmentation (random deformation, flipping, translation or rotation, etc.) or random patching. The reason is that we hope our neural network to learn the global motion artifact patterns, instead of learning a deblurring network. Since we want the training process to benefit the most from geometrical normalization, cropping the brain images into pieces will destroy the original distribution of the motion pattern. Moreover, since our network was trained with given motion parameters, the complexity of the problem was significantly reduced therefore data size was not an issue in this case.

Similar to other studies, we assumed that the motion time-scale is longer than the repetition time TR, and motion trajectory can be modeled as a piecewise constant function in each *k*-space segment. Under these assumptions, our approach can correct for arbitrary motion trajectories in six degrees of freedom. This algorithm can be easily extended to line-by-line *k*-space acquisition since the assumption of *k*-space segment is not necessary for either training or optimization processes. Also, the model can be extended to non-Cartesian acquisitions, e.g., radial or spiral trajectories. Since the training dataset is built according to the motion trajectories, thus the model is flexible enough to learn any kind the motion degradation and recover the motion-free image as long as we can get the motion parameters. In addition, an extension of the proposed method to parallel imaging is also straightforward. The only thing we need to do is to change the data fidelity term to fit multi-coil data.

An important application of the proposed algorithm is fetal neuroimaging because the movements are more “uncontrolled” for fetus. Some groups have applied retrospective motion correction techniques to recover through-plane motion with either a slice profile filter [25,46] or registration-based approaches [47,48]. It is quite straightforward to apply the CNN-based prior images into the existing retrospective motion correction techniques to solve the common ill-conditioning problem. A combination of snapshot multi-slice imaging with CNN prediction seems to be promising for the correction of through-plane motion artifacts in fetal neuroimaging.

There exist several limitations in our study. First, we only tested our algorithm on simulated motion images. Since the purpose of this study is to demonstrate the advantage of introducing deep learning into the motion problem and prove the added value of a priori image, we did all the testing with given motion parameters to avoid motion estimation errors. To take into consideration of some major physical effects that happens during motions, e.g., variations in B_0 and B_1 fields, we performed an additional experiment with repeated scans of the same subject in several different orientations. According to our experimental results, good motion correction performance was achieved regardless of the changes of physical effects. When putting into a practical situation, motion parameters can be easily obtained from navigators [8–11], external sensors [5–7] or autofocusing algorithms [12–16]. Another limitation is that we did not take into consideration the spin history, B_1-field variation and B_0-field inhomogeneity variation during the scan, which has been included in our future study. Another issue that we should take into account is the computational time. With the application of parallel computing, the training time of our CNN is shortened to 40 min for 2D and 5 h for 3D images. There are still many ways to further improve the training efficiency. For example, we can pre-train a neural network with random motion artifacts and then fine-tune these parameters to a specific motion model.

6. Conclusion

This study demonstrated the feasibility of applying deep learning technique to solve the out-of-FOV motion problem in both 2D and 3D imaging. With the incorporation of prior image into a motion model, the proposed algorithm is capable of correcting out-of-FOV motion artifacts with up to 20% data missed. Loss function of the neural network was defined with consideration of motion parameters into it to reduce the complexity of the problem. Significant improvement of image quality was observed by using CNN-based algorithm compared to AF-based algorithm.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mri.2020.05.004.

References

[1] Glover GH, Pauly JM. Projection reconstruction techniques for reduction of motion effects in MRI. Magn Reson Med 1992;28:275–89.
[2] Ooi MB, Krueger S, Thomas WJ, Swaminathan SV, Brown TR. Prospective real-time correction for arbitrary head motion using active markers. Magn Reson Med 2009;62:943–54.
[3] Tisdall MD, Hess AT, Reuter M, Meintjes EM, Fischl B, van der Kouwe AJ. Volumetric navigators for prospective motion correction and selective reacquisition in neuroanatomical MRI. Magn Reson Med 2012;68:389–99.
[4] White N, Roddey C, Shankaranarayanan A, Han E, Rettmann D, Santos J, et al. PROMO: real-time prospective motion correction in MRI using image-based tracking. Magn Reson Med 2010;63:91–105.
[5] Zaitsev M, Dold C, Sakas G, Hennig J, Speck O. Magnetic resonance imaging of freely moving objects: prospective real-time motion correction using an external optical motion tracking system. Neuroimage 2006;31:1038–50.
[6] Qin L, van Gelderen P, Derbyshire JA, Jin F, Lee J, de Zwart JA, et al. Prospective head-movement correction for high-resolution MRI using an in-bore optical tracking system. Magn Reson Med 2009;62:924–34.
[7] Marxen M, Marmurek J, Baker N, Graham SJ. Correcting magnetic resonance k-space data for in-plane motion using an optical position tracking system. Med Phys 2009;36:5580–5.
[8] Gallichan D, Marques JP, Gruetter R. Retrospective correction of involuntary microscopic head movement using highly accelerated fat image navigators (3D FatNavs) at 7T. Magn Reson Med 2016;75:1030–9.
[9] Ehman RL, Felmlee JP. Adaptive technique for high-definition MR imaging of moving structures. Radiology 1989;173:255–63.
[10] Fu Z, Wang Y, Grimm R, Rossman P, Felmlee J, Riederer S, et al. Orbital navigator echoes for motion measurements in magnetic resonance imaging. Magn Reson Med 1995;34:746–53.
[11] Welch EB, Manduca A, Grimm RC, Ward HA, Jack CR. Spherical navigator echoes for full 3-D rigid body motion measurement in MRI. Med Reson Med 2002;47:32–41.
[12] Manduca A, McGee K, Welch E, Felmlee J, Grimm R, Ehman R. Auto-correction in MR imaging: adaptive motion correction without navigator echoes. Radiology 2000;215:904–9.
[13] Atkinson D, Hill DL, Stoyle PN, Summers PE, Keevil SF. Automatic correction of motion artifacts in magnetic resonance images using an entropy focus criterion. IEEE Trans Med Imaging 1997;16:903–10.
[14] Atkinson D, Hill DL, Stoyle PN, Summers PE, Clare S, Bowtell R, et al. Automatic compensation of motion artifacts in MRI. Magn Reson Med 1999;41:163–70.
[15] Lin W, Ladinsky G, Wehrli F, Song HK. Image metric-based correction (auto-focusing) of motion artifacts in high-resolution trabecular bone imaging. J Magn Reson Imaging 2007;26:191–7.
[16] Haskell MW, Cauley SF, Wald LL. TArgeted motion estimation and reduction (TAMER): data consistency based motion mitigation for MRI using a reduced model joint optimization. IEEE Trans Med Imaging 2018;37:1253–65.

[17] Loktyushin A, Nickisch H, Pohmann R, Schölkopf B. Blind retrospective motion correction of MR images. Magn Reson Med 2013;70:1608–18.
[18] McGee KP, Manduca A, Felmlee JP, Riederer SJ, Ehman RL. Image metric-based correction (autocorrection) of motion effects: analysis of image metrics. J Magn Reson Imaging 2000;11:174–81.
[19] Cheng JY, Alley MT, Cunningham CH, Vasanawala SS, Pauly JM, Lustig M. Nonrigid motion correction in 3D using autofocusing with localized linear translations. Magn Reson Med 2012;68:1785–97.
[20] Maclaren J, Herbst M, Speck O, Zaitsev M. Prospective motion correction in brain imaging: a review. Magn Reson Med 2013;69:621–36.
[21] Zaitsev M, Maclaren J, Herbst M. Motion artifacts in MRI: a complex problem with many partial solutions. J Magn Reson Imaging 2015;42:887–901.
[22] Bracewell RN. The Fourier transform and its applications. New York: McGraw Hill; 1999.
[23] Atkinson D, Hill DLG. Reconstruction after rotational motion. Magn Reson Med 2003;49:183–7.
[24] Yang Z, Zhang C, Xie L. Sparse MRI for motion correction. IEEE 10th International Symposium on Biomedical Imaging. 2013. p. 962–5.
[25] Cordero-Grande L, Hughes EJ, Hutter J, Price AN, Hajnal JV. Three-dimensional motion corrected sensitivity encoding reconstruction for multi-shot multi-slice MRI: application to neonatal brain imaging. Magn Reson Med 2018;79:1365–76.
[26] Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain transform manifold learning. Nature 2018;555:487–92.
[27] Wang S, Ke Z, Cheng H, Jia S, Leslie Y, Zheng L, et al. DIMENSION: Dynamic MR Imaging with Both K-space and Spatial Prior Knowledge Obtained via Multi-Supervised Network Training. arXiv:1810.00302; 2018.
[28] Hammernik K, Klatzer T, Kobler E, Recht MP, Sodickson DK, Pock T, et al. Learning a Variational network for reconstruction of accelerated MRI data. Magn Reson Med 2018;79:3055–71.
[29] Bilgic B, Chatnuntawech I, Manhard MK, Tian Q, Liao C, Cauley SF, et al. Highly Accelerated Multishot EPI through Synergistic Combination of Machine Learning and Joint Reconstruction. arXiv:1808.02814; 2018.
[30] Chiang S, Balasingham S, Richmond L, Curpen B, Skarpathiotakis M, Martel A. Motion Corruption Detection in Breast DCE-MRI. Cham: International Workshop on Machine Learning in Medical Imaging. Springer; 2017. p. 10–8.
[31] Han Y, Yoo J, Kim HH, Shin HJ, Sung K, Ye JC. Deep learning with domain adaptation for accelerated projection-reconstruction MR. Magn Reson Med 2018;80:1189–205.
[32] Gurbani SS, Schreibmann E, Maudsley AA, Cordova JS, Soher BJ, Poptani H, et al. A convolutional neural network to filter artifacts in spectroscopic MRI. Magn Reson Med 2018;80:1765–75.
[33] Akçakaya M, Moeller S, Weingärtner S, Uğurbil K. Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging. Magn Reson Med 2018;81:439–53.
[34] Duffy BA, Zhang W, Tang H, Zhao L, Law M, Toga AW, et al. Retrospective correction of motion artifact affected structural MRI images using deep learning of simulated motion. 1st Conference on Medical Imaging with Deep Learning, Amsterdam, The Netherlands. 2018.
[35] Pawar K, Chen Z, Shah NJ, Egan GF. MoCoNet: Motion Correction in 3D MPRAGE images using a Convolutional Neural Network approach. arXiv:1807.10831; 2018.
[36] Sommer Karsten, Brosch Tom, Wiemker Rafael, Harder Tim, Saalbach Axel, Hall Christopher S, et al. Correction of motion artifacts using a multi-resolution fully convolutional neural network. Joint Annual Meeting ISMRM-ESMRMB. 2018.
[37] Batchelor PG, Atkinson D, Irarrazaval P, Hill DL, Hajnal J, Larkman D. Matrix description of general motion correction applied to multishot images. Magn Reson Med 2005;54:1273–80.
[38] The Human Connectome Project. Available at https://www.humanconnectome.org/.
[39] Van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K, et al. The WU-Minn human connectome project: an overview. Neuroimage 2013;80:62–79.
[40] Ashburner J. A fast diffeomorphic image registration algorithm. Neuroimage 2007;38:95–113.
[41] Jin KH, McCann MT, Froustey E, Unser M. Deep convolutional neural network for inverse problems in imaging. IEEE Trans Image Process 2017;26:4509–22.
[42] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv:1502.03167; 2015.
[43] Kingma Diederik, Adam Jimmy Ba. A method for stochastic optimization. arXiv:1412.6980; 2014.
[44] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 2004;13:600–12.
[45] Lee D, Yoo J, Ye JC. Deep residual learning for compressed sensing MRI. IEEE 14th International Symposium on IEEE. 2017. p. 15–8.
[46] Jiang S, Xue H, Glover A, Rutherford M, Rueckert D, Hajnal JV. MRI of moving subjects using multislice snapshot images with volume reconstruction (SVR): application to fetal, neonatal, and adult brain studies. IEEE Trans Med Imaging 2007;26:967–80.
[47] Holdsworth SJ, Aksoy M, Newbould RD, Yeom K, Van AT, Ooi MB, et al. Diffusion tensor imaging (DTI) with retrospective motion correction for large-scale pediatric imaging. J Magn Reson Imaging 2012;36:961–71.
[48] Rousseau F, Glenn OA, Iordanova B, Rodriguez-Carranza C, Vigneron DB, Barkovich JA, et al. Registration-based approach for reconstruction of high-resolution in utero fetal MR brain images. Acad Radiol 2006;13:1072–81.