Accelerated Magnetic Resonance Imaging by Adversarial Neural Network

Ohad Shitrit, Tammy Riklin Raviv

[1]*Department of Electrical Engineering,* [2]*The Zlotowski Center for Neuroscience Ben-Gurion University of the Negev, Israel*

Abstract

A main challenge in Magnetic Resonance Imaging (MRI) for clinical applications is speeding up scan time. Beyond the improvement of patient experience and the reduction of operational costs, faster scans are essential for time-sensitive imaging, where target movement is unavoidable, yet must be significantly lessened, e.g., fetal MRI, cardiac cine, and lungs imaging. Moreover, short scan time can enhance temporal resolution in dynamic scans, such as functional MRI or dynamic contrast enhanced MRI. Current imaging methods facilitate MRI acquisition at the price of lower spatial resolution and costly hardware solutions.

We introduce a practical, software-only framework, based on deep learning, for accelerating MRI scan time allows maintaining good quality imaging. This is accomplished by partial MRI sampling, while using an adversarial neural network to estimate the missing samples. The inter-play between the generator and the discriminator networks enables the introduction of an adversarial cost in addition to a fidelity loss used for optimizing the peak signal-to-noise ratio (PSNR). Promising image reconstruction results are obtained for 1.5T MRI where only 52% of the original data are used.

Keywords: MRI, GAN, Imaging

---

1. Introduction

Magnetic Resonance Imaging (MRI) is a non-ionizing imaging modality, and is therefore widely used in diagnostic medicine and biomedical research. The physical principles of MRI are based on a strong magnetic field and pulses of radio frequency (RF) electromagnetic radiation. Images are produced when hydrogen atoms, which are prevalent in living organisms, emit the absorbed RF energy that is then received by antennas in close proximity to the anatomy being examined. Spatial localization of the detected MRI signals is obtained by varying the magnetic field gradients. The discretized RF output is presented in a Fourier space (called K-space), where the x-axis is refers to the frequency and the y-axis to the phase. An inverse fast Fourier transform (IFFT) of the K-space is then used for generating anatomically meaningful MRI scans. Figure 1 presents K-space traversal patterns used in conventional imaging. Each row of the k-space is acquired after one RF excitation pulse. The number of rows multiplied by the number of slices (z-axis) determines the total scan time.

The duration of standard single structural MRI acquisition is approximately 5 minutes. Usually, several scans of different modalities or a sequence of scans are acquired such that the overall scan time is much longer. Lengthy imaging process reduces patient comfort and is more vulnerable to motion artifacts. In cases where motion is inevitable, e.g., fetal MRI, cardiac cine, and lungs imaging, scan time must be significantly shortened, otherwise the produced images might be useless. Moreover, in dynamic MRI sequences, acquisition must be brief such that the temporal resolution of the sequence would allow capturing significant temporal changes, e.g., instantaneous increment of the contrast-enhanced material concentration in DCE-MRI or differences in hemodynamic response expressed in fMRI [1].

A straight forward reduction of the scan time can be obtained by sampling fewer slices, thus reducing the spatial resolution in the z-axis. Spatial distances between adjacent slices of fetal MRI or fMRI, for example, are often as high as 0.5 centimeters. Therefore, a significant portion of the potential input is not conveyed through imaging. On the other hand, under-sampling in the x-y domain leads to aliasing, as predicted by the Nyquist sampling theorem.

Numerous research groups as well as leading MRI scanner manufacturers make significant efforts to accelerate the MRI acquisition process. Hardware solutions allow parallel imaging by using multiple coils [2] to sample k-space data. There exist two major approaches [3] that are currently implemented in commercial MRI machines. Both reconstruct an image from the under-sampled k-space data provided by each of the coils. The sensitivity encoder (SENSE) transforms the partial k-spaces into images, then merges the resulting aliased images into one coherent image [4]. The GeneRalized Autocalibrating Partial Parallel Acquisition (GRAPPA) techniques [5] operate on signal data within the complex frequency domain before the IFFT.

The compressed sensing (CS) technique [6] allows efficient acquisition and reconstruction of a signal with fewer samples than the Nyquist-Shannon sampling theorem requires, if the signal has sparse representation in a known transform domain. Using CS for MRI reconstruction by sampling a small subset of the k-space grid had been proposed in [7]. The underlying assumption is that the undersampling is random, such that the zero-filled Fourier reconstruction exhibits incoherent artifacts that behave similarly to additive random noise. This, however, would require specified pulse programming.

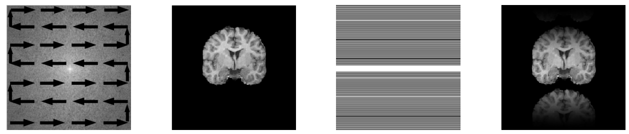
Recently machine learning techniques based on manifold learning [8, 9] and dictionary learning [10, 11] were suggested for MRI reconstruction. MRI reconstruction using convolutional neural networks (CNN) was introduced in [12]. The network learns the mapping between zero-filled and fully-sampled MR images. In [13], residual network was proposed for MRI super-resolution. Their model is able to receive multiple inputs acquired from different viewing planes for better image reconstruction. Both works address the reconstruction problems in the image domain rather than the k-space domain. The proposed framework utilizes recent advances in deep learning, while similarly to the CS methods addresses MRI reconstruction directly from the k-space. Specifically, we use generative adversarial networks (GAN) [14, 15, 16]. GANs are based on the inter-play between two networks: a generator and a discriminator. The generator is capable of learning the distribution over a data-base, and sample realizations of this distribution. The discriminator is trained to distinguish between 'generated' samples and real ones. This powerful combination has been used for generating Computed Tomography (CT)-like images from MRIs [17]. Here, the generator is used for reconstruction of the entire k-space grid from under-sampled data. Its loss is a combination of an adversarial loss, based on the discriminator output and a fidelity loss with respect to the fully sampled MRI. Promising results are obtained for brain MRI reconstruction using only 52% of the data.

The paper is organized as follows. Section 2 presents some theoretical foundation and our method. Section 3 describes the experimental results. Conclusions and future directions are describes in section 4.

2. Method

2.1. K-space

Let $\boldsymbol{u}$ denote the desired signal, a 2D MR image, obtained by the IFFT of the complex k-space signal s_0. Let M_F denote a full sampling mask such that the reconstructed MR



(a) raw k-space (b) fully sampled MRI (c) sampling mask (d) Zero-filled

Figure 1: Under-sampling artifacts: the arrows illustrate the sampling methodology

image is:

$$\boldsymbol{u} = F^H M_F \odot s_0 \tag{1}$$

where H is the Hermitian transpose operation, $\odot$ denotes element-wise multiplication, and F^H is an orthonormal 2D IFFT operator, such that $F^H F = I$. While sampling part of the k-space, using M_p as a sampling mask, the reconstructed MR image suffers from artifacts and aliasing. An example of the under-sampling (52%) artifacts is shown in Figure **1**.

2.2. Objective

Let $s_p = M_p \odot s_0$ denote the under-sampled k-space. Given a sampling mask and a model f, defined by the set of parameters Θ, our goal is to estimate the missing k-space samples such that:

$$\Theta = \arg\min_{\Theta} L(F^H f\left(s_p; \Theta\right), \boldsymbol{u}) \tag{2}$$

where $L\left(\cdot\right)$ is the loss function. While choosing the loss to be L2 norm is reasonable for natural images, for the k-space, which has different spatial features, this may not be enough. As mentioned in [16], L2 minimization provides a blurry solution. Averaging the high frequency details in the k-space domain results in very poor reconstruction. In order to address this problem, we used the adversarial loss, based on GAN.

We trained our model using the adversarial strategy, as described in [14, 15]. This method is based on a generator G, which takes noise z with uniform distribution $p_u\left(z\right)$ as input and generates samples from the data distribution. A discriminator D is trained to distinguish between "real" examples from the data and generated ("fake") examples from G. During the training process, we optimize G to maximize the discriminator's probability of error. Simultaneously, D is getting better and provides more accurate predictions.

Let s_0 denote a "real" k-space sample from the distribution $p_r(s_0)$. The following optimization process can be described by two-players min-max game:

$$\min_G \max_D \mathbb{E}_{\mathrm{s_0 \sim p_r(s_0)}} \log\left[\mathrm{D}\left(\mathrm{x}\right)\right] + \mathbb{E}_{\mathrm{z \sim p_u(z)}} \log\left[1 - \mathrm{D}\left(\mathrm{G(z)}\right)\right] \tag{3}$$

In equilibrium, the generator G is able to generate samples that look like the real data. In our case, G estimates the missing k-space samples from a linear combination of the sampled data and a uniform noise with distribution $p_u(z)$. An L2 fidelity constraint is added to the adverbial loss of the generator, as follows:

$$\begin{aligned} L_G = \alpha \cdot E_{\mathrm{z \sim p_u(z)}} \log\left[1 - \mathrm{D}\left(\mathrm{F}^{-1}\left(\hat{\mathrm{s}}_0\right)\right)\right] + \\ \beta \cdot \|\left(1 - M_p\right) \odot \left(\hat{s}_0 - s_0\right)\|_2^2 \end{aligned} \tag{4}$$

where $\hat{s}_0$ is the estimated k-space and $\alpha = 1$, $\beta = 0.8$ are hyperparameters tuned by a cross-validation process. The discriminator's input is the reconstructed MR image, i.e., after IFFT. By that, we are integrating the reconstruction phase in our optimization.

2.3. Network Architecture

The generator input is a two-channel image representing the real and the imaginary parts of the partially sampled k-space image, s_p. Each missing sample is initialized by uniform i.i.d. noise. The pixel (i,j) in the generator input image is:

$$G_{in}\left(i,j\right) = s_{p_{i,j}} + \left(1 - M_p\right)_{i,j} z_{i,j} \tag{5}$$

Due to the combination of the adversarial and the fidelity loss, G produces reasonable k-space samples from a given samples and noise distribution $p_u(z)$. In order to use the sampled data, s_p, and estimate only the missing samples we used a residual network [18] as used in [13], such that:

$$\hat{s}_0 = s_p + \left(1 - M_p\right) \odot G_{out} \tag{6}$$

where G_{out} is the generator output. Figure 2 describes our framework:

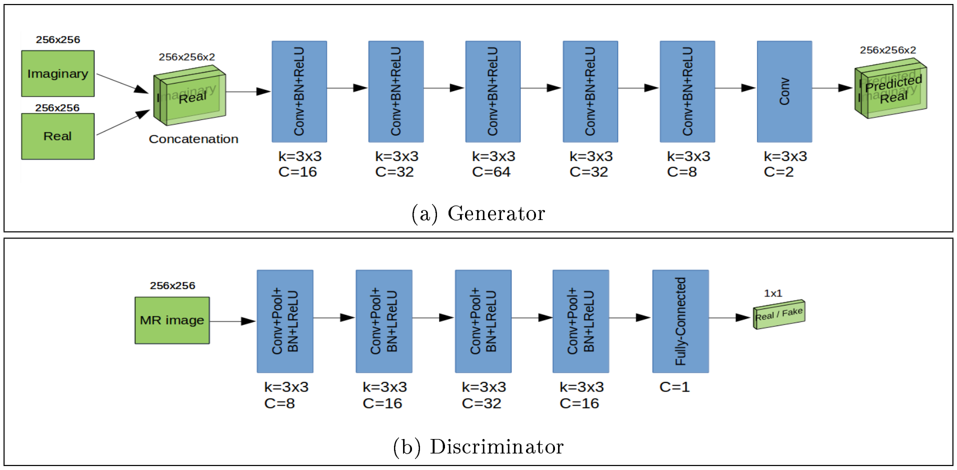


Figure 3: Networks architecture. The generator input is a two-channel signal, real and imaginary. For each layer, k is the kernel size and C is the number of output channels.

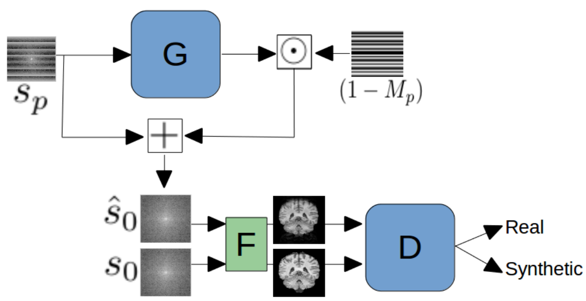


Figure 2: Framework architecture: G and D are the generator and discriminator networks, respectively. F is a $2D$ IFFT operator.

95 A common architecture is used for the discriminator, composed of convolutional layers, batch normalization, and leaky-ReLU as suggested in [15]. For the generator, we compose a dedicated architecture based on multi-channel input for representing the real and imaginary components. Both architectures are shown in Figure 3. The training methodology is doing k_g generator update steps for each discriminator single step.

3. Experiments

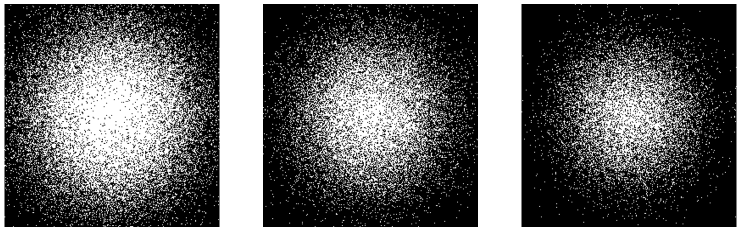
The training data consists of 500 $3D$ brain MRI (T1) scans of different patients from the IXI dataset[1]. The data has been acquired by three MR machines, Philips 1.5T, 3T and GE 1.5. All images padded to resolution of 256×256 pixels. From each $3D$ volume we extract 93 $2D$ saggital slices. We used $37.2k$ (80%) 2D slices for training and $9.2k$ (20%) for testing (100 $3D$ volumes). In order to create k-space images for training, inverse orthonormal 2D FFT is applied to the fully-sampled MR images. We sample the k-space using $2D$ Gaussian mask with sampling factor of 2.5, 4 and 6. Data augmentation is created by random offsets of the proposed mask and image flipping. This leads us to reconstruction of the MR image from 16.6%-40% (Figure 4) of the original k-space data.

The generator is composed of 5 blocks of CONV-BatchNorm-ReLU, with output channels $16, 32, 64, 32, 8$, respectively. The last layer is CONV with two outputs channels (for real and imaginary parts). The discriminator is composed of 4 blocks of CONV-Pool-BatchNorm-LReLU with output channels $8, 16, 32, 16$ and one fully-connected layer. All CONV layers kernel size is 3×3. All weights was initialized by Xavier [19]. We used RMSprop solver with fixed learning rate of $5\text{e}{-6}$ and set k_d to 1.

We compare the proposed method to reconstruction results obtained by using a conventional compressed sensing method CS-MRI [7] and Zero-filling. In addition, we trained a generator (G) using only L2 loss (CNN-L2). The same sampling masks was used for all cases.

A common metric used to quantify the reconstructed image quality is the PSNR. In order to provide a reliable and robust evaluation of the reconstruction quality, which will be used for medical diagnostic, we suggest the following test: PSNR, brain's shape measure and segmentation.

[1] `http://brain-development.org/ixi-dataset/`



(a) Factor 2.5 (40%) (b) Factor 4 (25%) (c) Factor 6 (16.66%)

Figure 4: $2D$ Gaussian sampling masks

3.1. Peak signal-to-noise ratio (PSNR)

The first measure and the most basic one. PSNR measures the Mean Squared Error (MSE) between the fully-sampled MR image and the reconstructed image. Therefore, this measure might no be a good metric for edges and general shape. For example, a model can provides a good reconstruction in the sense of PSNR but with blurry edges, results in a poor performance of algorithms that are based on them, for example segmentation. However, PSNR is still a good indication for the image quality, especially if an expert should view it. Quantitative evaluation is presented in Table 1 and a graphical visualization in Figure 3.1. The PSNR calculated on the whole image without masking. Note that the proposed method outperforms the other.

PSNR	Factor 2.5		Factor 4		Factor 6	
Method	Mean	std	Mean	std	Mean	std
Zero-filled	32.044	2.616	26.447	1.997	16.470	2.181
CS-MRI	39.053	2.479	33.273	2.452	26.951	3.380
CNN-L2	38.978	2.454	33.405	2.232	31.010	2.299
Proposed	39.802	2.489	34.595	2.519	31.555	2.487

Table 1: Error in PSNR, without masking

An extended validation of our model done by calculating the PSNR measure for different brain tissues. We applied an image segmentation algorithm, FAST [20], on the original fully-sampled MR image and then use it for calculating the masked-PSNR on gray matters, white matters and the Cerebrospinal Fluid (CSF). results are presented in Table 2.

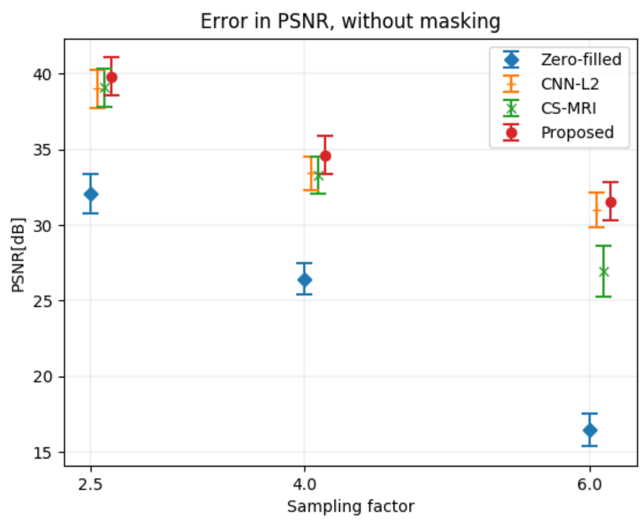


Figure 5: PSNR without masking error-bar

PSNR	Factor 2.5			Factor 4			Factor 6		
Method	White	Gray	CSF	White	Gray	CSF	White	Gray	CSF
Zero-filled	41.494	36.736	38.546	37.279	34.379	36.369	24.887	24.987	29.337
CS-MRI	44.374	40.042	40.744	41.655	37.247	39.119	35.595	33.111	36.648
CNN-L2	45.727	41.593	41.689	43.265	40.151	40.247	41.119	39.277	39.102
Proposed	46.206	41.942	41.616	43.857	40.109	40.496	41.919	39.014	39.366

Table 2: Error in PSNR, with masking

3.2. Brain Extraction - Skull Stripping

Brain extraction (Skull stripping) is a an algorithm that delineates the brain boundary (Figure 6). It is necessary for almost every brain analysis algorithm. In tissue segmentation for example, skull stripping is a pre-processing step which affects directly on the segments partition. We examine the different reconstruction methods by applying the Brain Extraction Tool (BET) [21] on each different MR reconstructed image. Then, we compared the skull stripping results to the fully-sampled result using the Modified Hausdorff Distance (MHD) [22].

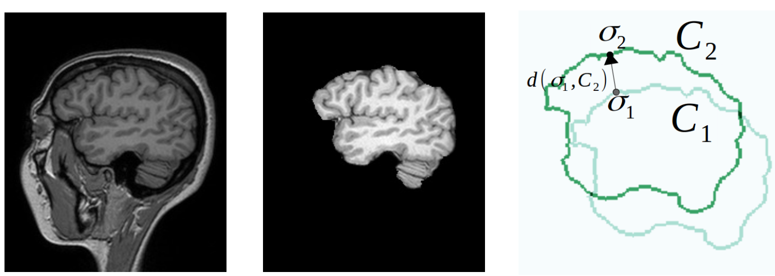
Let $C_1, C_2 \in R^2$ denotes the brain contours extract from skull stripping algorithm on two different reconstruction methods respectively. MHD measures the distance between the contours such that:

$$\begin{aligned} MHD(C_1, C_2) &= \max\left\{ \tfrac{1}{|C_1|} \textstyle\sum_{\sigma_1 \in C_1} d(\sigma_1, C_2),\ \tfrac{1}{|C_2|} \textstyle\sum_{\sigma_2 \in C_2} d(\sigma_2, C_1) \right\} \\ d(\sigma_1, C_2) &= \min_{\sigma_2 \in C_2} ||\sigma_1 - \sigma_2|| \\ d(\sigma_2, C_1) &= \min_{\sigma_1 \in C_1} ||\sigma_2 - \sigma_1|| \end{aligned} \tag{7}$$

where σ_1, σ_2 are points on the contours C_1, C_2 respectively (Figure 6.c). MHD values for different sampling ratios are presented in Table 3.

MHD	Factor 2.5		Factor 4		Factor 6	
Method	Mean	std	Mean	std	Mean	std
Zero-filled	1.111	0.563	2.617	1.214	3.121	1.279
CS-MRI	0.701	0.511	1.447	1.027	3.114	1.617
CNN-L2	0.420	0.270	0.715	0.561	1.083	1.052
Proposed	0.391	0.250	0.617	0.306	1.050	1.033

Table 3: MHD - brain extraction



(a) MR image (b) Brain extraction (c) Hausdorff distance

Figure 6: Example of brain extraction

3.3. Brain Segmentation

Talk about segmentation (FSL, FAST, etc..)

Why it is important

Noise and smoothing vs edges - important for segmentation algorithms

What is Dice score

Table

Images

4. Conclusions

We proposed a software-only framework, using GANs for accelerating MRI acquisition. Specifically, high-quality MRI reconstruction using only 52% of the original k-space data is demonstrated. The key idea is based on utilizing an adversarial loss in addition to L2 loss. It is worth mentioning that the proposed sampling mask is currently implemented in commercial MRI machines with no need for additional hardware or dedicated pulse programming. Future work will concentrate on generation of MRI in the presence of pathologies.

4.1. Acknowledgment

This research is partially supported by the Israel Science Foundation (T.R.R. 1638/16) and the IDF Medical Corps (T.R.R.)

References

[1] S. Moeller, E. Yacoub, C. A. Olman, E. Auerbach, J. Strupp, N. Harel, K. Uğurbil, Multiband multislice ge-epi at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fmri, Magnetic Resonance in Medicine 63 (5) (2010) 1144–1153.

[2] P. B. Roemer, W. A. Edelstein, C. E. Hayes, S. P. Souza, O. Mueller, The nmr phased array, Magnetic resonance in medicine 16 (2) (1990) 192–225.

[3] A. Deshmane, V. Gulani, M. A. Griswold, N. Seiberlich, Parallel mr imaging, Journal of Magnetic Resonance Imaging 36 (1) (2012) 55–72.

[4] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, P. Boesiger, et al., Sense: sensitivity encoding for fast mri, Magnetic resonance in medicine 42 (5) (1999) 952–962.

[5] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, A. Haase, Generalized autocalibrating partially parallel acquisitions (grappa), Magnetic resonance in medicine 47 (6) (2002) 1202–1210.

[6] D. L. Donoho, Compressed sensing, Information Theory, IEEE Transactions on 52 (4) (2006) 1289–1306.

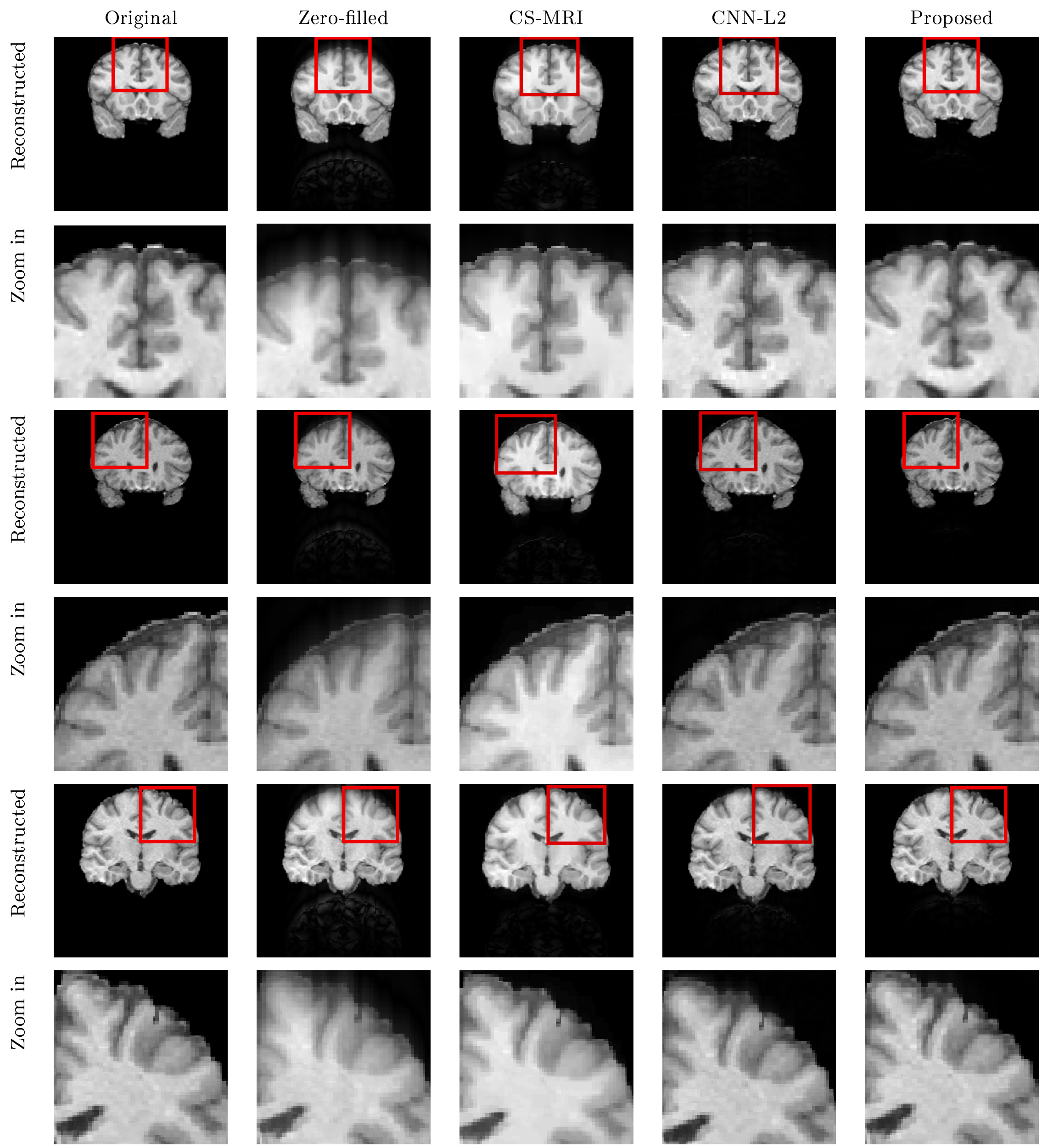


Figure 7: Examples of reconstructed MR images from under-sampled k-space.

[7] M. Lustig, D. Donoho, J. M. Pauly, Sparse mri: The application of compressed sensing for rapid mr imaging, Magnetic resonance in medicine 58 (6) (2007) 1182–1195.

[8] M. Usman, G. Vaillant, D. Atkinson, T. Schaeffter, C. Prieto, Compressive manifold learning: Estimating one-dimensional respiratory motion directly from undersampled k-space data, Magnetic Resonance in Medicine 72 (4) (2014) 1130–1140.

[9] K. K. Bhatia, J. Caballero, A. N. Price, Y. Sun, J. V. Hajnal, D. Rueckert, Fast reconstruction of accelerated dynamic mri using manifold kernel regression, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015, Springer, 2015, pp. 510–518.

[10] S. Ravishankar, Y. Bresler, Mr image reconstruction from highly undersampled k-space data by dictionary learning, Medical Imaging, IEEE Transactions on 30 (5) (2011) 1028–1041.

[11] J. Caballero, A. N. Price, D. Rueckert, J. V. Hajnal, Dictionary learning and time sparsity for dynamic mr data reconstruction, Medical Imaging, IEEE Transactions on 33 (4) (2014) 979–994.

[12] S. Wang, Z. Su, L. Ying, X. Peng, S. Zhu, F. Liang, D. Feng, D. Liang, Accelerating magnetic resonance imaging via deep learning, in: Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on, IEEE, 2016, pp. 514–517.

[13] O. Oktay, W. Bai, M. Lee, R. Guerrero, K. Kamnitsas, J. Caballero, A. de Marvao, S. Cook, D. OARegan, D.Regan, D. Rueckert, Multi-input cardiac image super-resolution using convolutional neural networks, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2016, pp. 246–254.

[14] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, in: Advances in neural information processing systems, 2014, pp. 2672–2680.

[15] A. Radford, L. Metz, S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks, arXiv preprint arXiv:1511.06434.

[16] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, A. A. Efros, Context encoders: Feature learning by inpainting, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2536–2544.

[17] D. Nie, R. Trullo, C. Petitjean, S. Ruan, D. Shen, Medical image synthesis with context-aware generative adversarial networks, arXiv preprint arXiv:1612.05362.

[18] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[19] X. Glorot, Y. Bengio, Understanding the difficulty of training deep feedforward neural networks., in: Aistats, Vol. 9, 2010, pp. 249–256.

[20] Y. Zhang, M. Brady, S. Smith, Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm, IEEE transactions on medical imaging 20 (1) (2001) 45–57.

[21] S. M. Smith, Fast robust automated brain extraction, Human brain mapping 17 (3) (2002) 143–155.

[22] M.-P. Dubuisson, A. K. Jain, A modified hausdorff distance for object matching, in: Pattern Recognition, 1994. Vol. 1-Conference A: Computer Vision & Image Processing., Proceedings of the 12th IAPR International Conference on, Vol. 1, IEEE, 1994, pp. 566–568.