

Joint Deep Model-based MR Image and Coil Sensitivity Reconstruction Network (Joint-ICNet) for Fast MRI

Yohan Jun Hyungseob Shin Taejoon Eo Dosik Hwang^{*}
 School of Electrical and Electronic Engineering, Yonsei University

Abstract

Magnetic resonance imaging (MRI) can provide diagnostic information with high-resolution and high-contrast images. However, MRI requires a relatively long scan time compared to other medical imaging techniques, where long scan time might occur patient's discomfort and limit the increase in resolution of magnetic resonance (MR) image. In this study, we propose a Joint Deep Model-based MR Image and Coil Sensitivity Reconstruction Network, called Joint-ICNet, which jointly reconstructs an MR image and coil sensitivity maps from undersampled multi-coil k-space data using deep learning networks combined with MR physical models. Joint-ICNet has two main blocks, where one is an MR image reconstruction block that reconstructs an MR image from undersampled multi-coil k-space data and the other is a coil sensitivity maps reconstruction block that estimates coil sensitivity maps from undersampled multi-coil k-space data. The desired MR image and coil sensitivity maps can be obtained by sequentially estimating them with two blocks based on the unrolled network architecture. To demonstrate the performance of Joint-ICNet, we performed experiments with a fastMRI brain dataset for two reduction factors ($R = 4$ and 8). With qualitative and quantitative results, we demonstrate that our proposed Joint-ICNet outperforms conventional parallel imaging and deep-learning-based methods in reconstructing MR images from undersampled multi-coil k-space data.

1. Introduction

Magnetic resonance imaging (MRI) is one of the medical imaging techniques widely used for diagnosis and research. However, MRI requires a long scan time that may cause patient discomfort and increase the probability of motion-related artifacts in the reconstructed image. One way to reduce the scan time of MRI is to acquire undersampled k-space data instead of acquiring whole k-space data. However, sampling below the Nyquist rate in a k-space domain

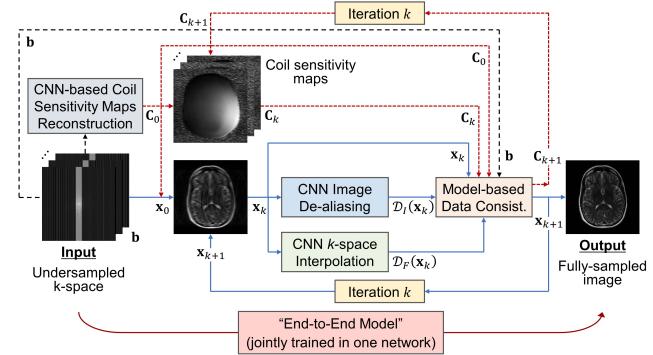


Figure 1. Overall framework of our proposed joint deep model-based magnetic resonance image and coil sensitivity reconstruction network (**Joint-ICNet**).

causes aliasing artifacts in a spatial domain. Several methods have been proposed to reconstruct an artifact-free MR image from undersampled k-space. One is a compressed sensing-based method that uses the sparsity of an MR image in the transformed domain, such as the wavelet domain and total variation (TV), to reconstruct the image [20, 21]. The other is a parallel imaging (PI)-based method that unfolds aliasing patterns in an image domain using pre-acquired coil sensitivity maps or interpolates unacquired k-space points using acquired multi-coil k-space points such as autocalibration signal (ACS) [8, 23, 29]. In an undersampled MR image reconstruction, a forward model of MRI that includes Fourier transform, coil sensitivity maps, and k-space sampling matrix can be used as prior information, which is a model-based reconstruction of MR image [7]. Model-based MR image reconstruction methods are based on iterative algorithms to reconstruct MR images from undersampled k-space data.

In recent years, deep learning methods demonstrated superior performance in the inverse problem of imaging including denoising, compressive sensing, super-resolution, and inpainting compared to conventional analytical methods [34, 19, 18, 16]. In MRI, several methods have been proposed that used deep learning algorithms in an MR image reconstruction. Deep-learning-based MR image re-

^{*}Corresponding author.

construction methods can be classified into two categories based on the role of the deep learning algorithms: 1) a direct mapping method that used deep learning algorithms for removing aliasing artifacts in an image domain [12, 17], interpolating missing k-space data using acquired data in a k-space domain [2, 10], and directly estimating a target image from undersampled k-space data using a manifold learning [6, 36]; and 2) a model-based method that incorporated deep learning algorithms as regularization functions with unrolled architectures [1, 5, 9, 25, 13].

Most of the current deep-learning-based MR reconstruction methods used the coil sensitivity maps that were pre-acquired in an acquisition protocol or pre-computed by coil sensitivity estimation methods such as ESPIRiT method [29]. However, with a high reduction factor or with fewer ACS lines of acquired multi-coil k-space data, estimated coil sensitivity maps may not be accurate and could affect the reconstruction performance of MR images [32]. To overcome the problem, several studies that jointly reconstruct both MR images and coil sensitivity maps have been proposed [32, 28]. They showed that jointly reconstructing both MR images and coil sensitivity maps increased the reconstruction performance of MR images. One study has been proposed to estimate coil sensitivity maps from undersampled multi-coil k-space data using deep learning networks in reconstructing MR images [27]; however, similar to conventional coil sensitivity estimation methods, there is a limitation that the initially estimated coil sensitivity maps were not jointly updated during the MR image reconstruction.

In this study, we propose a Joint Deep Model-based MR Image and Coil Sensitivity Reconstruction Network (**Joint-ICNet**) that jointly reconstructs an MR image and coil sensitivity maps from undersampled multi-coil k-space data using deep learning networks and interleaved MR model-based data consistency schemes. An overall framework of the proposed Joint-ICNet is presented in Fig. 1. The main contributions of this study are as follows.

First, Joint-ICNet is a joint MR image and coil sensitivity maps reconstruction network that combines a deep learning network with a model-based method. We merge a deep-learning-based method, which demonstrates superior performance in an MR image reconstruction, and a model-based method, which uses analytic MR signal models as prior information, to improve the reconstruction performance of an MR image and coil sensitivity maps.

Second, Joint-ICNet consists of two main blocks: 1) an MR image reconstruction block that reconstructs an MR image from undersampled multi-coil k-space data using deep learning networks and a model-based data consistency layer, and 2) a coil sensitivity maps reconstruction block that reconstructs coil sensitivity maps from undersampled multi-coil k-space data using a deep learning network and

a model-based data consistency layer. In this paper, we demonstrate the performance of Joint-ICNet with a fastMRI dataset [33]. The preliminary work for this study has been accepted for presentation at the ISMRM Conference 2021.

2. Theory

2.1. Problem Formulation

An MR image reconstruction problem is to recover a desired image from measured multi-coil k-space data. Mapping the MR image to the multi-coil k-space data can be represented using a forward operator \mathbf{A} :

$$\mathbf{Ax} + \mathbf{n} = \mathbf{b}, \quad (1)$$

where $\mathbf{x} \in \mathbb{C}^N$ denotes the desired MR image, $\mathbf{b} \in \mathbb{C}^{NN_c}$ denotes the measured multi-coil k-space, $N = N_x \times N_y$ denotes the matrix size of the image, N_c denotes the number of coils, $\mathbf{n} \in \mathbb{C}^{NN_c}$ is the complex additive Gaussian noise, and $\mathbf{A} : \mathbb{C}^N \mapsto \mathbb{C}^{NN_c}$ has a coil sensitivity map $\mathbf{C} \in \mathbb{C}^{N \times N}$, Fourier transform $\mathcal{F} \in \mathbb{C}^{N \times N}$, and a k-space sampling matrix $\mathbf{M} \in \mathbb{R}^{N \times N}$.

Eq. (1) is an ill-posed problem when the undersampled multi-coil k-space data \mathbf{b} were measured, and an explicit solution for \mathbf{x} is hard to obtain. In general, Eq. (1) can be solved in \mathbf{x} by minimizing the least squares problem with a regularization term

$$\min_{\mathbf{x}} \|\mathbf{Ax} - \mathbf{b}\|_2^2 + \lambda \mathcal{R}(\mathbf{x}), \quad (2)$$

where $\mathcal{R}(\mathbf{x}) : \mathbb{C}^N \mapsto \mathbb{C}^N$ is the regularization term of \mathbf{x} and λ is the regularization parameter.

In PI-based methods, the coil sensitivity maps \mathbf{C} are needed to reconstruct the image \mathbf{x} from the undersampled multi-coil k-space data \mathbf{b} . Eq. (2) assumes constant coil sensitivity maps \mathbf{C} that were pre-acquired or pre-computed. Coil sensitivity maps are usually obtained with calibration scan data; however, obtaining additional data can increase the total scan time, which might weaken the benefit of PI by reducing the scan time with undersampling. The method also has the problem of inconsistency between the acquired data and the calibration scan data [32]. Another method for acquiring coil sensitivity maps is to compute them from ACS lines of acquired multi-coil k-space data by a coil estimation method such as ESPIRiT method [29]. Unfortunately, the estimated coil sensitivity maps might not be accurate with fewer ACS lines especially when the reduction factor is high.

Considering that the estimated coil sensitivity maps were incomplete especially with fewer ACS lines, the coil sensitivity maps \mathbf{C} need to be updated with the image \mathbf{x} by minimizing the least squares problem with an additional regularization term of the coil sensitivity maps. Then, Eq. (2) can be extended as

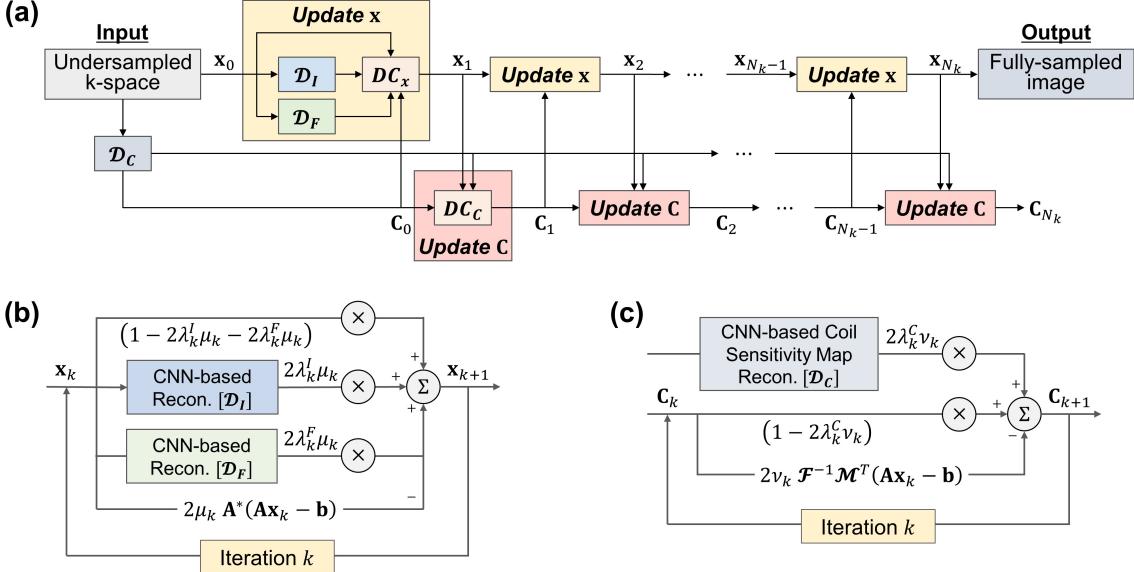


Figure 2. Detailed architecture of Joint-ICNet. (a) Unrolled architecture of Joint-ICNet that has two blocks, where one is a magnetic resonance (MR) image reconstruction block and the other is a coil sensitivity maps reconstruction block. (b) MR image reconstruction block has two convolutional neural network (CNN)-based regularizations and a model-based data consistency layer of MR image. (c) Coil sensitivity maps reconstruction block has a CNN-based regularization and a model-based data consistency layer of coil sensitivity map.

$$\min_{\mathbf{x}, \mathbf{C}} \|\mathbf{Ax} - \mathbf{b}\|_2^2 + \lambda_x \mathcal{R}(\mathbf{x}) + \lambda_C \mathcal{R}(\mathbf{C}), \quad (3)$$

where $\mathcal{R}(\mathbf{C}) : \mathbb{C}^{N \times N} \mapsto \mathbb{C}^{N \times N}$ is the regularization term of \mathbf{C} , and λ_x and λ_C are the regularization parameters that control the balance between the data consistency and the regularization terms of \mathbf{x} and \mathbf{C} , respectively.

2.2. Joint Reconstruction of Image and Coil Sensitivity Map

In general, sparsity transforms such as wavelet transform and TV are used for \mathcal{R} ; however, these regularizers have problems such as residual artifacts and over-smoothing in the reconstructed images [3, 15, 20, 9]. To overcome the problems, deep-learning-based regularizations have been proposed, which can learn optimized regularization functions by training and show better performance for inverse problems in imaging than conventional models [4, 35]. Thus, we proposed to use the deep-learning-based regularization that removes noise and aliasing artifacts in an image domain. Compared to a general image-based regularization, an additional k-space-based regularization term is used to compensate it. Then, the deep-learning-based regularization of \mathbf{x} can be represented as

$$\mathcal{R}(\mathbf{x}) = \|\mathbf{x} - \mathcal{D}_I(\mathbf{x})\|_2^2 + \|\mathbf{x} - \mathcal{F}^{-1}\mathcal{D}_F(\mathbf{f})\|_2^2, \quad (4)$$

where $\mathcal{F}^{-1} \in \mathbb{C}^{N \times N}$ represents the inverse Fourier transform, $\mathbf{f} \in \mathbb{C}^N$ represents the k-space of \mathbf{x} that is Fourier

transformed by \mathcal{F} , $\mathcal{D}_I : \mathbb{C}^N \mapsto \mathbb{C}^N$ is the de-aliasing model that reconstructs artifacts-free MR images from \mathbf{x} , and $\mathcal{D}_F : \mathbb{C}^N \mapsto \mathbb{C}^N$ is the k-space model that interpolates missing k-space data points from \mathbf{f} . One study has demonstrated that utilizing the dual-domain regularization, i.e. image domain and k-space domain, improved the performance of an MR image reconstruction [5]. The dual-domain regularization terms can be extended to a PI and expected to increase the reconstruction performance.

Furthermore, we proposed to use the deep-learning-based regularization of coil sensitivity maps. An estimation of coil sensitivity maps is difficult with highly undersampled multi-coil k-space data that do not have enough ACS lines. Incomplete coil sensitivity maps could affect the reconstruction performance of an MR image. To overcome the problem, several methods that jointly reconstructed coil sensitivity maps and an MR image with a regularization have been proposed. Previous studies used a TV as a regularization assuming that coil sensitivities are sparse in TV of the spatial domain or used prior knowledge of coil sensitivity maps (i.e. smoothness) for regularizing coil profiles such as Sobolev norm [26, 32, 28].

Compared to the previous studies, we proposed to use the deep-learning-based regularization of coil sensitivity maps that can find the optimal coil sensitivity maps by training to reconstruct the image. The regularization of \mathbf{C} can be represented as similar to the regularization of \mathbf{x} , which can be presented as follows:

$$\mathcal{R}(\mathbf{C}) = \|\mathbf{C} - \mathcal{D}_C(\mathcal{F}^{-1}\mathbf{b})\|_2^2, \quad (5)$$

where $\mathcal{D}_C : \mathbb{C}^{N \times N} \mapsto \mathbb{C}^{N \times N}$ is the coil sensitivity model that estimates coil sensitivity maps from the acquired k-space data \mathbf{b} . Recently, a deep-learning-based coil sensitivity maps estimation model has been proposed by [27]; however, the estimated initial coil sensitivity maps using the model were not jointly updated with the image. Here, we proposed to use the coil sensitivity model for regularizing the problem of Eq. (3) to find the optimal \mathbf{x} .

If we combine the proposed regularizations with the pre-defined problem formulation, Eq. (3) can be represented as

$$\begin{aligned} & \min_{\mathbf{x}, \mathbf{C}} \|\mathbf{Ax} - \mathbf{b}\|_2^2 + \lambda_I \|\mathbf{x} - \mathcal{D}_I(\mathbf{x})\|_2^2 \\ & + \lambda_F \|\mathbf{x} - \mathcal{F}^{-1}\mathcal{D}_F(\mathbf{f})\|_2^2 + \lambda_C \|\mathbf{C} - \mathcal{D}_C(\mathcal{F}^{-1}\mathbf{b})\|_2^2, \end{aligned} \quad (6)$$

where λ_I and λ_F represent the regularization parameters of \mathcal{D}_I and \mathcal{D}_F , respectively. The least squares problem of Eq. (6) can be solved sequentially in terms of \mathbf{x} and \mathbf{C} , respectively, as

$$\begin{aligned} & \min_{\mathbf{x}} \|\mathbf{Ax} - \mathbf{b}\|_2^2 + \lambda_I \|\mathbf{x} - \mathcal{D}_I(\mathbf{x})\|_2^2 \\ & + \lambda_F \|\mathbf{x} - \mathcal{F}^{-1}\mathcal{D}_F(\mathbf{f})\|_2^2, \end{aligned} \quad (7)$$

and

$$\min_{\mathbf{C}} \|\mathbf{Ax} - \mathbf{b}\|_2^2 + \lambda_C \|\mathbf{C} - \mathcal{D}_C(\mathcal{F}^{-1}\mathbf{b})\|_2^2. \quad (8)$$

The least squares problem of Eqs. (7) and (8) can be solved using a gradient descent method with an iterative algorithm.

$$\begin{aligned} \mathbf{x}_{k+1} = & \mathbf{x}_k - 2\mu_k [\mathbf{A}^*(\mathbf{Ax}_k - \mathbf{b}) + \lambda_k^I(\mathbf{x}_k - \mathcal{D}_I(\mathbf{x}_k)) \\ & + \lambda_k^F(\mathbf{x}_k - \mathcal{F}^{-1}\mathcal{D}_F(\mathbf{f}))], \end{aligned} \quad (9)$$

and

$$\begin{aligned} \mathbf{C}_{k+1} = & \mathbf{C}_k - 2\nu_k [\mathcal{F}^{-1}\mathbf{M}^\top(\mathbf{Ax}_k - \mathbf{b})\mathbf{x}_k^* \\ & + \lambda_k^C(\mathbf{C}_k - \mathcal{D}_C(\mathcal{F}^{-1}\mathbf{b}))], \end{aligned} \quad (10)$$

where \mathbf{x}_k and \mathbf{C}_k are the reconstructed image and coil sensitivity maps at iteration k ($k = 0, \dots, N_k$), N_k represents the number of iterations, μ_k and ν_k are the step size of \mathbf{x}_k and \mathbf{C}_k at iteration k , respectively, λ_k^I and λ_k^F are the regularization parameters of \mathcal{D}_I and \mathcal{D}_F at iteration k , respectively, $\mathbf{A}^* : \mathbb{C}^{NN_c} \mapsto \mathbb{C}^N$ is the adjoint of \mathbf{A} , \mathbf{x}_k^* is the adjoint of \mathbf{x}_k , and \mathbf{M}^\top is the transpose of \mathbf{M} . The gradient of $\mathcal{D}_I(\mathbf{x})$, $\mathcal{F}^{-1}\mathcal{D}_F(\mathbf{f})$, and $\mathcal{D}_C(\mathcal{F}^{-1}\mathbf{b})$ can be approximated as zero for a small perturbation of \mathbf{x} and \mathbf{C} [1].

\mathbf{x}_k and \mathbf{C}_k are sequentially reconstructed from the Eqs. (9) and (10).

A deep model-based MR image and coil sensitivity reconstruction network is structured based on Eqs. (9) and (10). In each iteration of the unrolled network, the model first reconstructs the MR image from the acquired undersampled multi-coil k-space data using the deep-learning-based MR image regularizations operating both in image and k-space domains, and then reconstructs the coil sensitivity maps using the previously reconstructed image and the deep-learning-based coil sensitivity maps regularization.

3. Methods

3.1. Proposed Model: Joint-ICNet

The detailed architecture of Joint-ICNet is presented in Fig. 2. Joint-ICNet consists of two main blocks: 1) an **MR image reconstruction block** that reconstructs an MR image from undersampled multi-coil k-space data with convolutional neural network (CNN)-based regularizations and a model-based data consistency layer of an MR image, and 2) a **coil sensitivity maps reconstruction block** that estimates coil sensitivity maps from undersampled multi-coil k-space data with a CNN-based regularization and a model-based data consistency layer of coil sensitivity maps. Those two blocks are constructed based on the Eqs. (9) and (10). Two blocks are sequentially placed between the input and the output. The output of the MR image reconstruction block is fed into the next coil sensitivity maps reconstruction block, and the output of the coil sensitivity maps reconstruction block is then fed into the next image reconstruction block. The desired MR image and coil sensitivity maps can be obtained by sequentially estimating them with two blocks.

3.2. MR Image Reconstruction Block

Fig. 2(b) presents the detailed architecture of the MR image reconstruction block that has the CNN-based regularizations \mathcal{D}_I that operates on the image domain and \mathcal{D}_F that operates on the k-space domain, and the embedded model-based data consistency layer. The CNN-based MR image reconstruction block is structured based on Eq. (9), which can be rewritten as follows:

$$\begin{aligned} \mathbf{x}_{k+1} = & (1 - 2\lambda_k^I\mu_k - 2\lambda_k^F\mu_k)\mathbf{x}_k \\ & + 2\mu_k(\lambda_k^I\mathcal{D}_I(\mathbf{x}_k) + \lambda_k^F\mathcal{F}^{-1}\mathcal{D}_F(\mathbf{f})) \\ & - 2\mu_k\mathbf{A}^*(\mathbf{Ax}_k - \mathbf{b}), \end{aligned} \quad (11)$$

where λ_k^I , λ_k^F , and μ_k are the trainable parameters of the regularizations and the step size in iteration k , respectively. The model-based data consistency layer of \mathbf{x} is structured based on Eq. (11), and $\mathbf{x}_{k+1} \in \mathbb{C}^N$ (i.e., the $(k+1)$ -th iteration output) is reconstructed using the previous k -th

iteration output \mathbf{x}_k , $\mathcal{D}_I(\mathbf{x}_k)$, $\mathcal{F}^{-1}\mathcal{D}_F(\mathbf{f})$, and $\mathbf{A}^*(\mathbf{Ax}_k - \mathbf{b})$.

The CNN-based regularization \mathcal{D}_I , which removes aliasing artifacts in the image, and \mathcal{D}_F , which interpolates k-space data, are structured based on U-net architecture [24]. The encoder and decoder of U-net consist of several convolutional blocks, each of which has a convolutional layer, regularization layer, and nonlinear activation function. In addition, \mathcal{D}_I and \mathcal{D}_F have a residual learning structure [11], which is implemented by adding the output from the input of \mathcal{D}_I and \mathcal{D}_F . To train \mathcal{D}_F that operates on the k-space domain, Fourier transform is applied to the input image and inverse Fourier transform is applied to the output, respectively.

3.3. Coil Sensitivity Maps Reconstruction Block

The detailed architecture of the coil sensitivity maps reconstruction block that has the CNN-based regularization \mathcal{D}_C and the embedded model-based data consistency layer is illustrated in Fig. 2(c). The CNN-based coil sensitivity maps reconstruction block is structured based on Eq. (10), which can be rewritten as follows:

$$\begin{aligned} \mathbf{C}_{k+1} = & (1 - 2\lambda_k^C \nu_k) \mathbf{C}_k + 2\lambda_k^C \nu_k \mathcal{D}_C(\mathcal{F}^{-1}\mathbf{b}) \\ & - 2\nu_k \mathcal{F}^{-1}\mathbf{M}^\top(\mathbf{Ax}_k - \mathbf{b}) \mathbf{x}_k^*, \end{aligned} \quad (12)$$

where λ_k^C and ν_k are the trainable parameters of the regularizations and the step size in iteration k , respectively. The model-based data consistency layer of \mathbf{C} is structured based on Eq. (12), and $\mathbf{C}_{k+1} \in \mathbb{C}^{N \times N}$ (i.e., the $(k+1)$ -th iteration output) is reconstructed using the previous k -th iteration output \mathbf{C}_k , $\mathcal{D}_C(\mathcal{F}^{-1}\mathbf{b})$, and $\mathcal{F}^{-1}\mathbf{M}^\top(\mathbf{Ax}_k - \mathbf{b}) \mathbf{x}_k^*$.

Similar to \mathcal{D}_I and \mathcal{D}_F , the CNN-based regularization \mathcal{D}_C , which reconstructs the coil sensitivity maps from the undersampled multi-coil k-space data, is structured based on U-net architecture [24]. The input of \mathcal{D}_C is the undersampled multi-coil k-space data and the output is the estimated coil sensitivity maps. Similar to other coil sensitivity estimation methods, low frequency lines (i.e. ACS lines) of k-space were extracted from the undersampled multi-coil k-space data before feeding into the network.

3.4. Unfolded Network

The two proposed blocks, i.e. MR image reconstruction block and coil sensitivity maps reconstruction block, are sequentially placed between the input and the output. The output of the MR image reconstruction block is fed into the next coil sensitivity maps reconstruction block, and the output of the coil sensitivity maps reconstruction is fed into the next MR image reconstruction block. The MR image and coil sensitivity maps are jointly reconstructed using those two blocks in one network.

The Joint-ICNet is implemented in one network and trained in an end-to-end manner with a structural similarity (SSIM) loss function [31]. The loss function \mathcal{L} that calculates the error between the output of the network and the target is given by

$$\mathcal{L} = \sum_{N_S} (1 - \text{SSIM}(|\mathbf{x}_{N_k}|, |\mathbf{x}_T|)), \quad (13)$$

where N_S is the number of training data, $\mathbf{x}_{N_k} \in \mathbb{C}^N$ is the output of the network that is the output of the MR image reconstruction block at the final iteration N_k , and $\mathbf{x}_T \in \mathbb{C}^N$ represents the target image that is the fully sampled MR image.

4. Experiments

4.1. Implementation Details

Implementation of Joint-ICNet. We converted the complex data into the real-valued data that concatenates real and imaginary channels. Then, the data and networks can be redefined as follows: $\mathbf{x}_k \in \mathbb{R}^{2N}$ for $(k = 0, \dots, N_k)$, $\mathbf{C}_k \in \mathbb{R}^{2N \times N}$ for $(k = 0, \dots, N_k)$, $\mathcal{D}_I: \mathbb{R}^{2N} \mapsto \mathbb{R}^{2N}$, $\mathcal{D}_F: \mathbb{R}^{2N} \mapsto \mathbb{R}^{2N}$, and $\mathcal{D}_C: \mathbb{R}^{2N \times N} \mapsto \mathbb{R}^{2N \times N}$.

We used U-net based architectures for the CNN-based regularizations \mathcal{D}_I , \mathcal{D}_F , and \mathcal{D}_C , where U-net has four pooling and up-sampling layers. Each convolutional block located in U-net consisted of a 2D convolutional layer followed by a leaky rectified linear unit (leaky ReLU) with 0.2 negative slope coefficient and instance normalization [30]. The kernel size of all convolutional layers was 3×3 , and the number of feature maps starts from 32, 32, and 4, which are doubled after max-pooling layer and halved after up-sampling layer for \mathcal{D}_I , \mathcal{D}_F , and \mathcal{D}_C , respectively. The trainable parameters of regularization λ_k , and step sizes μ_k and ν_k were initialized as 1. A total of 10 iterations (i.e., $N_k = 10$) was performed in the unrolled network. Before calculating the loss \mathcal{L} , a root sum-of-squares (RSS) operation was applied to the output and the target.

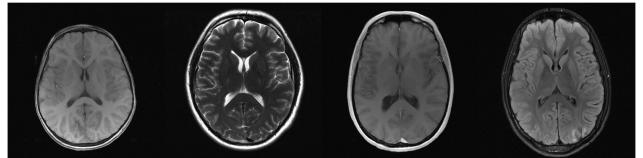


Figure 3. Representative examples of T1 weighted, T2 weighted, T1 weighted with contrast agent, and FLAIR images from the fastMRI dataset [33] used in this study.

We implemented Joint-ICNet using a Pytorch library [22]. Joint-ICNet was trained using an Adam optimizer [14] with $\beta_1 = 0.9$ and $\beta_2 = 0.999$ for 50 epochs with a learning rate of 0.0005. The training took approximately 48 h using 8 NVIDIA TITAN RTX GPUs.

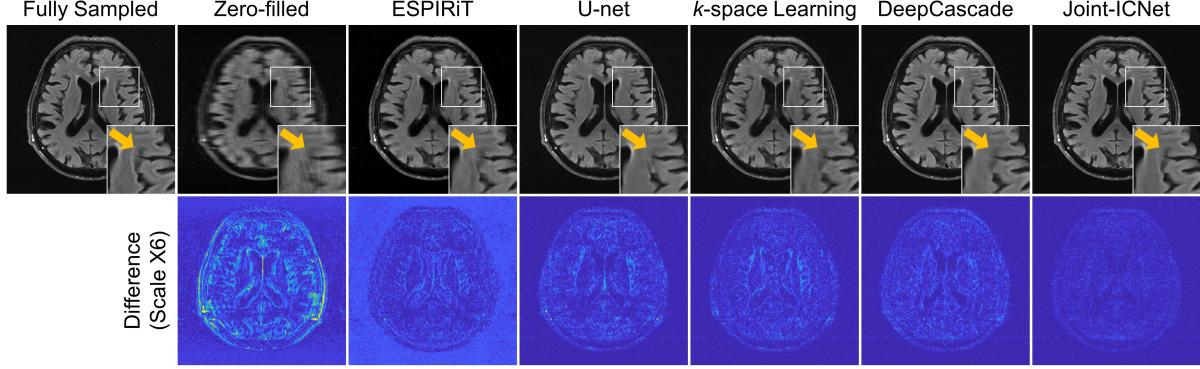


Figure 4. Fully sampled and reconstructed FLAIR images using zero-filling, ESPIRiT, U-net, k-space learning, DeepCascade, and Joint-ICNet, with the reduction factor $R = 4$.

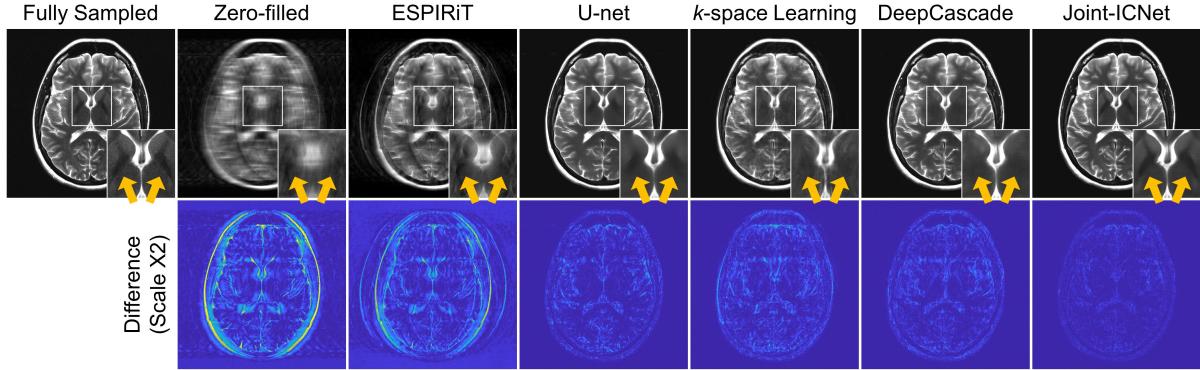


Figure 5. Fully sampled and reconstructed T2 images using zero-filling, ESPIRiT, U-net, k-space learning, DeepCascade, and Joint-ICNet, with the reduction factor $R = 8$.

Dataset and Sampling Mask. We used multi-coil k-space data of the fastMRI open dataset [33]. The dataset has brain MRI scans including T1 weighted, T1 weighted with contrast agent (T1POST), T2 weighted, and FLAIR images taken from various vendors. The dataset includes 4,469 scans for the training (70,748 axial slices), 1,378 scans for the validation (21,842 axial slices), and 558 scans for the test (8,852 axial slices). The representative examples are presented in Fig. 3. During the training, 20% of the slices were used for the training and the validation, which were randomly selected from the dataset. For the test, the whole test slices were used. The multi-coil k-space data were retrospectively undersampled using 1D Cartesian equispaced sampling masks. The reduction factors were $R = 4$ and 8 .

Evaluation Metrics. Three metrics including a normalized root mean squared error (NRMSE), a peak signal-to-noise ratio (PSNR), and SSIM [31] were used to quantitatively evaluate the reconstructed images.

4.2. Comparison Studies

Comparison Studies. We compared Joint-ICNet with a conventional PI-based method, l_1 -ESPIRiT [29], and deep-learning-based MR reconstruction methods, which were U-

net [33], k-space learning [10], and DeepCascade [25]. U-net method reconstructs the image from the RSS undersampled image. k-space learning method uses U-net based architecture for reconstructing the image from the undersampled multi-coil k-space data. DeepCascade method has 10 convolutional blocks with interleaved data consistency layers. Each convolutional block has 5 convolutional layers with a residual learning layer.

Ablation Studies. To evaluate the performance of Joint-ICNet, we compared the original Joint-ICNet model with five different models as follows: 1) Joint-ICNet with a different number of iterations ($N_k = 1, 2, 4, 6, 8$, and 10); 2) Joint-ICNet with \mathcal{D}_C replaced by a conventional coil sensitivity estimation method (ESPIRiT), called *Replace \mathcal{D}_C with ESPIRiT*; 3) Joint-ICNet without a k-space regularization \mathcal{D}_F , called *w/o \mathcal{D}_F* ; and 4) Joint-ICNet without data consistency layers, called *w/o DC layers*.

The first model has a different number of iterations N_k to see the effectiveness of the iteration. The second model (*Replace \mathcal{D}_C with ESPIRiT*) reconstructs coil sensitivity maps with a conventional coil sensitivity estimation method that is ESPIRiT instead of \mathcal{D}_C in Joint-ICNet to evaluate the effectiveness of the \mathcal{D}_C . The third model (*w/o \mathcal{D}_F*) recon-

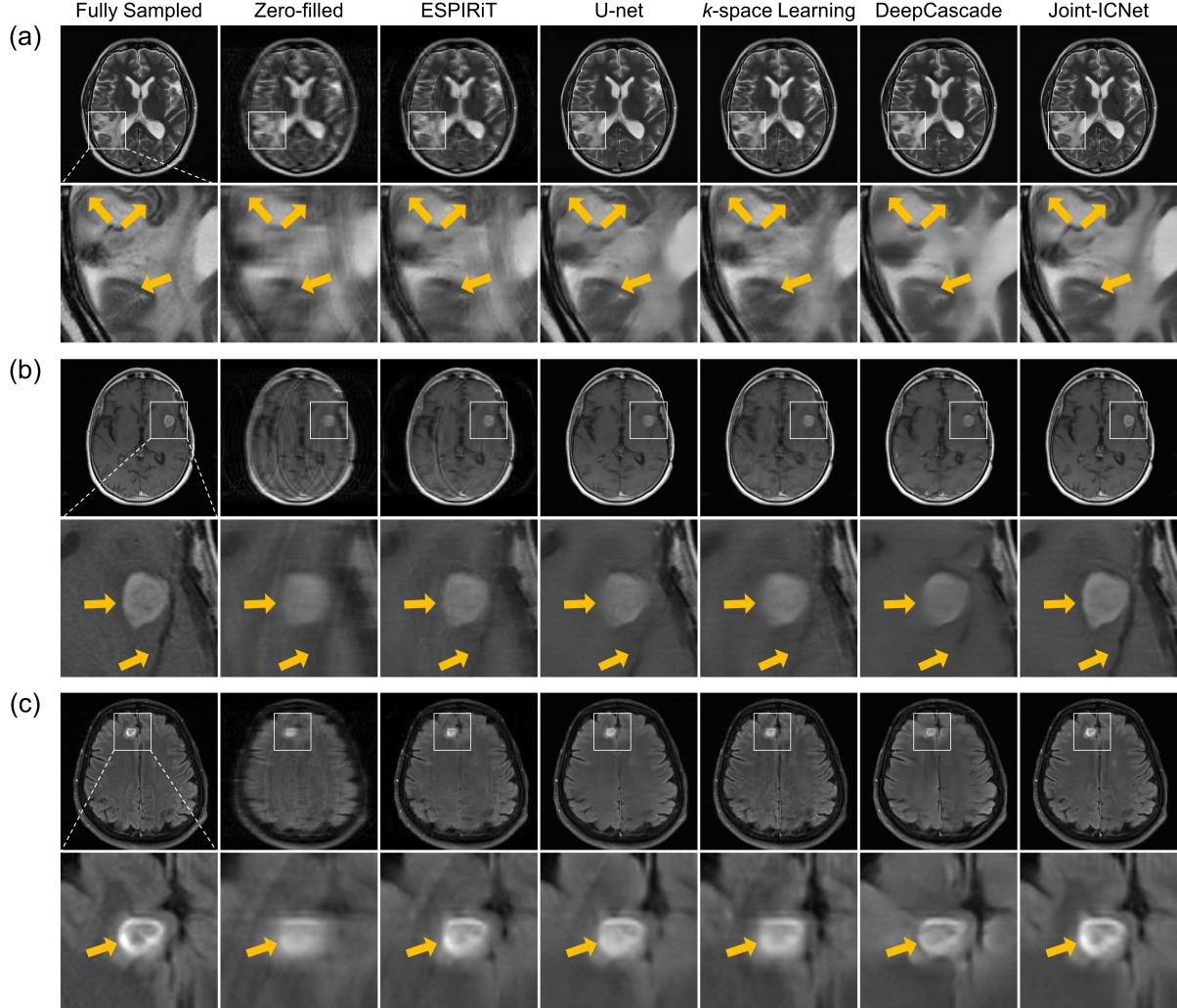


Figure 6. Fully sampled and reconstructed (a) T2, (b) T1POST, and (c) FLAIR images with abnormal cases using zero-filling, ESPIRiT, U-net, k-space learning, DeepCascade, and Joint-ICNet, with the reduction factor $R = 4$.

structs MR images without \mathcal{D}_F to see the effectiveness of the k-space regularization. The fourth model (*w/o DC layers*) does not have data consistency layers in the model.

4.3. Experimental Results

Comparison with Other Reconstruction Methods. Fig. 4 presents the fully sampled and reconstructed FLAIR images using zero-filling, ESPIRiT, U-net, k-space learning, DeepCascade, and Joint-ICNet, with the reduction factor $R = 4$. The zero-filled image shows aliasing and blurring artifacts that obscured detailed brain structures. ESPIRiT, U-net, k-space learning, and DeepCascade methods removed those artifacts, but could not recover the detailed structures due to over-smoothing problems and still had residual aliasing artifacts. Compared to other methods, Joint-ICNet shows superior performance in reconstructing images and removing artifacts as shown on the magnified and difference im-

ages. Similar results are presented in Fig. 5 that shows fully sampled and reconstructed T2 images, with the reduction factor $R = 8$. With a high reduction factor, ESPIRiT method could not recover the image due to fewer ACS lines. U-net and DeepCascade show severe blurring artifacts and k-space learning shows residual aliasing artifacts.

The reconstructed results with abnormal cases are shown in Fig. 6, which presents fully sampled and reconstructed (a) T2, (b) T1POST, and (c) FLAIR images using zero-filling, ESPIRiT, U-net, k-space learning, DeepCascade, and Joint-ICNet, with the reduction factor $R = 4$. Whereas the conventional PI-based and other deep-learning-based methods could not recover the blurred abnormal tissues, which were obscured in the zero-filled images, Joint-ICNet recovered those similar to the fully sampled images, as shown on the magnified images.

The quantitative analysis of the reconstructed images is

Reduction Factor	Model	NRMSE				PSNR				SSIM			
		T1	T2	T1 POST	FLAIR	T1	T2	T1 POST	FLAIR	T1	T2	T1 POST	FLAIR
$R = 4$	Zero-filled	0.0421	0.0521	0.0410	0.0341	30.3	29.1	30.7	30.1	0.816	0.791	0.820	0.778
	ESPIRiT	0.0259	0.0299	0.0209	0.0367	32.5	31.8	33.7	29.9	0.653	0.645	0.717	0.602
	U-net	0.0075	0.0083	0.0069	0.0095	38.0	37.2	38.6	36.3	0.945	0.944	0.948	0.911
	k-space Learning	0.0104	0.0089	0.0070	0.0095	38.1	37.3	38.5	36.5	0.945	0.946	0.949	0.918
	DeepCascade	0.0075	0.0067	0.0076	0.0091	37.7	38.2	38.2	36.8	0.945	0.951	0.950	0.920
	Joint-ICNet	0.0035	0.0040	0.0037	0.0052	41.2	40.6	41.4	39.1	0.958	0.959	0.962	0.935
$R = 8$	Zero-filled	0.0992	0.1163	0.1060	0.0735	26.7	25.5	26.4	26.9	0.723	0.672	0.717	0.647
	ESPIRiT	0.0561	0.0828	0.0614	0.0697	29.2	27.2	28.8	27.2	0.598	0.553	0.631	0.506
	U-net	0.0118	0.0198	0.0134	0.0181	36.0	33.4	35.6	33.2	0.928	0.914	0.929	0.862
	k-space Learning	0.0129	0.0248	0.0160	0.0197	35.8	32.4	34.9	32.9	0.927	0.908	0.926	0.863
	DeepCascade	0.0134	0.0176	0.0167	0.0167	35.4	34.0	34.7	33.6	0.927	0.922	0.927	0.870
	Joint-ICNet	0.0063	0.0096	0.0075	0.0099	38.8	36.9	38.1	35.9	0.946	0.939	0.948	0.895

Table 1. Quantitative evaluation of reconstructed magnetic resonance images with two reduction factors.

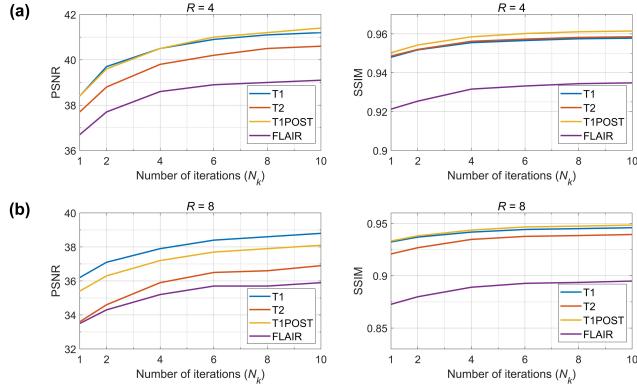


Figure 7. Quantitative results of Joint-ICNet with a different number of iterations (N_k).

presented in Table 1, which shows NRMSE, PSNR, and SSIM values of the reconstructed MR images including T1, T2, T1POST, and FLAIR using zero-filling, ESPIRiT, U-net, k-space learning, DeepCascade, and Joint-ICNet, with the reduction factors $R = 4$ and $R = 8$. With two reduction factors and four different MR images, Joint-ICNet had the lowest NRMSE, and the highest PSNR and SSIM values compared to other reconstruction methods.

Ablation Study of Joint-ICNet. Fig. 7 shows the quantitative results of Joint-ICNet with a different number of N_k with the reduction factors $R = 4$ and 8. As N_k increased, the PSNR and SSIM values increased in all MR images including T1, T2, T1POST, and FLAIR images for two reduction factors. Table 2 shows the quantitative results of the reconstructed images with different models of Joint-ICNet. In overall, the original Joint-ICNet shows the lowest NRMSE and the highest PSNR and SSIM values compared to other models.

Reconstruction of Coil Sensitivity Maps. Fig. 8 presents the reconstructed coil sensitivity maps using zero-filling and Joint-ICNet, with the reduction factor $R = 8$. With a high reduction factor, zero-filled coil sensitivity maps shows severe blurring and aliasing artifacts due to fewer ACS lines, but Joint-ICNet clearly removed those artifacts.

Reduction Factor	Model	NRMSE	PSNR	SSIM
$R = 4$	Joint-ICNet	0.0040	40.7	0.957
	Replace \mathcal{D}_C with ESPIRiT w/o \mathcal{D}_F	0.0054	39.4	0.949
	w/o DC layers	0.0041	40.5	0.956
$R = 8$	Joint-ICNet	0.0087	37.3	0.938
	Replace \mathcal{D}_C with ESPIRiT w/o \mathcal{D}_F	0.0120	35.8	0.927
	w/o DC layers	0.0093	37.0	0.937
		0.0116	36.0	0.932

Table 2. Quantitative evaluation of reconstructed magnetic resonance images using four different models of Joint-ICNet with two reduction factors.

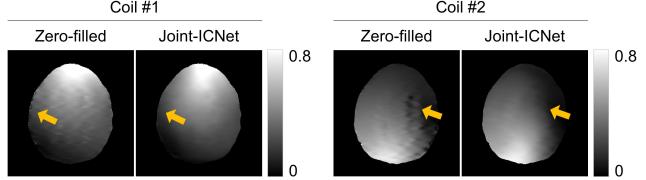


Figure 8. Reconstructed results of coil sensitivity maps using zero-filling and Joint-ICNet, with the reduction factor $R = 8$.

5. Conclusion

In this study, we proposed a joint deep model-based MR image and coil sensitivity reconstruction network, called Joint-ICNet, that jointly reconstructs an MR image and coil sensitivity maps from undersampled multi-coil k-space data. Experiments with various MR images and reduction factors demonstrated that our proposed Joint-ICNet showed superior performance compared to conventional PI-based and state-of-the-art deep-learning-based reconstruction methods in reconstructing the MR images. Joint-ICNet can be applied to various MRI applications to reduce scan time.

Acknowledgements. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (2019R1A2B5B01070488), Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Ministry of Science and ICT (NRF-2018M3A9H6081483), Y-BASE R&E Institute a Brain Korea 21, Yonsei University, and partially supported by the Graduate School of YONSEI University Research Scholarship Grants in 2017.

References

- [1] Hemant K Aggarwal, Merry P Mani, and Mathews Jacob. MoDL: Model-based deep learning architecture for inverse problems. *IEEE transactions on medical imaging*, 38(2):394–405, 2018.
- [2] Mehmet Akçakaya, Steen Moeller, Sebastian Weingärtner, and Kâmil Uğurbil. Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging. *Magnetic resonance in medicine*, 81(1):439–453, 2019.
- [3] Kai Tobias Block, Martin Uecker, and Jens Frahm. Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint. *Magnetic Resonance in Medicine*, 57(6):1086–1098, 2007.
- [4] Steven Diamond, Vincent Sitzmann, Felix Heide, and Gordon Wetzstein. Unrolled optimization with deep priors. *arXiv preprint arXiv:1705.08041*, 2017.
- [5] Taejoon Eo, Yohan Jun, Taeseong Kim, Jinseong Jang, Ho-Joon Lee, and Dosik Hwang. KIKI-net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images. *Magnetic resonance in medicine*, 80(5):2188–2201, 2018.
- [6] Taejoon Eo, Hyungseob Shin, Yohan Jun, Taeseong Kim, and Dosik Hwang. Accelerating Cartesian MRI by domain-transform manifold learning in phase-encoding direction. *Medical Image Analysis*, page 101689, 2020.
- [7] Jeffrey A Fessler. Model-based image reconstruction for MRI. *IEEE signal processing magazine*, 27(4):81–89, 2010.
- [8] Mark A Griswold, Peter M Jakob, Robin M Heidemann, Matthias Nittka, Vladimir Jellus, Jianmin Wang, Berthold Kiefer, and Axel Haase. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magnetic Resonance in Medicine*, 47(6):1202–1210, 2002.
- [9] Kerstin Hammernik, Teresa Klatzer, Erich Kobler, Michael P Recht, Daniel K Sodickson, Thomas Pock, and Florian Knoll. Learning a variational network for reconstruction of accelerated MRI data. *Magnetic resonance in medicine*, 79(6):3055–3071, 2018.
- [10] Yoseo Han, Leonard Sunwoo, and Jong Chul Ye. k -space deep learning for accelerated MRI. *IEEE transactions on medical imaging*, 39(2):377–386, 2019.
- [11] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [12] Yohan Jun, Taejoon Eo, Hyungseob Shin, Taeseong Kim, Ho-Joon Lee, and Dosik Hwang. Parallel imaging in time-of-flight magnetic resonance angiography using deep multi-stream convolutional neural networks. *Magnetic resonance in medicine*, 81(6):3840–3853, 2019.
- [13] Yohan Jun, Hyungseob Shin, Taejoon Eo, Taeseong Kim, and Dosik Hwang. Deep model-based magnetic resonance parameter mapping network (DOPAMINE) for fast T1 mapping using variable flip angle method. *Medical Image Analysis*, page 102017, 2021.
- [14] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [15] Florian Knoll, Kristian Bredies, Thomas Pock, and Rudolf Stollberger. Second order total generalized variation (TGV) for MRI. *Magnetic resonance in medicine*, 65(2):480–491, 2011.
- [16] Kuldeep Kulkarni, Suhas Lohit, Pavan Turaga, Ronan Keriven, and Amit Ashok. Reconnet: Non-iterative reconstruction of images from compressively sensed measurements. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 449–458, 2016.
- [17] Dongwook Lee, Jaejun Yoo, Sungho Tak, and Jong Chul Ye. Deep residual learning for accelerated MRI using magnitude and phase networks. *IEEE Transactions on Biomedical Engineering*, 65(9):1985–1995, 2018.
- [18] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 136–144, 2017.
- [19] Guilin Liu, Fitsum A Reda, Kevin J Shih, Ting-Chun Wang, Andrew Tao, and Bryan Catanzaro. Image inpainting for irregular holes using partial convolutions. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 85–100, 2018.
- [20] Michael Lustig, David Donoho, and John M Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 58(6):1182–1195, 2007.
- [21] Michael Lustig, David L Donoho, Juan M Santos, and John M Pauly. Compressed sensing MRI. *IEEE signal processing magazine*, 25(2):72–82, 2008.
- [22] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. In *Advances in neural information processing systems*, pages 8026–8037, 2019.
- [23] Klaas P Pruessmann, Markus Weiger, Markus B Scheidegger, and Peter Boesiger. SENSE: sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine*, 42(5):952–962, 1999.
- [24] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [25] Jo Schlemper, Jose Caballero, Joseph V Hajnal, Anthony N Price, and Daniel Rueckert. A deep cascade of convolutional neural networks for dynamic MR image reconstruction. *IEEE transactions on Medical Imaging*, 37(2):491–503, 2017.
- [26] Huajun She, Rong-Rong Chen, Dong Liang, Edward VR DiBella, and Leslie Ying. Sparse BLIP: BLind Iterative Parallel imaging reconstruction using compressed sensing. *Magnetic Resonance in Medicine*, 71(2):645–660, 2014.

- [27] Anuroop Sriram, Jure Zbontar, Tullie Murrell, Aaron Defazio, C Lawrence Zitnick, Nafissa Yakubova, Florian Knoll, and Patricia Johnson. End-to-end variational networks for accelerated MRI reconstruction. *arXiv preprint arXiv:2004.06688*, 2020.
- [28] Martin Uecker, Thorsten Hohage, Kai Tobias Block, and Jens Frahm. Image reconstruction by regularized nonlinear inversion—joint estimation of coil sensitivities and image content. *Magnetic Resonance in Medicine*, 60(3):674–682, 2008.
- [29] Martin Uecker, Peng Lai, Mark J Murphy, Patrick Virtue, Michael Elad, John M Pauly, Shreyas S Vasanawala, and Michael Lustig. ESPIRiT—an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA. *Magnetic resonance in medicine*, 71(3):990–1001, 2014.
- [30] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022*, 2016.
- [31] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.
- [32] Leslie Ying and Jinhua Sheng. Joint image reconstruction and sensitivity estimation in sense (JSENSE). *Magnetic Resonance in Medicine*, 57(6):1196–1202, 2007.
- [33] Jure Zbontar, Florian Knoll, Anuroop Sriram, Matthew J Muckley, Mary Bruno, Aaron Defazio, Marc Parente, Krzysztof J Geras, Joe Katsnelson, Hersh Chandarana, et al. fastMRI: An open dataset and benchmarks for accelerated MRI. *arXiv preprint arXiv:1811.08839*, 2018.
- [34] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, 2017.
- [35] Kai Zhang, Wangmeng Zuo, Shuhang Gu, and Lei Zhang. Learning deep CNN denoiser prior for image restoration. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3929–3938, 2017.
- [36] Bo Zhu, Jeremiah Z Liu, Stephen F Cauley, Bruce R Rosen, and Matthew S Rosen. Image reconstruction by domain-transform manifold learning. *Nature*, 555(7697):487–492, 2018.