

OPEN

Retrospective Motion Correction in Multishot MRI using Generative Adversarial Network

Muhammad Usman^{1,2,5}, Siddique Latif^{3,6}, Muhammad Asim¹, Byoung-Dai Lee^{4*} & Junaid Qadir¹

Multishot Magnetic Resonance Imaging (MRI) is a promising data acquisition technique that can produce a high-resolution image with relatively less data acquisition time than the standard spin echo. The downside of multishot MRI is that it is very sensitive to subject motion and even small levels of motion during the scan can produce artifacts in the final magnetic resonance (MR) image, which may result in a misdiagnosis. Numerous efforts have focused on addressing this issue; however, all of these proposals are limited in terms of how much motion they can correct and require excessive computational time. In this paper, we propose a novel generative adversarial network (GAN)-based conjugate gradient SENSE (CG-SENSE) reconstruction framework for motion correction in multishot MRI. First CG-SENSE reconstruction is employed to reconstruct an image from the motion-corrupted k -space data and then the GAN-based proposed framework is applied to correct the motion artifacts. The proposed method has been rigorously evaluated on synthetically corrupted data on varying degrees of motion, numbers of shots, and encoding trajectories. Our analyses (both quantitative as well as qualitative/visual analysis) establish that the proposed method is robust and reduces several-fold the computational time reported by the current state-of-the-art technique.

Magnetic resonance imaging (MRI) is a safe, non-ionizing, and non-invasive imaging modality that provides high resolution and excellent contrast of soft tissues. It has emerged as a powerful and effective technique for early diagnosis of many common but potentially treatable diseases including stroke, cancer, and ischemic heart disease. Despite these advantages, the prolonged data acquisition time of MRI causes many difficulties in its clinical applications, and various research efforts have been proposed in response to expedite the data acquisition process including the use of parallel imaging (PI)¹, compressed sensing (CS)², and echo-planar imaging (EPI)³.

In single-shot echo-planar imaging (EPI), all the k -space data necessary to reconstruct the final magnetic resonance (MR) image is acquired in a single excitation pulse. It significantly accelerates the data acquisition time and minimizes the possibility of motion artifacts in MR images^{4,5}. However, MR images reconstructed using single-shot EPI suffers from low resolution and susceptibility artifacts. To overcome these limitations, segmented EPI or multishot MRI is used⁶, which is a compromise between echo-planar and standard spin-echo imaging. It significantly reduces the demands on gradient performance and allows the in-plane spatial resolution to be improved to a level comparable to that of standard spin echo pulse sequences⁷. In multi-shot MRI, the k -space data is acquired in using a large number of shots at different time instances to obtain the high-resolution volumetric image. As a result, the image may be severely degraded due to subject motion between consecutive shots. This makes the multishot sequences very sensitive to shot-to-shot variabilities caused by the motion.

On the basis of the source of motion, motion in MRI is classified into two categories. *Rigid motion* is caused by the movement of a solid part of body, in which deformation is zero or so small it can be neglected, such as arm, knee, and head motion, while *non-rigid* motion arises from those parts of body, which does not retain any consistent shape, like cardiac motion⁸. The rigid motion produces acute artifacts⁹, which may cause suboptimal image quality, especially intra-brain scan where the contribution of rigid motion is more significant in contrast to non-rigid motion. Subsequently, it may negatively impact radiologic interpretation¹⁰, which affects patient safety

¹Information Technology University (ITU)-Punjab, Lahore, 54700, Pakistan. ²Center for Artificial Intelligence in Medicine and Imaging, HealthHub Co. Ltd., Seoul, 06524, South Korea. ³University of Southern Queensland, Springfield, 4300, Australia. ⁴Kyonggi University, Suwon, 16227, South Korea. ⁵Department of Computer Science & Engineering, Seoul National University, Seoul, 08826, South Korea. ⁶Distributed Sensing Systems Group, Data61, CSIRO, Pullenvale Queensland, 4069, Australia. *email: blee@kyonggi.ac.kr

and enhances the medico-legal risks related to the interpretation of motion degraded images. Therefore, motion correction techniques are considered as an imperative part of MRI reconstruction processes¹¹.

Previously, the problem of motion correction has been solved mostly in an iterative manner¹², which is time-consuming as well as computationally extensive. Researchers are now increasingly interested in leveraging recent advances in deep learning (DL) for improving the state-of-the-art performances in the healthcare^{13,14}. In particular, the use of generative adversarial networks (GANs)¹⁵ is interesting due to its capability of generating data without the explicit modelling of the probability density function and also due to its robustness to over-fitting. The adversarial loss brought by the discriminator formulated in GANs provides a clever way of forcing the generative network to produce sharp and highly continuous data that can be useful for motion correction in MRI.

In this paper, we propose using a GAN-enhanced framework to correct rigid motion in multishot MRI. We focus on brain structural scans due to their frequent use and significance in clinical settings¹⁶. This work is the extension of our previous preliminary work¹⁷, where we empirically showed the suitability of GAN for motion correction in multishot MRI. In particular, we are proposing a GAN-based conjugate gradient (CG) SENSE¹⁸ reconstruction model to correct the motion in multishot MRI. The proposed techniques involve the use of CG-SENSE for the reconstruction of the motion-corrupted multishot k -space data, which is then fed to a GAN to produce an artifact-free image. The proposed technique is effective in reducing motion artifacts and in reducing the computation time, which makes our technique attractive for clinical applications. We have validated our method on publicly available data by changing various parameters of multishot MRI—such as the amount of motion, the number of shots, and the encoding trajectories—with our results showing impressive performance in producing artifact-free image across different parameters in significantly less reconstruction time as compared to traditional iterative techniques.

Background and Related Work

MRI is highly sensitive to subject motion during the k -space data acquisition, which can reduce image quality significantly by inducing motion artifacts. Such artifacts, particularly those produced by rigid motion, are widely observed in multishot MR images during the clinical examination¹⁶. The application of motion correction techniques during or after the reconstruction process, therefore, becomes essential to ensure that an artifact-free image is obtained.

Retrospective motion correction (RMC) techniques are applied to the rigid motion correction^{19,20}. RMC techniques are post-processing techniques employed after the acquisition of the k -space data, while the data is acquired without considering the potential motion⁸ and the object motion is estimated from acquired k -space data. Researchers have proposed a number of RMC-based method for rigid motion correction. For instance, Bydder *et al.*²¹ studied the inconsistencies of k -space caused by subject motion using parallel imaging (PI) technique. The inconsistent data is discarded and replaced with consistent data generated by the PI technique to compensate the motion artifacts. This method produces an image with fewer motion artifacts, albeit at a cost of a lower signal to noise ratio (SNR).

Loktyushin *et al.*²² proposed a joint reconstruction and motion correction technique to iteratively search for motion trajectory. Gradient-based optimization approach has been opted to efficiently explore the search space. The same authors extended their work in a subsequent work²³ by disintegrating the image into small windows that contain local rigid motion and used their own forward model to construct an objective function that optimizes the unknown motion parameters. Similarly, Cordero *et al.*²⁴ proposed the use of a forward model to correct motion artifacts. However, this technique utilises the full reconstruction inverse to integrate the information of multi-coils for estimation and correction of motion. In another study²⁵, authors extended their framework to correct three-dimensional motion (i.e., in-plane and through-plane motion). Through the plane, the motion is corrected by sampling the slices in an overlapped manner.

Conventional techniques such as those just discussed estimate the motion iteratively, which makes them computationally extensive and time-consuming and therefore unsuitable for use in time-critical medical applications. Recent advancements in the field of DL has facilitated significant advance in the medical imaging research community—but very limited attempts have been made for motion correction in MRI. Loktyushin *et al.*²⁶ studied the performance of convolution neural network (CNN) for retrospective motion correction in MR images and proposed training of a model for learning a mapping from motion-corrupted data to motion-free images. The study indicated the potential application of deep neural networks (DNNs) to solve the motion problem in MRI; however, the study did not provide detailed quantitative results or a detailed investigation of the utilized technique. Similarly, Duffy *et al.*²⁷ used CNN to correct motion-corrupted MR images. The work has been compared with traditional Gaussian smoothing²⁸ and significant improvement has been reported but comparison with the advanced state-of-the-art iterative motion correction techniques was unaccounted. Importantly, previous DL-based *motion correction* studies have not exploited GANs despite the fact that GANs have shown excellent performance in MRI reconstruction in particular^{29,30}, and more broadly in modelling natural images^{31,32} and in biomedical image analysis³³.

In our previous work¹⁷, we proposed the use of GAN for multishot MRI motion correction. This work presented the preliminary results on motion correction by notably reducing the computational time. However, the study did not perform a detailed performance evaluation of the proposed multishot MRI framework against the various parameters such as the *number of shots* and the *encoding trajectories*. Building on our previous work, we propose an adversarial CG-SENSE reconstruction framework for the correction of the motion. A detailed analysis of the proposed framework has been presented with respect to different parameters of multishot imaging such as the *levels of motion*, the *number of shots*, and the *encoding trajectories*.

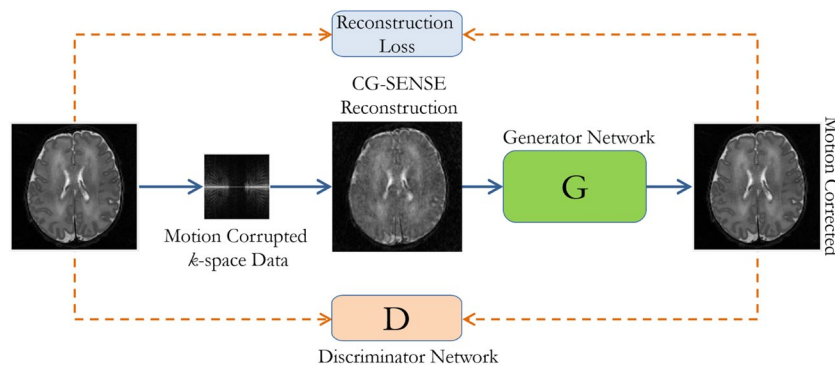


Figure 1. The proposed motion correction framework for multishot MRI, where CG-SENSE is used to reconstruct motion-corrupted images, and the generator network of the GAN, in conjunction with the discriminator network, is tasked with motion correction (Figure Credit: Latif *et al.*¹⁷).

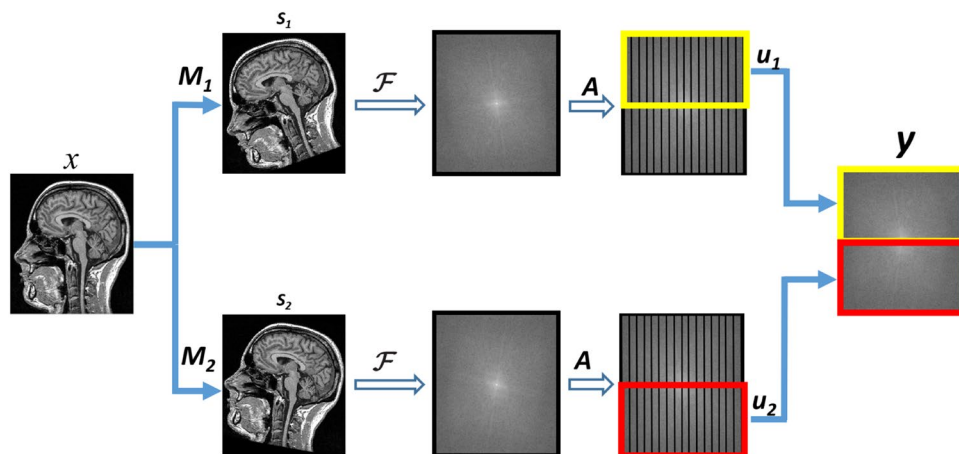


Figure 2. Forward motion corruption model (in 2D) for single coil and two shots MRI: x is the motion-free image; M_s is responsible for introducing the motion in particular shot; F and A employs DFT and sampling; and u_s extracts the k -space segment for each shot.

Methodology

In our proposed method, reconstruction and motion correction are performed, independently. Standard CG-SENSE is employed to reconstruct k -space data which provides a motion-corrupted image in the spatial domain. Motion-corrupted images are given to the GAN for reduction of motion artifacts in the second stage. Figure 1 shows the overall proposed architecture.

Motion model for multishot MRI. In Multishot MRI, k -space data is acquired in multiple shots (i.e., 2, 4 or 8 shots) in order to cover the whole k -space. The MRI scanners capture Fourier coefficients along encoding trajectories that are directed by the gradient shapes of the MRI sequence. For generating motion-corrupted data, we opted the same model as followed by^{24,26}, originally proposed by Batchelor *et al.*¹⁹

In this model, motion M_s is introduced for each s^{th} shot in a motion-free image x . Subsequently, Fourier transform F and sampling matrix A is applied to achieve the k -space representation. Finally, the segment u_s of k -space is extracted for each shot and eventually, all the segments are combined to obtain the full k -space data. Mathematically, it can be written as:

$$y = \sum_{s=1}^N x M_s F A u_s \quad (1)$$

where, N represents the number of shots, M_s the translation as well rotational motion for s^{th} shot, and y the motion-corrupted k -space data. Figure 2 shows the forward motion model for single coil and two shots.

Conjugate gradient SENSE (CG-SENSE) reconstruction. In our proposed technique, we employ CG-SENSE reconstruction technique to reconstruct motion-corrupted k -space data. It utilises conjugate gradient (CG)³⁴ algorithm to efficiently solve the SENSE equations³⁵, which relates the gradient encoding, sensitivities and aliased images to unaliased ones. CG-SENSE algorithm relates the object to be imaged x_m , the encoding matrix E and the acquired k -space data y as follows:

$$Ex_m = y \quad (2)$$

The acquired data y has size $n_c n_k$, where n_c and n_k are the number of coils and the number of sampled positions in k -space, respectively. The size of reconstructed image x_m is N^2 , while N is the matrix size of the image. The spatial encoding information of gradients and coil sensitivities, is presented by the encoding matrix E .

To solve Eq. (2), E has to be inverted, which is a difficult task due to its large size. CG algorithm is used to iteratively solve Eq. (2) for the unaliased image, due to its fast convergence compared to other methods³⁶. To facilitate the formulation of the CG-SENSE reconstruction, another matrix Z is introduced to inverse the encoding as follows:

$$ZE = I_d \quad (3)$$

where, Z and I_d represents the reconstruction matrix and the identity matrix, respectively. Multiplying both sides of Eq. (2) by the F matrix results into an unaliased image which can be described as:

$$x_m = Zy \quad (4)$$

The reconstruction matrix Z can be computed by employing Moore-Penrose inversion:

$$Z = (E^H E)^{-1} E^H \quad (5)$$

Now the set of equations can be solved without finding the inverse of the E matrix by employing CG algorithm. To efficiently perform the CG-SENSE reconstruction process pre-conditioning is performed for better initial estimation of x ³⁶.

Generative adversarial framework. GANs¹⁵ are latent variable generative models that learn via an adversarial process to produce realistic samples from some latent variable code. It includes a generator G and a discriminator D which play the following two-player min-max game:

$$\min_G \max_D P_x[\log(D(x))] + P_z[\log(1 - D(G(z)))] \quad (6)$$

In a simple vanilla GAN, the generator G maps the latent vectors drawn from some known prior p_z (simple distribution e.g. Gaussian) to the sample space. The discriminator D is tasked with differentiating between samples generated $G(z)$ (fake) and data samples (real).

Here, we use conditional GAN³⁷, where instead of random samples, G is fed corrupted MRI images x_m and is trained to produce motion corrected image x_c . The adversarial training loss \mathcal{L}_{adv} for G is defined as

$$\mathcal{L}_{adv} = \log(1 - D(G(x_m))) \quad (7)$$

To facilitate the generator, in addition to the adversarial loss, we also incorporate data mismatch term.

$$\mathcal{L}_{data} = \|x_c - G(x_m)\|_2 \quad (8)$$

Adversarial training encourages the network to produce sharp images, which is of crucial importance in MRI imaging, whereas data mismatch loss forces the network to correctly map degraded images to the original ones. Thus the final loss for G , dubbed generator, is a weighted sum of \mathcal{L}_{data} and \mathcal{L}_{adv} .

$$\mathcal{L} = \mathcal{L}_{data} + \lambda \mathcal{L}_{adv} \quad (9)$$

where λ is a hyper-parameter that controls the weight of each loss term. As training progresses, G and D are trained iteratively.

Experimental Setup

Dataset. For the evaluation of the proposed method, publicly available data is utilized. The data is obtained from the MICCAI Challenge on Multimodal Brain Tumor Segmentation (BraTS) organized by B. Menze, A. Jakab, S. Bauer, M. Reyes, M. Prastawa, and K. Van Leemput^{38–42}. The challenge database contains fully anonymized images from the following institutions: ETH Zurich, University of Bern, University of Debrecen, and University of Utah. We followed the data usage agreement provided by BraTS (<https://www.med.upenn.edu/sbia/brats2018/registration.html>) and all the experiments were carried out in accordance with relevant guidelines and regulations.

We used T2 FLAIR images of high grade (HG) tumor scans; the BraTS 2015 dataset contains 274 HG scans of different subjects. We divided the scans into three subsets including training, validation, and testing sets that contain 191, 25 and 58 scans respectively. Each scan in the dataset has already been normalized into a standard size (i.e., $240 \times 240 \times 255$). However, we further refined the scan while extracting each slice by cropping it from the center and resizing it into 128×128 size. The blank slices in each scan are discarded and a total of 37627, 4875, and 11484 images for training, validation, and testing are produced, respectively. Images of BraTS dataset are considered as motion-free images and motion is introduced by employing the model described in Section 2. The same perturbation technique has been employed in other works^{25,26}. As BraTS contains spatial domain images, we used a reference scan to estimate the coil sensitivity maps as in Allison *et al.*⁴³. For our work, we produce data with varying degrees of angular motion, number of shots, and trajectories to validate the robustness of our proposed technique.

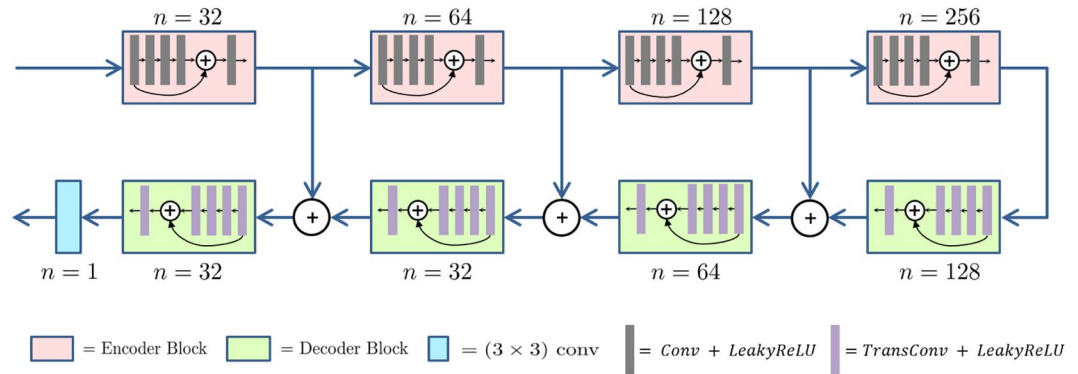


Figure 3. U-Net like Model Architecture used as Generator and Discriminator in GAN.

Model architecture. We adopt a U-Net like architecture (shown in Fig. 3) because of its recent success in image restoration task^{2,44}. It has an hour-glass like structure that involves encoder and decoder networks. The encoder bottlenecks the important information from the corrupted image by reducing the motion artifacts and the decoder is responsible to restore motion free image. In our paper, the encoder consists of convolutional blocks, where each block consists of convolutional layers following by non-linear activation; decoder blocks are composed of transposed convolution layers.

The U-Net architecture also contains symmetric skip connections from encoder blocks to the decoder blocks. It helps to recover fine details for better image restoration: encoder learns to compress image into the high-level features necessary for image restoration, but may remove fine details along with the corruptions, whereas the skip connections from encoder to decoder transfer low-level features from the encoding path to the decoding path to recover the details of the image. In addition to these skip connections, we employ residual connections⁴⁵ inside each encoder and decoder block like Milletari *et al.*⁴⁶. These residual connections along with skip connections allow efficient gradient flow, which helps in alleviating issues such as vanishing gradients and slow convergence.

The high-level model architecture is described in Fig. 3. Each encoder block consists of 5 convolution layers, each with n feature maps except for the layer in the middle with $n/2$ feature maps. Padding is employed to keep the dimension of feature maps same inside each block. We set the strides equal to 1 for all layers except the first one, where we choose it to be 2. This stride 2 convolution serves to down-sample feature maps using a learned kernel. Inside each encoder block, a residual connection is used between the first layer and the last layer. Decoder block has the same structure as the encoder except that we replace all the convolutional layers with transposed convolutions and use a stride of 2 at the last layer instead of the first layer. Here stride 2 transposed convolution serves to up-sample the feature maps along the U-Net architecture. The discriminator is exactly the same as the encoder part of the generator.

Model training. We train our network using training data and validation data was used for parameters selection. We evaluated the model in the testing phase using held out data i.e., testing set. We selected the best value of γ by evaluating the model on different values (0.1, 0.2, 0.4, 0.5, 0.6, 0.8, 1.0) of γ . The value of γ giving the best results on the validation set is used for evaluations in testing phase. For all experiments in this paper, we achieve best results using $\gamma > 0.5$. We optimized the model using RMSProp with the learning rate being 1×10^{-4} until convergence. We use a batch size of 16. For each update of G , we update D twice. We pre-train the generator G using Adam optimizer with same learning rate and batch size. This allows the training of G to converge faster.

Quantifying parameters. We used the following three parameters to measure the quantitative performance of our proposed framework.

Peak signal to noise ratio (PSNR). It is the ratio of maximum possible value (power) of a signal and the power of distorting noise that affects the quality of its representation. We calculated the PSNR of our resultant images by using formulation as follows:

$$PSNR = 20 \log_{10} \left(\frac{\max(r)}{\sqrt{MSE}} \right) \tag{10}$$

where MSE is mean square error which can be calculated as

$$MSE = \left(\frac{1}{mn} \right) \sum_{i=0}^{1-m-1} \sum_{j=0}^{1-n-1} ||r(i, j) - x(i, j)||^2 \tag{11}$$

where r represents the reference image, x denotes the reconstructed image, m and n are the numbers of rows and columns of the reconstructed image, and max function computes the maximum value.

| Degree of motion | 2° | 5° | 8° | 10° | 12° | 14° |
|------------------------------------|---|---|---|---|---|---|
| Peak signal to noise ratio (PSNR) | 32.31 ± 1.93 | 31.57 ± 1.57 | 30.89 ± 1.85 | 28.18 ± 1.92 | 27.85 ± 1.63 | 27.25 ± 1.74 |
| Structural similarity index (SSIM) | 0.96 ± 0.061 | 0.96 ± 0.067 | 0.94 ± 0.047 | 0.92 ± 0.052 | 0.91 ± 0.068 | 0.90 ± 0.063 |
| Artifact power (AP) | $2.47 \times 10^{-3} \pm 0.13 \times 10^{-3}$ | $4.52 \times 10^{-3} \pm 0.17 \times 10^{-3}$ | $6.57 \times 10^{-3} \pm 0.29 \times 10^{-3}$ | $7.31 \times 10^{-3} \pm 0.27 \times 10^{-3}$ | $8.08 \times 10^{-3} \pm 0.29 \times 10^{-3}$ | $9.10 \times 10^{-3} \pm 0.38 \times 10^{-3}$ |

Table 1. Performance metrics of our approach on different amount of motion with 16-shots and random trajectory.

Structural similarity index (SSIM). This is a very common method for predicting the quality of a reconstructed image by checking its similarity with the reference image. The SSIM index is calculated on various windows of an image⁴⁷ which can be formulated as

$$SSIM(r, x) = \frac{(2\mu_r\mu_x + c_1)(2\sigma_{rx} + c_2)}{(\mu_r^2 + \mu_x^2 + c_1)(\sigma_r^2 + \sigma_x^2 + c_2)} \quad (12)$$

Where,

- r is the reference image
- x is the reconstructed image
- μ_r is the mean value of reference image
- μ_x is the mean value of reconstructed image
- σ_r^2 is the variance of r
- σ_x^2 is the variance of x
- σ_{rx} is the covariance of r and x ;

Artifact power (AP). It represents the level of artifacts in any given image with reference to the ground truth. AP can be defined as⁴⁸

$$AP = \frac{\sum |r(i, j) - x|^2}{\sum |r(i, j)|^2} \quad (13)$$

Higher the value of the AP, the more the artifacts; therefore, the reduction of APs is attempted to achieve an artifact-free image.

Results and Discussion

In this section, we have performed a detailed investigation of our proposed technique for the reconstruction of motion-free images in the presence of varying levels of motion, number of shots, and encoding trajectories. Since motion is corrected in the spatial domain which allows the solution to be employed to any kind of motion and encoding/sampling scheme. However, considering the immense range of potential sampling trajectories, acquisition orderings, patterns of motion and number of shots, we restrict our evaluation to a limited set of encoding trajectories, number shots, and degrees of rotational motion. Further, we also perform a comparison of our results with the state-of-the-art technique by Cordero *et al.*²⁵, which has been selected for comparison because of the closeness of its reconstruction process to the one proposed here. Particularly, Cordero *et al.* also used the same forward model of the acquisition process to add perturbation into the motion free image. Most importantly, they demonstrated a significant improvement in terms of reconstruction error as compared to the previous state-of-the-art technique²⁶. For validation, we used peak signal to noise ratio (PSNR), structural similarity index (SSIM), and artifact power (AP) as quantification parameters.

Effect of the levels of motion. Our work is focused on rigid motion correction, specifically for intra-brain scan head motion⁹, which is mostly rotational motion and it causes austere effects in the reconstructed image. Therefore, to evaluate the effect of motion, different rotational motion artifacts have been introduced into motion-free images with 16-shots and random trajectory. The motion-corrupted k -space data has been reconstructed using CG-SENSE (without motion correction) and then fed to the adversarial network, which is tasked to generate motion-free images. Table 1 summarizes the average results obtained for varying degrees of rotational motion ($\Delta\theta = \{2^\circ, 5^\circ, 8^\circ, 10^\circ, 12^\circ, 14^\circ\}$) on test data. It can be noted from Table 1 that the proposed framework shows excellent performance for a small amount of motion by capturing the underlying statistical properties of MR images, and recover sharp and excellent images. However, with the increase in the amount of motion, a smooth decay in the performance of model is observed, as expected, because with higher degree of inter-scan motion (i.e., 14°) MRI scans get severely degraded and it becomes very difficult to recover the motion free image.

Moreover, the performance of our technique is better than the previous state-of-the-art iterative technique²⁴ for higher levels of motion (i.e., $\Delta\theta = 14^\circ$) (see Fig. 4). For a small amount of motion, the approach of Cordero *et al.*²⁴ performs slightly better in terms of AP, however, the long computational time restrains its efficiency.

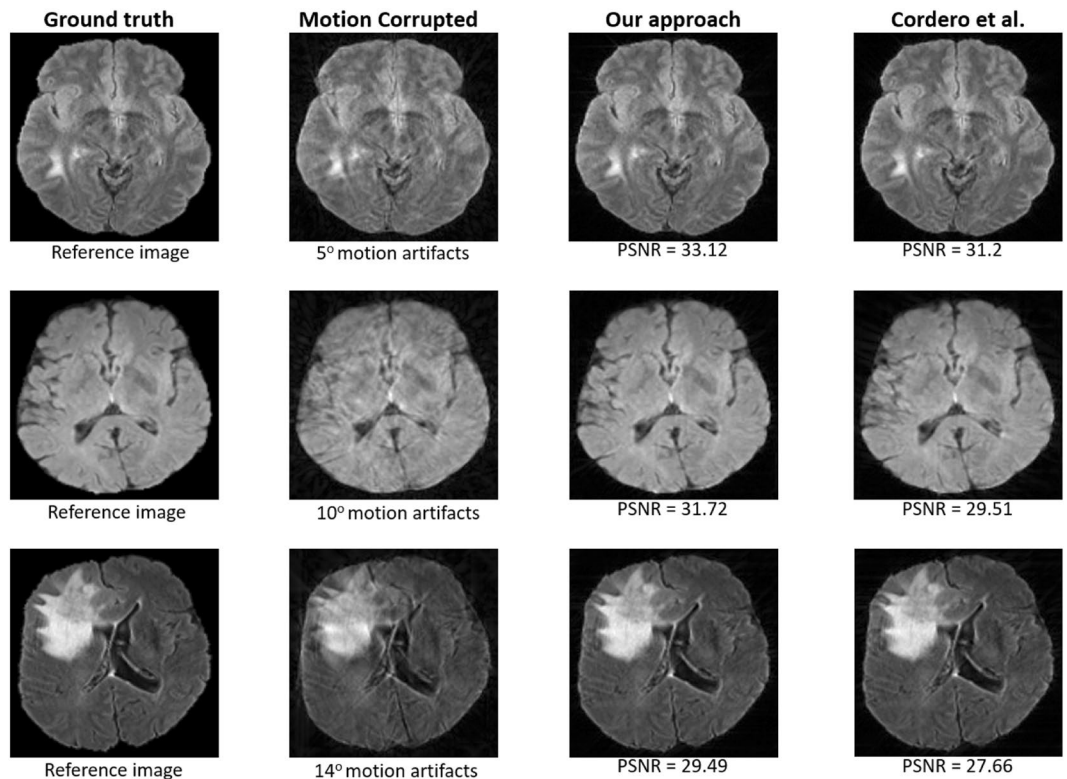


Figure 4. Resultant images produced by our approach compared to those produced by Cordero *et al.*²⁵ for $\Delta\theta = \{5^\circ, 10^\circ, 14^\circ\}$ with 16-shot and random trajectory.

Influence of the number of shots. In this experiment, we investigate the performance of the proposed framework for different number of shots. We generated motion-corrupted data for various number of shots, (i.e., $S = \{2, 4, 8, 16, 32, 64, 128\}$) with *five degree* of motion and *the random trajectory*. We trained our model individually for each *number of shots* and evaluated the performance. The results are summarized in Table 2, which describes the mean values of results obtained on all the test scans. It can be seen from Table 2 that the network is able to learn the artifact pattern and provides significantly promising results for all the number of shots. Encouragingly, our network produces sharp images with high values of PSNR and SSIM even for a higher number of shots. In contrast, state-of-the-art iterative technique²⁴ were only able to correct the motion for lower number of shots effectively. In Fig. 5(a) a snippet of performance comparison against the different number of shots has been shown on fifty randomly selected test images with 2° of motion. It can be witnessed that our method has similar performance for each number of shots, while the conventional technique²⁴ gradually reduces performance with the increase in the number of shots. For higher number of shots ($S > 32$), the convergence of such iterative techniques^{24,26} becomes very difficult. In our case, motion is corrected in the spatial domain after the full reconstruction of the motion-corrupted image, which enables the adversarial network to correct the motion artifacts in the image domain without encountering such convergence challenges.

We also evaluated the robustness of the proposed model. We trained the model using a higher number of shots and testing is performed using a lower number of shots. The model slightly improves the PSNR of the reconstructed image, which is not suitable for real-time applications as shown in Fig. 6. This is due to the fact that the motion artifacts produced in the images reconstructed from the lower number of shots are different than the artifacts produced from the higher number of shots. However, initializing the model with the weights of higher number of shots and fine-tuning with the lower number of shots helps improve convergence. Therefore, we followed the same method for each number of shots to expedite the convergence.

Influence of the encoding trajectory. From the vast range of trajectories, we restricted ourselves to the four trajectories (as shown in Fig. 7) to validate the performance of the proposed framework. The motion-corrupted data of each encoding trajectory is generated with eight number of shots ($S = 8$) and a relative rotation of $\Delta\theta = 5^\circ$ is assumed between shots. We first performed the full reconstruction of motion-corrupted k -space data for each encoding trajectory and then trained GANs with the resultant motion artifact-corrupted images, individually for each trajectory.

Table 3 describes the mean results of our proposed framework for each encoding trajectory. The results show that our approach performs significantly well for all the encoding trajectories. However, it can be noted through close observation that the performance of the proposed technique is slightly better for the *random trajectory* since the random trajectory is least affected by the motion. The same reasoning can be applied for slightly degraded performance for Cartesian sequential trajectory as this trajectory is most affected by the motion artifacts. On the

| Number of Shots | 2 | 4 | 8 | 16 | 32 | 64 | 128 |
|------------------------------------|---|---|---|---|---|---|---|
| Peak signal to noise ratio (PSNR) | 31.82 ± 2.23 | 31.92 ± 2.55 | 31.55 ± 1.89 | 31.57 ± 2.21 | 31.93 ± 1.92 | 32.02 ± 2.08 | 32.08 ± 1.76 |
| Structural similarity index (SSIM) | 0.95 ± 0.057 | 0.96 ± 0.053 | 0.96 ± 0.048 | 0.96 ± 0.045 | 0.96 ± 0.043 | 0.96 ± 0.041 | 0.96 ± 0.042 |
| Artifact power (AP) | $4.52 \times 10^{-3} \pm 0.29 \times 10^{-3}$ | $4.42 \times 10^{-3} \pm 0.26 \times 10^{-3}$ | $4.32 \times 10^{-3} \pm 0.25 \times 10^{-3}$ | $4.13 \times 10^{-3} \pm 0.18 \times 10^{-3}$ | $3.56 \times 10^{-3} \pm 0.18 \times 10^{-3}$ | $3.59 \times 10^{-3} \pm 0.23 \times 10^{-3}$ | $3.46 \times 10^{-3} \pm 0.19 \times 10^{-3}$ |

Table 2. Performance metrics of our approach for varying shots at 5 degree.

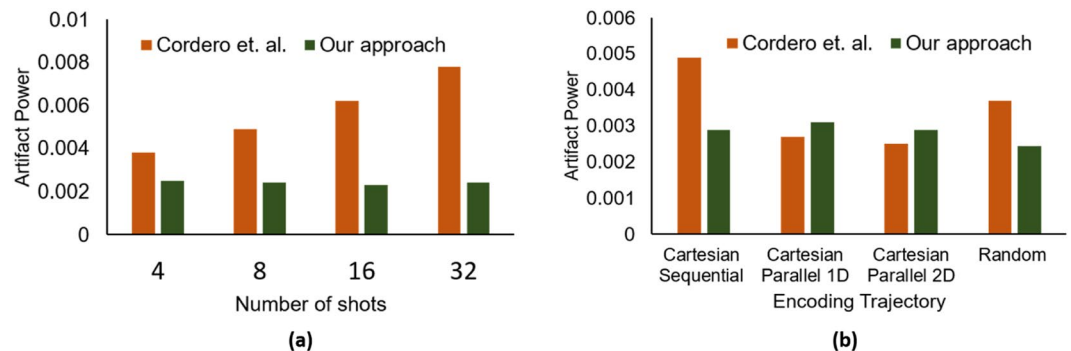


Figure 5. Comparison of our framework with the state of the art iterative technique²⁴ for fifty randomly selected test images in terms of the (a) number of shots and (b) encoding trajectories.

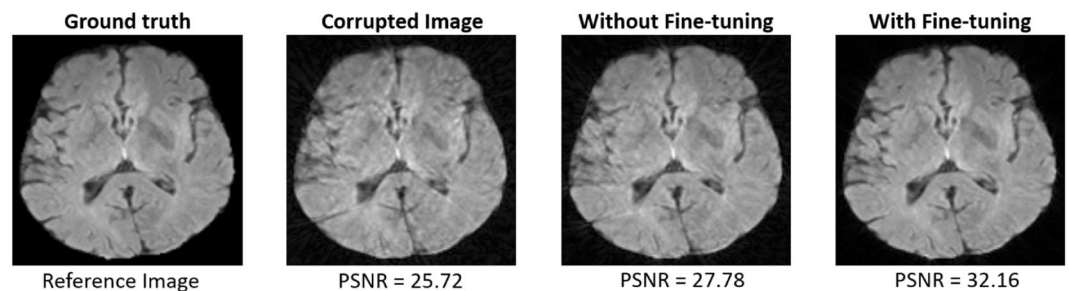


Figure 6. Results from model, without and with fine-tuning, have been shown on a motion-corrupted image (with 16 number of shots).

other hand, the iterative technique²⁴ vigorously changes its performances against different encoding trajectories as depicted in Fig. 5(b) where we compare the proposed technique with the solution proposed by Cordero *et al.*²⁴ for different encoding trajectories on fifty randomly selected test images with 2 degrees of motion. For *Cartesian sequential trajectory*, this technique takes an extraordinarily large number of iterations to reach the convergence, while the proposed technique has universal acceptance and it can be employed to any encoding trajectory.

Computational time analysis. In this section, we present the results of the comparison of the computational time of our technique with the state-of-the-art iterative technique²⁴. To keep our analysis fair, we performed the motion correction of same motion-corrupted k -space data on the same hardware—specifically, an Intel® Core™ i3-2120 CPU with 3.5GHz speed with 16GB of memory and NVIDIA® Quadro M5000 Graphic Processing Unit (GPU) with 8GB GDDR5 memory—by employing both techniques. Since the proposed technique involves two steps (CG-SENSE reconstruction and motion correction), we added the reconstruction and motion correction time to compute the total computational time. Table 4 provides a relative summary of the computational time analysis of our technique compared with the solution proposed by Cordero *et al.*²⁴ for a varying number of shots for 50 randomly selected test images. It can be seen that our technique is several times faster than the previous iterative approach²⁴. The previous technique is an iterative method that first iteratively estimates the motion and then corrects for that motion, which needs extra computational time. With the increase in the number of shots, it becomes difficult to estimate the motion between two consecutive shots, subsequently, it further increases the time required to correct the motion for higher numbers of shots. Moreover, changing the encoding trajectory also significantly affects the computational performance of the conventional iterative technique²⁴. Alternatively, in our proposed technique, motion correction is independent of the reconstruction process and it

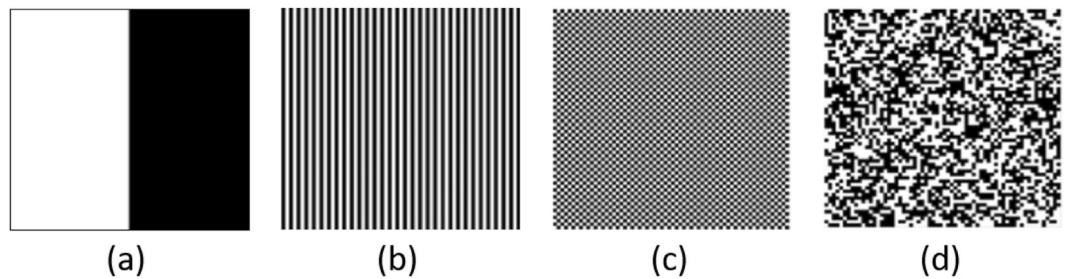


Figure 7. Encoding strategies used for experiments, depicted for $S = 2$ shots: (a) Cartesian sequential, (b) Cartesian parallel 1D, (c) Cartesian parallel 2D, (d) Random; samples corresponding to one of the shots are in white, otherwise samples are in black.

| Sampling trajectory | Cartesian sequential | Cartesian parallel 1D | Cartesian parallel 2D | Random |
|------------------------------------|---|--|--|---|
| Peak signal to noise ratio (PSNR) | 30.11 ± 2.10 | 30.14 ± 1.80 | 30.72 ± 1.53 | 31.55 ± 1.89 |
| Structural similarity index (SSIM) | 0.95 ± 0.057 | 0.95 ± 0.047 | 0.95 ± 0.57 | 0.96 ± 0.062 |
| Artifact power (AP) | $5.51 \times 10^{-3} \pm 0.44 \times 10^{-3}$ | $5.425 \times 10^{-3} \pm 0.37 \times 10^{-3}$ | $5.025 \times 10^{-3} \pm 0.25 \times 10^{-3}$ | $4.32 \times 10^{-3} \pm 0.30 \times 10^{-3}$ |

Table 3. Performance of our approach for different trajectories of multishot MR imaging for eight number of shot ($S = 8$) and 5° of motion.

| Number of shots | 4 | 8 | 16 | 32 | 64 | 128 |
|-----------------------|-------|--------|--------|--------|--------|--------|
| Cordero <i>et al.</i> | 9.00s | 12.80s | 14.59s | 27.83s | 50.07s | 89.19s |
| Our approach | 0.23s | 0.28s | 0.34s | 0.48s | 0.79s | 1.40s |

Table 4. Comparing computational time of our approach with the current state-of-art technique²⁴ against different number of shots.

| Degree of motion | 5° | 8° | 10° | 12° | 14° |
|-----------------------|-----------|-----------|------------|------------|------------|
| Cordero <i>et al.</i> | 25.23s | 39.27s | 57.93s | 134.27s | 152.03s |
| Our approach | 0.28s | 0.28s | 0.28s | 0.28s | 0.28s |

Table 5. Comparing computational time of our approach with the current state-of-art technique²⁴ for various levels of motion.

is performed after full reconstruction of k -space data. Therefore, the motion correction for all the number of shots takes the same computational time. However, the CG-SENSE reconstruction takes more time for the higher number of shots, which slightly increases the overall motion corrected reconstruction time (see Table 4). In Table 5, we summarize the computational time of our technique and iterative technique²⁴, against different levels of motion. The time required to correct for motion in our technique is not dependent upon the amount of motion, therefore, it remains the same for all levels of motion. Alternatively, the conventional technique takes longer time to estimate the higher amount of motion, thus it takes more time to correct such motion.

Conclusion

We introduced a flexible yet robust retrospective motion correction technique that employs generative adversarial networks (GANs) to correct motion artifacts in multishot Magnetic Resonance Imaging (MRI). This work is an extension of our previous preliminary work, where we empirically showed the suitability of GAN for motion correction in multishot MRI. The proposed technique first performs the full reconstruction of motion-corrupted k -space data and then the resultant artifact-affected image is fed into the deep generative networks that learns the mapping from motion artifact-affected images to the artifacts free images. Our GAN based framework significantly reduces the motion artifacts without any prior estimation of motion during the data acquisition or reconstruction process in contrast to the previous iterative methods. Such a parameter-free technique can be employed to any encoding scheme without introducing modifications in the acquisition sequence. To validate our method, we carried out comprehensive experimentation by varying different parameters, such as different levels of motion, the number of shots, and encoding schemes, of multishot MRI. Based on the results, we demonstrated that the performance of the proposed technique is more robust against these parameters and it also

reduced the computational time significantly in contrast to the state-of-the-art techniques. Future plans include the extension of framework to perform end-to-end learning using the generative network from motion-corrupted under-sampled coil information (k -space data) to artifacts free image.

Received: 18 October 2019; Accepted: 2 March 2020;

Published online: 16 March 2020

References

- Larkman, D. J. & Nunes, R. G. Parallel magnetic resonance imaging. *Physics in Medicine & Biology* **52**(7), R15 (2007).
- Lustig, M., Donoho, D. L., Santos, J. M. & Pauly, J. M. Compressed sensing MRI. *IEEE Signal Processing Magazine* **25**(2), 72–82 (2008).
- Mansfield, P. Multi-planar image formation using NMR spin echoes. *Journal of Physics C: Solid State Physics* **10**(3), L55 (1977).
- Rzedzian, R., Mansfield, P., Doyle, M., Guilfoyle, D. & Chapman, B. Real-time nuclear magnetic resonance clinical imaging in paediatrics. *The Lancet* **322**(8362), 1281–1282 (1983).
- Farzaneh, F., Riederer, S. J. & Pelc, N. J. Analysis of T2 limitations and off-resonance effects on spatial resolution and artifacts in echo-planar imaging. *Magnetic Resonance in Medicine* **14**(1), 123–139 (1990).
- Edelman, R. R., Wielopolski, P. & Schmitt, F. Echo-planar MR imaging. *Radiology* **192**(3), 600–612 (1994).
- Bernstein, M. A., King, K. F. & Zhou, X. J. *Handbook of MRI pulse sequences* (Elsevier, 2004).
- Zaitsev, M., Maclaren, J. & Herbst, M. Motion artifacts in MRI: a complex problem with many partial solutions. *Journal of Magnetic Resonance Imaging* **42**(4), 887–901 (2015).
- Gedamu, E. L. & Gedamu, A. Subject movement during multislice interleaved MR acquisitions: Prevalence and potential effect on MRI-derived brain pathology measurements and multicenter clinical trials of therapeutics for multiple sclerosis **36**(2), 332–343 (2012).
- Brown, T. T., Kuperman, J. M., Erhart, M., White, N. S. & Roddey Prospective motion correction of high-resolution magnetic resonance imaging data in children. *Neuroimage* **53**(1), 139–145 (2010).
- Budde, J., Shajan, G., Scheffler, K. & Pohmann, R. Ultra-high resolution imaging of the human brain using acquisition-weighted imaging at 9.4 t. *Neuroimage* **86**, 592–598 (2014).
- Godenschweger, F., Kägebein, U., Stucht, D., Yarach, U. & Sciarra Motion correction in MRI of the brain. *Physics in Medicine and Biology* **61**(5), R32 (2016).
- Usman, M., Latif, S. & Qadir, J. Using deep autoencoders for facial expression recognition. In *2017 13th International Conference on Emerging Technologies (ICET)*, 1–6 (IEEE, 2017).
- Latif, S., Usman, M., Rana, R. & Qadir, J. Phonocardiographic sensing using deep learning for abnormal heartbeat detection. *IEEE Sensors Journal* **18**(22), 9393–9400 (2018).
- Goodfellow, I. et al. Generative adversarial nets. In *Advances in neural information processing systems*, 2672–2680 (2014).
- Andre, J. B. et al. Toward quantifying the prevalence, severity, and cost associated with patient motion during clinical MR examinations. *Journal of the American College of Radiology* **12**(7), 689–695 (2015).
- Latif, S., Asim, M., Usman, M., Qadir, J. & Rana, R. Automating motion correction in multishot MRI using generative adversarial networks. *MED-NIPS* (2018).
- Pruessmann, K. P., Weiger, M., Börner, P. & Boesiger, P. Advances in sensitivity encoding with arbitrary k-space trajectories. *Magnetic Resonance in Medicine* **46**(4), 638–651 (2001).
- Batchelor, P. et al. Matrix description of general motion correction applied to multishot images. *Magnetic Resonance in Medicine* **54**(5), 1273–1280 (2005).
- Samsonov, A. A. et al. POCS-enhanced correction of motion artifacts in parallel MRI. *Magnetic Resonance in Medicine* **63**(4), 1104–1110 (2010).
- Bydder, M., Larkman, D. J. & Hajnal, J. V. Detection and elimination of motion artifacts by regeneration of k-space. *Magnetic Resonance in Medicine* **47**(4), 677–686 (2002).
- Loktyushin, A., Nickisch, H., Pohmann, R. & Schölkopf, B. Blind retrospective motion correction of MR images. *Magnetic Resonance in Medicine* **70**(6), 1608–1618 (2013).
- Loktyushin, A., Nickisch, H., Pohmann, R. & Schölkopf, B. Blind multirigid retrospective motion correction of MR images. *Magnetic Resonance in Medicine* **73**(4), 1457–1468 (2015).
- Cordero-Grande, L. et al. Sensitivity encoding for aligned multishot magnetic resonance reconstruction. *IEEE Transactions on Computational Imaging* **2**(3), 266–280 (2016).
- Cordero-Grande, L., Hughes, E. J., Hutter, J., Price, A. N. & Hajnal, J. V. Three-dimensional motion corrected sensitivity encoding reconstruction for multi-shot multi-slice MRI: Application to neonatal brain imaging. *Magnetic Resonance in Medicine* **79**, 1365–1376 (2018).
- Loktyushin, A., Schuler, C., Scheffler, K. & Schölkopf, B. Retrospective motion correction of magnitude-input MR images. In *Medical Learning Meets Medical Imaging*, 3–12 (Springer, 2015).
- Duffy, B. A. et al. Retrospective correction of motion artifact affected structural MRI images using deep learning of simulated motion. *MIDL* (2018).
- Haddad, R. A. & Akansu, A. N. A class of fast Gaussian binomial filters for speech and image processing. *IEEE Transactions on Signal Processing* **39**(3), 723–727 (1991).
- Mardani, M. et al. Deep generative adversarial networks for compressed sensing automates MRI. *arXiv preprint arXiv:1706.00051* (2017).
- Yang, G. et al. Dagan: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. *IEEE Transactions on Medical Imaging* **37**(6), 1310–1321 (2018).
- Radford, A., Metz, L. & Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434* (2015).
- Ledig, C. et al. Photo-realistic single image super-resolution using a generative adversarial network. In *CVPR*, vol. 2, 4 (2017)
- Wolterink, J. M., Kamnitsas, K., Ledig, C. & Išgum, I. Generative adversarial networks and adversarial methods in biomedical image analysis. *arXiv preprint arXiv:1810.10352* (2018).
- Hestenes, M. R. & Stiefel, E. *Methods of conjugate gradients for solving linear systems*, vol. 49 (NBS Washington, DC, 1952).
- Pruessmann, K. P., Weiger, M., Scheidegger, M. B. & Boesiger, P. Sense: sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine* **42**(5), 952–962 (1999).
- Wright, K. L., Hamilton, J. I., Griswold, M. A., Gulani, V. & Seiberlich, N. Non-cartesian parallel imaging reconstruction. *Journal of Magnetic Resonance Imaging* **40**(5), 1022–1040 (2014).
- Isola, P., Zhu, J.-Y., Zhou, T. & Efros, A. A. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1125–1134 (2017).
- Bjoern, M. et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BraTS). *IEEE Transactions on Medical Imaging*, page 33 (2014).

39. Liu, L. *et al.* Periacetabular osteotomy through the pararectus approach: technical feasibility and control of fragment mobility by a validated surgical navigation system in a cadaver experiment. *International Orthopaedics* **40**(7), 1389–1396 (2016).
40. Lloyd, C. T., Sorichetta, A. & Tatem, A. J. High resolution global gridded data for use in population studies. *Scientific Data* **4**(1), 1–17 (2017).
41. Bakas, S. *et al.* Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge. *arXiv preprint arXiv:1811.02629* (2018).
42. Bakas, S. *et al.* Segmentation labels and radiomic features for the pre-operative scans of the TCGA-LGG collection. *The Cancer Imaging Archive* **286** (2017).
43. Allison, M. J., Ramani, S. & Fessler, J. A. Accelerated regularized estimation of MR coil sensitivities using augmented lagrangian methods. *IEEE Transactions on Medical Imaging* **32**(3), 556–564 (2013).
44. Mao, X.-J., Shen, C. & Yang, Y.-B. Image restoration using convolutional auto-encoders with symmetric skip connections. *arXiv preprint arXiv:1606.08921* (2016).
45. He, K., Zhang, X., Ren, S. & Sun, J. Identity mappings in deep residual networks. In *European conference on computer vision*, 630–645 (Springer, 2016).
46. Milletari, F., Navab, N. & Ahmadi, S.-A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 Fourth International Conference on 3D Vision (3DV)*, 565–571 (IEEE, 2016).
47. Wang, Z., Simoncelli, E. P. & Bovik, A. C. Multiscale structural similarity for image quality assessment. In *The Thirty-Seventh Asilomar Conference on Signals, Systems and Computers, 2003*, vol. 2, 1398–1402 (IEEE, 2003).
48. Omer, H. & Dickinson, R. A graphical generalized implementation of sense reconstruction using MATLAB. *Concepts in Magnetic Resonance Part A* **36**(3), 178–186 (2010).

Acknowledgements

This work was supported by the GRRC program of Gyeonggi province. [GRRC KGU 2017-B04, Image/Network-based Intellectual Information Manufacturing Service Research].

Author contributions

M. Usman and S. Latif conceived the main idea, M. Usman and M. Asim performed all the necessary coding/implementation. M. Usman and S. Latif conducted the experiments, J. Qadir and B. Lee supervised the whole project, analysed the results and provided vital guidance. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to B.-D.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020