DuDoRNet: Learning a Dual-Domain Recurrent Network for Fast MRI Reconstruction with Deep T1 Prior

Bo Zhou[1]

[1]Department of Biomedical Engineering,
Yale University

bo.zhou@yale.edu

S. Kevin Zhou[2,3]

[2]Chinese Academy of Sciences
[3]Peng Cheng Laboratory, Shenzhen

s.kevin.zhou@gmail.com

Abstract

Magnetic Resonance Imaging (MRI) with multiple protocols is commonly used for diagnosis, but it suffers from a long acquisition time, which yields the image quality vulnerable to say motion artifacts. To accelerate, various methods have been proposed to reconstruct full images from under-sampled k-space data. However, these algorithms are inadequate for two main reasons. Firstly, aliasing artifacts generated in the image domain are structural and non-local, so that sole image domain restoration is insufficient. Secondly, though MRI comprises multiple protocols during one exam, almost all previous studies only employ the reconstruction of an individual protocol using a highly distorted undersampled image as input, leaving the use of fully-sampled short protocol (say T1) as complementary information highly underexplored. In this work, we address the above two limitations by proposing a Dual Domain Recurrent Network (DuDoRNet) with deep T1 prior embedded to simultaneously recover k-space and images for accelerating the acquisition of MRI with a long imaging protocol. Specifically, a Dilated Residual Dense Network (DRDNet) is customized for dual domain restorations from undersampled MRI data. Extensive experiments on different sampling patterns and acceleration rates demonstrate that our method consistently outperforms state-of-the-art methods, and can reconstruct high quality MRI.

1. Introduction

MRI is one of the most applied imaging procedures for disease diagnosis and treatment planning. As a non-invasive, radiation-free, and in-vivo imaging modality, it provides significantly better soft-tissue contrast than many other imaging modalities and offers accurate measurements of both anatomical and functional signals. However, its long acquisition time, owing to sampling full k-space, especially under the multiple protocols requiring long Echo Time (TE)

and Repetition Time (TR), could lead to significant artifacts in the reconstructed image caused by patient or physiological motions during acquisitions, *i.e.* cardiac motion and respiration. Furthermore, a long acquisition time also limits the availability of MR scanner for patients and causes delayed patient care in the medical system.

To accelerate the MRI acquisition, various efforts have been made for improving the reconstruction image quality with undersampled k-space data. Previously, Compressed Sensing (CS) and Parallel Imaging have achieved significant progresses in fast MRI [20, 4, 8]. In CS-MRI, one assumes that images have a sparse representation in certain domains [16, 19, 12]. In conventional CS-MRI, previous works focus on using sparse coefficients in sparsifying transforms, *i.e.* wavelet transform [22, 32] and contourlet transform [23], combined with regularization parameters, *i.e.* total variation, to solve the ill-posed inverse problems in an iterative manner. However, these iterative minimization process based on sparsifying transforms tends to generate smaller sparse coefficient values, and lead to loss of details and unwanted artifacts in the reconstruction when undersampling rate is high [25]. Thus, current CS-MRI is limited to a rate of $2 \sim 3$ [16, 25]. Furthermore, reconstruction based on these iterative algorithms is time-consuming; thus it is challenging to deploy them in near real-time MRI scenarios, *i.e.* Cardiac-MRI and Functional-MRI. More recently, triggered by the success of computer vision [11], deep learning based algorithms have been developed for fast MRI reconstruction and demonstrated significant advantages [29, 27, 28, 5, 14, 15, 6, 18, 34, 35]. Wang *et al.* [29] first proposed to train a multi-layer Convolutional Neural Network (CNN) to recover the fully sampled MRI image from undersampled MRI image using supervised training with paired data. Similarly, Jin *et al.* [9] proposed to use U-Net [26] to solve the inverse problem in imaging. Yang *et al.* [31] developed a De-Aliasing GAN that uses a U-Net as the generator with a loss function consisting of four components: an image domain loss, a frequency domain loss, a perceptual loss, and an adversarial loss. Quan *et al.* [24]
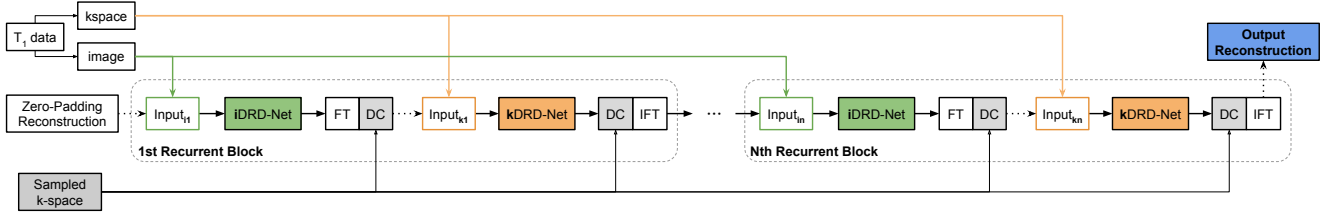
Figure 1: The architecture of our proposed **Du**al **Do**main **R**ecurrent **Net**work (**DuDoRNet**). DC denotes the data consistency layer. Each recurrent block contains one network for image domain restoration and one network for k-space domain restoration with two interleaved DC.

introduced a RefineGAN that uses a U-Net structure based generator with a cyclic loss for MRI de-aliasing. However, individual training is required for different sampling patterns and undersampling rates. Schlemper *et al*. [27, 21] came up with the data consistency layer in a deep cascade of CNN, which ensures the consistency of the reconstruction image in k-space and potentially reduces the issue of overfitting in deep training. Other than learning from pre-defined undersampling patterns, Zhang *et al*. [34] proposed to use active learning to determine the 1D Cartesian undersampling pattern on the fly during the MRI acquisition. These pioneering deep learning based methods surpass conventional CS algorithms owing to the high-nonlinearity properties of data-driven feature extraction and achieve significantly lower computation during run-time. However, existing deep learning based methods have three major limitations. In this work, we aim to break these limitations.

Firstly, the above-mentioned algorithms are principally learned in the image domain alone, with a few amendments that use the frequency domain in the loss design [31] or in the data consistency layer [27]. All of these deep learning based algorithms are designed to receive an image reconstructed from undersampled k-space as input and output an image as if reconstructed from fully sampled k-space; but unfortunately in the input images to the CNNs, detailed structures are likely distorted or even disappear. While there are recent attempts on restoring fully sampled image from undersampled k-space [3], the large scale trainable parameters limits the model can only be trained in an incremental fashion which could be further improved. As the first step to tackle this issue, we develop a dual domain learning scheme in MRI, which allows the network to restore the data in both image and frequency domains in a recurrent fashion. Previous work on CT metal artifact reduction also demonstrated the advantages of cross domain learning [13].

Secondly, the previous studies are limited to conventional network design, such as multi-layer CNN and U-Net, and there are few attempts to design a customized network structure for undersampled MRI reconstruction. Inspired by the recent development of super-resolution imaging techniques [2, 10, 33, 7], we propose a Dilated Residual Dense

Network (DRD-Net) with a Squeeze-and-Excitation Dilated Dense Residual Block (SDRDB) as the building module. The DRD-Net is used for both image and frequency domain restorations. Our SDRDB is customized for MRI reconstruction task. In fast MRI acquisitions, we observe significant sparsity in undersampled k-space, especially when undersampling rate is high. Previous studies have demonstrated that better k-space recovery can be achieved via using non-local information interpolation, such as GRAPPA [4]. Based on this assumption, the first motivation of our DRD Block design in k-space is synthesis of missing k-space from a large receptive field by utilizing non-local k-space data, thus bringing more robustness and reliability. In the image domain, human organ anatomy is correlated in different regions. Our DRD Block with a large receptive field in the image domain can better capture this correlation between anatomical regions and synthesize the missing anatomical information even when the signal is highly distorted.

Thirdly, previous studies have not fully explored the use of MRI protocol that requires a short acquisition time as deep prior to guide the MRI reconstruction process. In clinical routines, typical total scanning time for T1 Weighted Imaging (T1WI), T2 Weighted Imaging (T2WI), and Fluid-Attenuated Inversion Recovery (FLAIR) is $\sim$ 20 mins, in which T2WI and FLAIR take the majority due to their long TR and TE. However, using the undersampled image with the detailed structures may already disappeared as input, the existing methods could synthesize artificial structure that does not belong to the patient. Recently, Xiang *et al*. explored the merits of using T1WI as additional channel input in the image domain to aid T2WI reconstruction [30]. To further address this issue, we propose to use T1WI as deep prior in both image domain and k-space domain for improving the MRI reconstruction fidelity, given that the structural information in T1WI is highly correlated with that in different MRI protocols.

In summary, we propose a **Du**al **Do**main **R**ecurrent **Net**work (DuDoRNet) embedding with T1 priors to address these problems by learning two DRD-Nets on dual domains in a recurrent fashion to restore k-space and image domains

simultaneously. Our method (Figure 2) consists of three major parts: recurrent blocks comprised of image restoration network (iDRD-Net) and k-space restoration network (kDRD-Net), recurrent T1 priors embedded in image and k-space domains, and recurrent data consistency regularization. A correct reconstruction should ensure the consistency in both domains linked by the linear operation of Fourier Transform (FT). Our intuition is that image domain restoration can be enhanced by fusing signal back-propagated from the k-space restoration, vice versa. Given sparse signal in both domains, DRD-Net with a large receptive field can sense more signal for a better restoration. Our recurrent learning can better avoid overfitting in directly optimizing restoration networks in dual domains. Extensive experiments on MRI patients with different sampling patterns and acceleration rates demonstrate that our DuDoRNet generates superior reconstructions.

2. Problem Formulation

Denoting a 2D k-space with complex values as k, and a 2D image reconstructed from k as x, we need to reconstruct a fully sampled image x from both the undersampled k-space (k_u) and reconstructed image (x_u). The relationship between x and k can be written as:

$$k_u = M \odot k_f = M \odot \mathcal{F}(x_f), \quad (1)$$

$$x_u = \mathcal{F}^{-1}(k_u) = \mathcal{F}^{-1}(M \odot \mathcal{F}(x_f)), \quad (2)$$

where M represents the binary undersampling mask used for accelerations; k_f and k_u denote the fully sampled and the undersampled k-space, respectively; x_f and x_u denote the images reconstructed from fully sampled and undersampled k-space, respectively; $\odot$ is the element-wise multiplication operation; and $\mathcal{F}$ and $\mathcal{F}^{-1}$ are FT and Inverse Fourier Transform (IFT).

To tackle the ill-posed inverse problem of reconstructing x_f from limited sampled data, we propose to restore both image domain and k-space domain. In image domain restoration, the optimization can be expressed as minimizing the prediction error:

$$\arg\min_{\tilde{x}}||x_f - \tilde{x}||_2^2 = \arg\min_{\theta_x}||x_f - \mathcal{P}_x(x_u; \theta_x)||_2^2, \quad (3)$$

where $\tilde{x}$ is the prediction of the fully sampled image (x_f) generated by the estimation function ($\mathcal{P}_x$) with its parameters (θ_x) using the undersampled image (x_u) as input. Furthermore, the data consistency constraint [27] is often used in addition to the prediction error:

$$\arg\min_{\theta_x}(||x_f - \mathcal{P}_x(x_u; \theta_x)||_2^2 \quad (4a)$$

$$+ \lambda||k_u - M \odot \mathcal{F}(\mathcal{P}_x(x_u; \theta_x))||_2^2), \quad (4b)$$

where (4a) is the same as (3), and (4b) is the regularization term for data consistency. Similarly, we can formulate the

optimization target for k-space restoration as:

$$\arg\min_{\tilde{k}}||k_f - \tilde{k}||_2^2 = \arg\min_{\theta_k}||k_f - \mathcal{P}_k(k_u; \theta_k)||_2^2, \quad (5)$$

where $\tilde{k}$ is the prediction of the fully sampled k-space (k_f) generated by the estimation function ($\mathcal{P}_k$) with its parameters (θ_k) using the undersampled k-space (k_u) as input. Combining the image domain and k-space domain restoration with data consistency, the target function can thus be formulated as:

$$\arg\min_{\theta_k, \theta_x}(||k_f - \mathcal{P}_k(\mathcal{F}(\mathcal{P}_x(x_u; \theta_x)); \theta_k)||_2^2 \quad (6a)$$

$$+ ||x_f - \mathcal{P}_x(\mathcal{F}^{-1}(\mathcal{P}_k(k_u; \theta_k)); \theta_x)||_2^2 \quad (6b)$$

$$+ \lambda||k_u - M \odot \mathcal{F}(\mathcal{P}_x(\mathcal{F}^{-1}(\mathcal{P}_k(k_u; \theta_k)); \theta_x))||_2^2), \quad (6c)$$

Directly optimizing multiple terms in the above target function is challenging in traditional network design, owing to its high computational complexity, overfitting, and local optima problems. Thus, we propose a dual-domain recurrent learning strategy that optimizes θ_x and θ_k recurrently. Our proposed approach is illustrated in details in the following sections.

3. Methods

The overall pipeline of our network is illustrated in Figure 1. It consists of three major parts: 1) the dual domain recurrent network consisting of recurrent blocks with image and k-space restoration networks in it; 2) the deep prior information generated from T1 data in the image and k-space domains for feeding into the recurrent blocks of the network; and 3) the recurrent data consistency regulations using sampled k-space data. In each recurrent block, we propose to use a Dilated Residual Dense Network (DRD-Net) for both image and k-space restorations.

3.1. Dilated Residual Dense Network

Figure 2: The architecture of our **D**ilated **R**esidual **D**ense Network (**DRD-Net**) with building modules of SDRDB shown in Figure 3. The input can be either x_u in image domain or k_u in k-space domain. Convolution operation is followed by ReLU.

We develop a network structure for both MRI image and k-space restoration, called Dilated Residual Dense Network (DRD-Net). The idea is to use a Squeeze-and-excitation Dilated Residual Dense Block (SDRDB) as the backbone in our DRD-Net. The DRD-Net design and SDRDB are

shown in Figure 2 and Figure 3, respectively. Compared to RDN [33], we customize the local and global structure design for our MRI reconstruction task.

3.1.1 Global Structure

Our DRD-Net consists of three parts: initial feature extraction (IFE) via two sequential 3×3 convolution layers, multiple SDRDBs followed by global feature fusion, and global residual learning. The overall pipeline is as follow:

$$F_{-1} = \mathcal{P}_{IFE_1}(X_u), \qquad (7)$$

$$F_0 = \mathcal{P}_{IFE_2}(F_{-1}), \qquad (8)$$

where $\mathcal{P}_{IFE_1}$ and $\mathcal{P}_{IFE_2}$ denote the first and second convolutional operations in IFE, respectively. The first extracted feature F_{-1} is used for global residual learning in the third part. The second extracted feature F_0 is used as SDRDB input. If there are n SDRDBs, the n-th output F_n can be written as:

$$F_n = \mathcal{P}_{SDRDB_n}(F_{n-1}), \qquad (9)$$

where $\mathcal{P}_{SDRDB_n}$ represents the n-th SDRDB operation with $n \geq 1$. Given the extracted local features from a set of SDRDBs, we apply global feature fusion (GFF) to extract the global feature via:

$$F_{GF} = \mathcal{P}_{GFF}(\{F_1, F_2, \ldots, F_n\}), \qquad (10)$$

where $\{\}$ denotes the concatenation operation along feature channel. Our global feature fusion function $\mathcal{P}_{GFF}$ consists of a 1×1 and 3×3 convolution layers to fuse the extracted local features from different levels of SDRDB. The GFF output is used as input for global residual learning:

$$X_f = \mathcal{P}_{final}(F_{GF} + F_{-1}), \qquad (11)$$

The element-wise addition of global feature and initial feature are fed into our final 3×3 convolution layer for reconstruction output.

3.1.2 SDRD Block

The design of our SDRDB is shown in Figure 3. It contains four densely connected atrous convolution layers [1], local feature fusion, Squeeze-and-Excitation (SE) [7], and local residual learning. Expanding the expression of SDRDB, the t-th convolution output in n-th SDRDB can be written as:

$$F_n^t = \mathcal{H}_n^t \{F_{n-1}, F_n^1, \ldots, F_n^{t-1}\}, \qquad (12)$$

where $\mathcal{H}_n^t$ denotes the t-th convolution followed by Leaky-ReLU in the n-th SDRDB, $\{\}$ is the concatenation operation along feature channel, and the number of convolution is set to $t \leq 4$. Our SDRDB begins by composing a feature pyramid using 4 atrous convolution layers with dilation rates of 1, 2, 4, and 4. For an atrous convolution layer with kernel
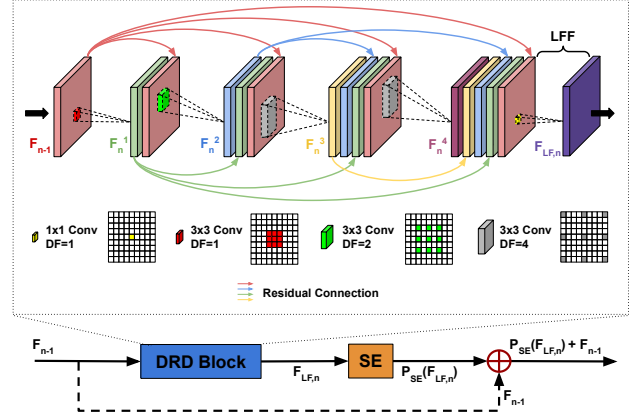
Figure 3: The structure of our **S**queeze-and-excitation **D**ilated **R**esidual **D**ense Block (**SDRDB**).

size (K) and kernel dilation (D), the receptive field (R) can be written as $R = K + (K - 1) \times (D - 1)$. The combination of two atrous convolution layers creates a new receptive field of $R_{comb} = R_1 + R_2 - 1$. With that being said, the dense connection over 4 atrous layers enables diverse combinations of layers with various receptive fields, which can more efficiently extract features from different scales than traditional dilation approaches. Figure 4 demonstrates the feature pyramid from our 4 densely connected atrous convolution layers.
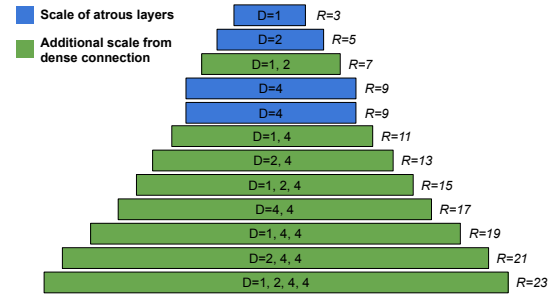
Figure 4: Illustration of the scale pyramid using our densely connected atrous layers with kernel size of 3×3 and dilation factors of 1,2,4, and 4. Our densely connection setting creates diverse kernel sizes with much larger receptive fields (blue + green) than naive sequential dilated layers (blue only). R denotes the size of receptive field created from the dilation (D) combination in each block.

Then, we apply our local feature fusion (LFF), consisting of 1×1 convolution layer and SE layer, to fuse the output from the last SDRDB and all convolution layers in current SDRDB. Thus, the LFF output can be expressed as:

$$F_{LF,n} = \mathcal{P}_{LFF,n}(\{F_{n-1}, F_n^1, F_n^2, F_n^3, F_n^4\}), \qquad (13)$$

where $\mathcal{P}_{LFF,n}$ denotes the LFF operation. Finally, we ap-

Table 1: Quantitative comparison of T2 reconstructions from different undersampling patterns and methods at R = 5. Best results with and without T1 prior are marked in red and blue, respectively.

Cartesian	ZP	GRAPPA[4]	TV[17]	Wang[29]	DeepCas[27]	RefGAN[24]	Ours w/o Prior	UF-T2 [30]	Ours
PSNR [dB]	22.498	23.318	23.026	26.061	26.993	25.848	27.834	30.594	32.511
SSIM	0.667	0.730	0.725	0.828	0.859	0.819	0.898	0.929	0.957
MSE($\times 10^2$)	0.622	0.508	0.555	0.306	0.221	0.340	0.178	0.097	0.066
Radial	ZP	GRAPPA	TV	Wang	DeepCas	RefGAN	Ours w/o Prior	UF-T2	Ours
PSNR [dB]	24.294	29.548	27.222	32.586	34.955	35.110	37.270	33.318	40.815
SSIM	0.581	0.822	0.727	0.889	0.957	0.959	0.974	0.939	0.981
MSE($\times 10^2$)	0.412	0.122	0.216	0.102	0.025	0.023	0.021	0.049	0.010
Spiral	ZP	GRAPPA	TV	Wang	DeepCas	RefGAN	Ours w/o Prior	UF-T2	Ours
PSNR [dB]	26.181	31.893	33.890	35.987	43.867	36.049	48.418	37.793	49.186
SSIM	0.675	0.882	0.909	0.932	0.972	0.960	0.991	0.961	0.993
MSE($\times 10^2$)	0.266	0.076	0.046	0.029	0.0046	0.024	0.0018	0.019	0.0014

ply SE and the local residual learning to LFF output by adding the residual connection from SDRDB input. Thus, the SDRDB output is:

$$F_n = \mathcal{P}_{SE}(F_{LF,n}) + F_{n-1}, \quad (14)$$

3.2. Dual-Domain Recurrent Learning

In this section, we present details about our proposed DuDoRNet. As illustrated in Figure 1, each of the recurrent blocks of DuDoRNet contains one image restoration DRD-Net (iDRD-Net), one k-space restoration DRD-Net (kDRD-Net), and two data consistency layers interleaved. The T1 deep prior provides data in both image space and k-space that are fed into DuDoRNet recurrent blocks. In the n-th recurrent block, denoting the image input as x_{u_n}, the image restoration optimization target can be written as:

$$\arg\min_{\theta_{iDRD}}(||x_f - \mathcal{P}_{iDRD}(x_{u_n}, x_{T1}; \theta_{iDRD})||_2^2 \quad (15a)$$
$$+\lambda||k_{u_n} - M \odot \mathcal{F}(\mathcal{P}_{iDRD}(x_{u_n}, x_{T1}; \theta_{iDRD}))||_2^2),$$

where $\mathcal{P}_{iDRD}$ is the image restoration network based on DRD-Net with parameters θ_{iDRD}. The image output from the last recurrent block x_{u_n} and T1 image prior x_{T1} are concatenated channel-wise for inputting into iDRD-Net. M is the binary undersampling mask used for accelerations. The k-space values in M can be altered after inference through iDRD-Net since it only optimizes the first term. To maintain the k-space fidelity at sampled locations z of M, we add the data consistency as the second term. Denoting the output of iDRD-Net as x_{iDRD_n} and its FT as $k_{iDRD_n} = \mathcal{F}(x_{iDRD_n})$, the corresponding output from data consistency layer can be thus formulated as:

$$k_{iDRD_n} = \begin{cases} \frac{\lambda k_{iDRD_n}(z) + k_{u_n}(z)}{\lambda+1} & \text{if } M(z) = 1 \\ k_{iDRD_n}(z) & \text{if } M(z) = 0 \end{cases} \quad (16)$$

where λ controls the level of linear combination between sampled k-space values and predicted values. When $\lambda = 0$, the sampled k-space directly substitutes the prediction at z

in k-space. Denoting this output as $k_{u_n} = k_{iDRD_n}$, then k_{u_n} and T1 k-space prior k_{T1} are concatenated channelwise to feed into the kDRD-Net for k-space restoration. Similarly, the k-space restoration optimization target can be written as:

$$\arg\min_{\theta_{kDRD}}(||k_f - \mathcal{P}_{kDRD}(k_{u_n}, k_{T1}; \theta_{kDRD})||_2^2 \quad (17a)$$
$$+\lambda||k_{u_n} - M \odot \mathcal{P}_{kDRD}(k_{u_n}, k_{T1}; \theta_{kDRD}))||_2^2),$$

where $\mathcal{P}_{kDRD}$ is the k-space restoration network based on DRD-Net with network parameters θ_{kDRD}. Similarly here, the second term is to ensure the data consistency in the restored k-space. Thus, the loss function for each recurrent block is $\mathcal{L}_{i_n} + \mathcal{L}_{k_n}$. The final loss is the summation of recurrent blocks losses: $\sum_{n=1}^{N_{rec}}(\mathcal{L}_{i_n} + \mathcal{L}_{k_n})$, where N_{rec} denotes the number of recurrent blocks.

In our experiments, λ is set to 0.01, N_{rec} is set to 5, and the number of SDRDB is set to 3. Each recurrent block in our DuDoRNet share the same network parameters. The final reconstruction output during testing is obtained by applying IFT to the last kDRD-Net output after data consistency operation.

4. Experiments

4.1. Experimental Settings

Dataset and Training We acquired an in-house MRI dataset consisting of 20 patients. We scanned each patient using three different protocols (T1, T2, and FLAIR) with full k-space sampled, resulting in three 3D volumes of $320 \times 230 \times 18$ for each patient in both image and k-space domains. 360 2D images are generated for each protocol. We split the dataset patient-wise into training/validation/test sets with a ratio of $7 : 1 : 2$. As a result, our dataset consists of 252 training images, 36 validation images, and 72 test images for each protocol. Three different k-space undersampling patterns are examined in our experiments, including Cartesian, radial, and spiral trajectories. Examples
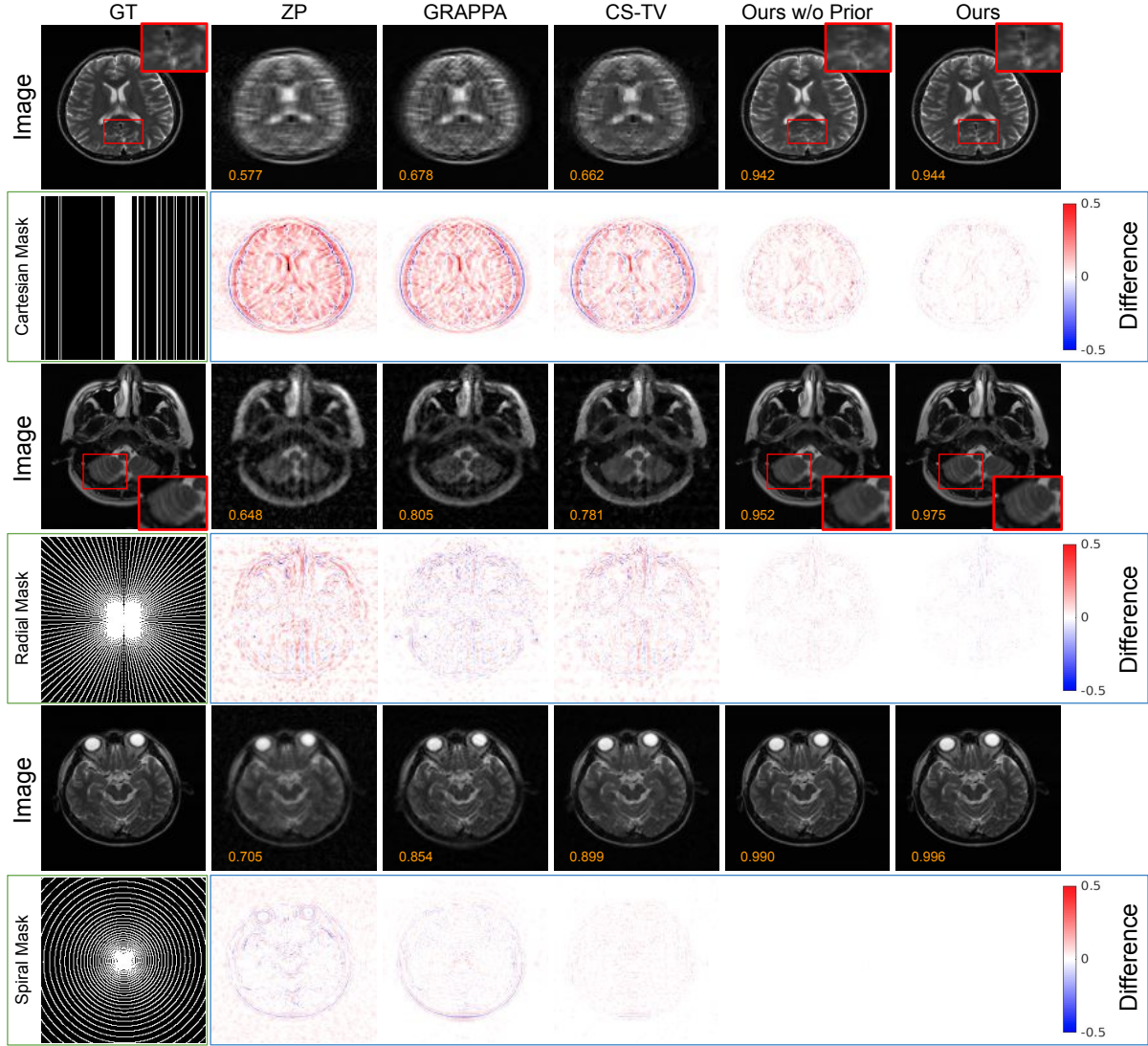
Figure 5: Reconstruction results from Cartesian/Radial/Spiral trajectory at under-sampling rate R = 5. The sampling pattern mask and difference images are shown on the second, fourth, and sixth row. Red boxes illustrate the enlarged views on details. The SSIM is indicated on the bottom left of the image.

of the sampling patterns are illustrated in Figure 5 (green box). The acceleration factor (R) is set to a value between 1 and 6 for all three patterns, corresponding to acceleration in acquisition time. Using these patterns, randomly undersampled T2/FLAIR and fully-sampled T1 are used as input to our model for training. During training, we also randomly augment the image data by flipping horizontally or vertically, and by rotating at different angles.

Performance Evaluation The final evaluation is performed on our 72 test images. For quantitative evaluation, we evaluate our image reconstruction results using Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Mean Square Error (MSE). For comparative study, we

compared our results with the following algorithms in previous works: total variation with penalized compress-sensing (TV-CS) [16], GRAPPA [4], and four CNN-based algorithms including a sequential convolutional model [29], a cascade image restoration model called DeepCas [27], a structural refinement GAN model called RefGAN [24], and a DenseUNet model with complementary T1 image called UF-T2 [30]. All CNN-based methods are trained using the same settings as to our method.

4.2. Results

Image Quality Evaluation and Comparison We evaluated our reconstruction results by computing their SSIM,
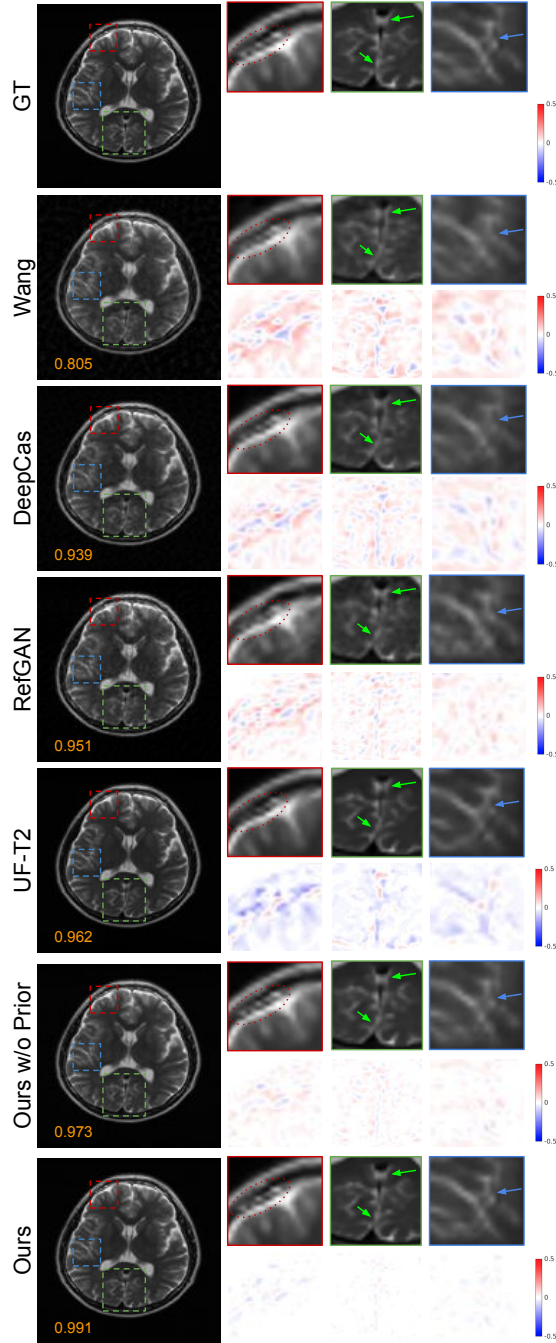
Figure 6: Comparison of reconstructions using radial trajectory at an acceleration rate $R = 5$. Three enlarged sub-regions and corresponding difference images are shown on the right. SSIM is indicated on the bottom-left.

PSNR, and MSE, where the fully sampled images are used as ground-truth for this calculations. In Table 1, we demonstrated our T2 reconstruction evaluations on three sampling patterns with a high acceleration rate of $R = 5$. The first

sub-table summarized the image quality when Cartesian-based acceleration is applied. Our DuDoRNet without T1 prior achieved PSNR up to 27.834 dB, and boosted to 32.511 dB when T1 prior is given. As compared to the previous state-of-the-art method with T1 prior called UF-T2 [30], we improved the reconstruction from 30.594 dB to 32.511 dB. Similarly for radial-based acceleration, our DuDoRNet without T1 prior achieved PSNR up to 37.27 dB, and further boosted to 40.815 dB when T1 prior is provided. As compared to the best results with PSNR = 35.11 from RefGAN [24] without T1 prior, our DuDoRNet achieves significantly better results. In the last sub-table, we found our DuDoRNet with spiral pattern yields the best image quality among all sampling patterns and reconstruction methods. Under the spiral pattern, our DuDoRNet without T1 prior achieved PSNR = 48.418 dB and reinforced to 49.186 dB when T1 prior is present. The qualitative comparison results with non-CNN based methods are shown in Figure 5. As we can see, the reconstructions with zero padding (ZP) at a high acceleration rate create significant aliasing artifacts and lose anatomical details. Non-CNN based methods can improve the reconstruction as compared to ZP, but it is hard to see a significant improvement when a significant level of aliasing artifact is presented. In comparison, our reconstructions are robust to aliasing artifacts and structural loss in the input. The qualitative comparison results with CNN based methods are shown in Figure 6. At a high acceleration rate, the CNN based methods achieve better results than non-CNN based methods. Among them, our method restores information in both image and k-space better preserved the important anatomical details, as demonstrated in Figure 6 with arrows and ellipses. More results on FLAIR reconstruction are summarized in supplemental materials with similar performances.

Reconstruction Stability Evaluation To evaluate the performance stability at different acceleration rates, we recorded the reconstruction performance by varying R from 2 to 6 on all three patterns. The evaluation results are summarized in Figure 7. For the Cartesian pattern, our method can consistently maintain the SSIM to above 0.95 even when an aggressive acceleration rate is used, *i.e.* $R = 6$. Due to the large aliasing artifact created from the Cartesian sampling pattern as the acceleration rate increases, the reconstruction performance is more challenging to remain stable without T1 prior, but we were still able to maintain SSIM to above 0.89 and consistently outperforms the other methods for all undersampling rates. For radial pattern, all methods performed approximately the same at low acceleration rates $R < 4$. However, for a more aggressive under-sampling rate, *i.e.* $R > 4$, our DuDoRNet is able to reduce the structural loss by a considerable margin. At $R = 6$, Our DuDoRNet with and without T1 prior can consistently maintain the SSIM to above 0.98 and 0.97, respectively.
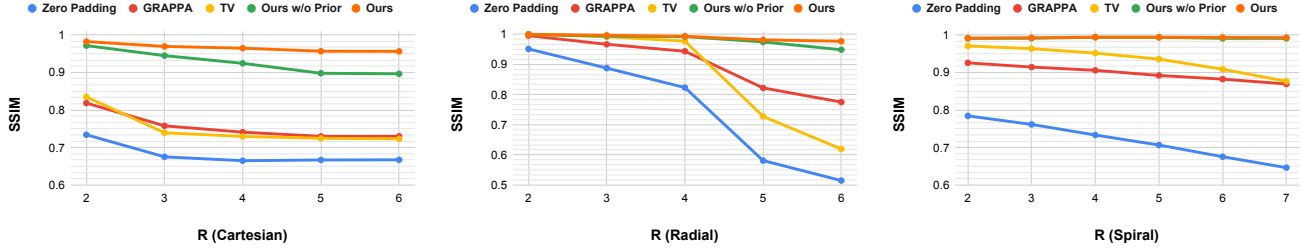
Figure 7: Comparison of reconstruction performance at different sampling rate using Cartesian, radial, and spiral patterns.

Lastly, we found the best performance stability when the spiral pattern is applied. Our DuDoRNet keeps the SSIM to above 0.99 over the whole undersampling range regardless of T1 prior. Under the same acceleration rate, radial and spiral patterns sample the k-space more uniformly than random Cartesian, thus leading to less aliasing artifact in the initial reconstruction input for the models, as demonstrated in Figure 5. As a result, radial and spiral patterns generate less aliasing artifacted input and are able to output more stable reconstruction in our experiments.

4.3. Ablation Studies

Table 2: Quantitative evaluations for Rec, DD, and DIL components in DuDoRNet. For the average, we use † to indicate if difference were statistically significant ($p < 0.05$) compared to (A).

SSIM	Cartesian	Radial	Spiral	Average
(A) Net-baseline	0.839	0.902	0.951	0.897
(B) Net-Rec	0.860	0.959	0.973	0.931†
(C) Net-DD	0.851	0.929	0.962	0.914†
(D) Net-DIL	0.844	0.909	0.956	0.903
(E) Net-Rec-DD	0.891	0.968	0.989	0.949†
(F) Net-Rec-DIL	0.869	0.962	0.978	0.936†
(G) Net-DD-DIL	0.859	0.935	0.969	0.921†
(H) Net-Rec-DD-DIL	0.898	0.974	0.991	0.954†

Firstly, we evaluated the effectiveness of different components in our DuDoRNet. Without loss of generality, the reconstruction performance is evaluated on $R = 5$ for all three patterns. We evaluate three key components, including: dual domain learning (DD), dilated residual dense learning (DIL), and recurrent learning (Rec) without T1 prior. N_{rec} in Rec is set to 5. The component analysis is summarized in Table 2. As we can observe, recurrent learning (B) and dual domain learning (C) each improve the performance over the baseline (A) by 0.015, which are more significant than dilated residual dense learning (D). Combining recurrent learning and dual domain learning (E), the performance achieves the largest boost as compared to the other two component combinations (F and G). Our DuDoRNet, equipping all components (H), produces the best

reconstruction results. Overall, all three components help DuDoRNet to enhance the performance.
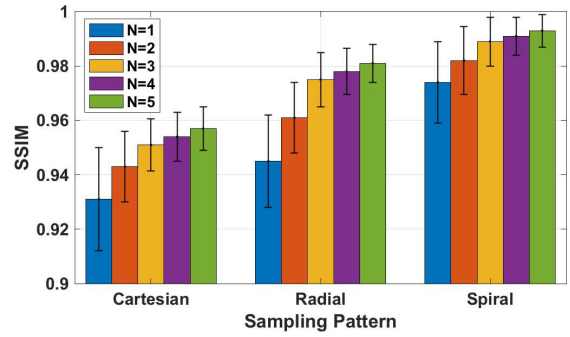

Figure 8: The effect of increasing the number of recurrent blocks (n) in our DuDoRNet under three sampling patterns.

Secondly, we evaluated the effect of N_{rec} in our DuDoR-Net. As shown in Figure 8, the reconstruction performance, measured by SSIM, increases monotonically when N_{rec} increases, while the rate of improvement starts to converge after $N_{rec} = 3$ for all three patterns. We also found that our DuDoRNet achieves the best performance when the spiral pattern is used even when different N_{rec} is implemented.

5. Conclusion

We present a dual domain recurrent network for fast MRI reconstruction with T1 prior embedded. Specifically, we propose to restore both image and k-space domains recurrently through DRD-Nets with large receptive fields. The T1 prior is embedded at each recurrent block to deeply guide restorations for both domains. Extensive experimental results demonstrate that while previous fast MRI methods on single domain for individual protocol have limited capability of directly reducing aliasing artifacts in the image domain, our DuDoRNet can efficiently restore the reconstruction, and the T1 prior can further significantly improve the structural recovery. Future work includes exploring DuDoRNet on pathological MR data and the application to other signal recovery tasks, such as noise reduction and super resolution.

References

[1] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4):834–848, 2017. 4

[2] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2):295–307, 2015. 2

[3] 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. 2

[4] Mark A Griswold, Peter M Jakob, Robin M Heidemann, Mathias Nittka, Vladimir Jellus, Jianmin Wang, Berthold Kiefer, and Axel Haase. Generalized autocalibrating partially parallel acquisitions (grappa). *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 47(6):1202–1210, 2002. 1, 2, 5, 6

[5] 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. 1

[6] Yoseob Han, Leonard Sunwoo, and Jong Chul Ye. k-space deep learning for accelerated mri. *IEEE transactions on medical imaging*, 2019. 1

[7] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018. 2, 4

[8] Feng Huang, James Akao, Sathya Vijayakumar, George R Duensing, and Mark Limkeman. k-t grappa: A k-space implementation for dynamic mri with high reduction factor. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 54(5):1172–1184, 2005. 1

[9] Kyong Hwan Jin, Michael T McCann, Emmanuel Froustey, and Michael Unser. Deep convolutional neural network for inverse problems in imaging. *IEEE Transactions on Image Processing*, 26(9):4509–4522, 2017. 1

[10] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Accurate image super-resolution using very deep convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1646–1654, 2016. 2

[11] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015. 1

[12] Dong Liang, Bo Liu, Jiunjie Wang, and Leslie Ying. Accelerating sense using compressed sensing. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 62(6):1574–1584, 2009. 1

[13] Wei-An Lin, Haofu Liao, Cheng Peng, Xiaohang Sun, Jingdan Zhang, Jiebo Luo, Rama Chellappa, and Shaohua Kevin Zhou. Dudonet: Dual domain network for ct metal artifact reduction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 10512–10521, 2019. 2

[14] Risheng Liu, Yuxi Zhang, Shichao Cheng, Xin Fan, and Zhongxuan Luo. A theoretically guaranteed deep optimization framework for robust compressive sensing mri. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 4368–4375, 2019. 1

[15] Kai Lønning, Patrick Putzky, Jan-Jakob Sonke, Liesbeth Reneman, Matthan WA Caan, and Max Welling. Recurrent inference machines for reconstructing heterogeneous mri data. *Medical image analysis*, 53:64–78, 2019. 1

[16] Michael Lustig, David Donoho, and John M Pauly. Sparse mri: The application of compressed sensing for rapid mr imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 58(6):1182–1195, 2007. 1, 6

[17] Shiqian Ma, Wotao Yin, Yin Zhang, and Amit Chakraborty. An efficient algorithm for compressed mr imaging using total variation and wavelets. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2008. 5

[18] Morteza Mardani, Enhao Gong, Joseph Y Cheng, Shreyas S Vasanawala, Greg Zaharchuk, Lei Xing, and John M Pauly. Deep generative adversarial neural networks for compressive sensing mri. *IEEE transactions on medical imaging*, 38(1):167–179, 2018. 1

[19] Mark Murphy, Marcus Alley, James Demmel, Kurt Keutzer, Shreyas Vasanawala, and Michael Lustig. Fast spirit compressed sensing parallel imaging mri: Scalable parallel implementation and clinically feasible runtime. *IEEE transactions on medical imaging*, 31(6):1250–1262, 2012. 1

[20] Ricardo Otazo, Daniel Kim, Leon Axel, and Daniel K Sodickson. Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion mri. *Magnetic resonance in medicine*, 64(3):767–776, 2010. 1

[21] Chen Qin, Jo Schlemper, Jose Caballero, Anthony N Price, Joseph V Hajnal, and Daniel Rueckert. Convolutional recurrent neural networks for dynamic mr image reconstruction. *IEEE transactions on medical imaging*, 38(1):280–290, 2018. 2

[22] Xiaobo Qu, Xue Cao, Di Guo, Changwei Hu, and Zhong Chen. Combined sparsifying transforms for compressed sensing mri. *Electronics letters*, 46(2):121–123, 2010. 1

[23] Xiaobo Qu, Weiru Zhang, Di Guo, Congbo Cai, Shuhui Cai, and Zhong Chen. Iterative thresholding compressed sensing mri based on contourlet transform. *Inverse Problems in Science and Engineering*, 18(6):737–758, 2010. 1

[24] Tran Minh Quan, Thanh Nguyen-Duc, and Won-Ki Jeong. Compressed sensing mri reconstruction using a generative adversarial network with a cyclic loss. *IEEE transactions on medical imaging*, 37(6):1488–1497, 2018. 1, 5, 6, 7

[25] Saiprasad Ravishankar and Yoram Bresler. Mr image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE transactions on medical imaging*, 30(5):1028–1041, 2010. 1

[26] 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. 1

[27] 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. 1, 2, 3, 5, 6

[28] Jian Sun, Huibin Li, Zongben Xu, et al. Deep admm-net for compressive sensing mri. In *Advances in neural information processing systems*, pages 10–18, 2016. 1

[29] Shanshan Wang, Zhenghang Su, Leslie Ying, Xi Peng, Shun Zhu, Feng Liang, Dagan Feng, and Dong Liang. Accelerating magnetic resonance imaging via deep learning. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 514–517. IEEE, 2016. 1, 5, 6

[30] Lei Xiang, Yong Chen, Weitang Chang, Yiqiang Zhan, Weili Lin, Qian Wang, and Dinggang Shen. Ultra-fast t2-weighted mr reconstruction using complementary t1-weighted information. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 215–223. Springer, 2018. 2, 5, 6, 7

[31] Guang Yang, Simiao Yu, Hao Dong, Greg Slabaugh, Pier Luigi Dragotti, Xujiong Ye, Fangde Liu, Simon Arridge, Jennifer Keegan, Yike Guo, et al. Dagan: deep de-aliasing generative adversarial networks for fast compressed sensing mri reconstruction. *IEEE transactions on medical imaging*, 37(6):1310–1321, 2017. 1, 2

[32] Yudong Zhang, Zhengchao Dong, Preetha Phillips, Shuihua Wang, Genlin Ji, and Jiquan Yang. Exponential wavelet iterative shrinkage thresholding algorithm for compressed sensing magnetic resonance imaging. *Information Sciences*, 322:115–132, 2015. 1

[33] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu. Residual dense network for image super-resolution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2472–2481, 2018. 2, 4

[34] Zizhao Zhang, Adriana Romero, Matthew J Muckley, Pascal Vincent, Lin Yang, and Michal Drozdzal. Reducing uncertainty in undersampled mri reconstruction with active acquisition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2049–2058, 2019. 1, 2

[35] 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, 2018. 1