Nonuniform Variational Network: Deep Learning for Accelerated Nonuniform MR Image Reconstruction

Jo Schlemper[†12], Seyed Sadegh Mohseni Salehi[†1], Prantik Kundu[1], Carole Lazarus[1], Hadrien Dyvorne[1], Daniel Rueckert[2], and Michal Sofka[1]

[1] Hyperfine Research, CT, USA
[2] Biomedical Image Analysis Group, Imperial College London, UK

Abstract. Deep learning for accelerated magnetic resonance (MR) image reconstruction is a fast growing field, which has so far shown promising results. However, most works are limited in the sense that they assume equidistant rectilinear (Cartesian) data acquisition in 2D or 3D. In practice, a reconstruction from nonuniform samplings such as radial and spiral is an attractive choice for more efficient acquisitions. Nevertheless, it has less been explored as the reconstruction process is complicated by the necessity to handle non-Cartesian samples. In this work, we present a novel approach for reconstructing from nonuniform undersampled MR data. The proposed approach, termed *nonuniform variational network* (NVN), is a convolutional neural network architecture based on the unrolling of a traditional iterative nonlinear reconstruction, where the knowledge of the nonuniform forward and adjoint sampling operators are efficiently incorporated. Our extensive evaluation shows that the proposed method outperforms existing state-of-the-art deep learning methods, hence offering a method that is widely applicable to different imaging protocols for both research and clinical deployments.

1 Introduction

Magnetic resonance imaging (MRI) has a fundamentally slow acquisition speed due to underlying physical and physiological constraints. Slow acquisitions result in reduced patient comfort as well as degraded image quality, where the latter is largely due to patient motion and a long-term accumulation of system imperfections. As such, accelerating the data acquisition has become an important research topic in the last decades. In MRI, an image is acquired indirectly through its Fourier transform, referred to as the *k-space*. Currently, a typical MR imaging technique employs rectilinear (Cartesian) sampling on a uniform k-space grid. Cartesian sampling is attractive due to its simplicity; once the k-space is acquired at Nyquist sampling rate, the image can be obtained by applying an inverse Fourier transform. However, more efficient 2D or 3D sampling patterns

[†]Co-first authors. Emails: {jschlemper,ssalehi}@hyperfine-research.com

may be designed to speed up MR acquisition time. For example, alternative trajectories include radial [8], spiral [1] and variable density [9] as well as optimized sampling patterns [10]. Non-Cartesian sampling patterns are attractive due to their motion robustness [14, 3]. In addition, when undersampled, non-Cartesian sampling patterns are particularly suitable for compressed sensing (CS) [11] and deep learning (DL) based approaches as the aliasing patterns show higher incoherence than Cartesian sampling, which can potentially offer faster acceleration factors.

Recently, deep learning based approaches for accelerated MR image reconstruction have gained popularity due to their promising performances, often exceeding the quality of the CS approaches [5, 15, 12, 17]. However, many of the novel approaches have so far been developed for Cartesian sampling trajectories and only a handful of work has considered nonuniform MR data [18, 6]. In this work, we present a convolutional neural network (CNN) architecture based on the generalization of variational network [5] for non-Cartesian cases, termed *nonuniform variational network* (NVN). Similar to variational networks, the proposed method is based on the *unrolling* of iterative nonlinear reconstruction used in CS, however, we generalize the regularization functional as well as adapting the data fidelity term for nonuniform case. The extensive evaluation reveals that the proposed algorithm is superior over existing deep learning approaches for nonuniform data.

2 Problem Formulation

Let $\boldsymbol{x} \in \mathbb{C}^N$ denote a complex-valued MR image to be reconstructed, represented as a vector with $N = N_x N_y$ where N_x and N_y are width and height of the image. Let $\boldsymbol{y} \in \mathbb{C}^M$ ($M << N$) represent the undersampled k-space measurements. Our problem is to reconstruct $\boldsymbol{x}$ from $\boldsymbol{y}$, formulated as an unconstrained optimization:

$$\underset{\boldsymbol{x}}{\operatorname{argmin}} \quad \frac{\lambda}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \mathcal{R}(\boldsymbol{x}) \tag{1}$$

Here $\boldsymbol{A}$ is a nonuniform Fourier sampling operator, $\mathcal{R}$ expresses regularization terms on $\boldsymbol{x}$ and λ is a hyper-parameter often associated to the noise level. In Cartesian case, $\boldsymbol{A} = \boldsymbol{M}\boldsymbol{F}$ where $\boldsymbol{M}$ is a sampling mask, $\boldsymbol{F}$ is discrete Fourier transform. In non-Cartesian case, the measurements no longer fall on a uniform k-space grid and hence generalization is required. In essence, the sampling operator $\boldsymbol{A}$ is now given by the nonuniform discrete Fourier transform of type I *(NUDFT-I)*:

$$\boldsymbol{y}\left((k_x, k_y)\right) = \sum_{l=0}^{N_x} \sum_{m=0}^{N_y} \boldsymbol{x}_{lm} e^{2\pi i (\frac{l}{N_x} k_x + \frac{m}{N_y} k_y)} \tag{2}$$

where $(k_x, k_y) \in \mathbb{R}^2$ (rather than $(k_x, k_y) \in \mathbb{Z}^2$). An efficient implementation of the above forward model exists, which is called nonuniform Fast Fourier

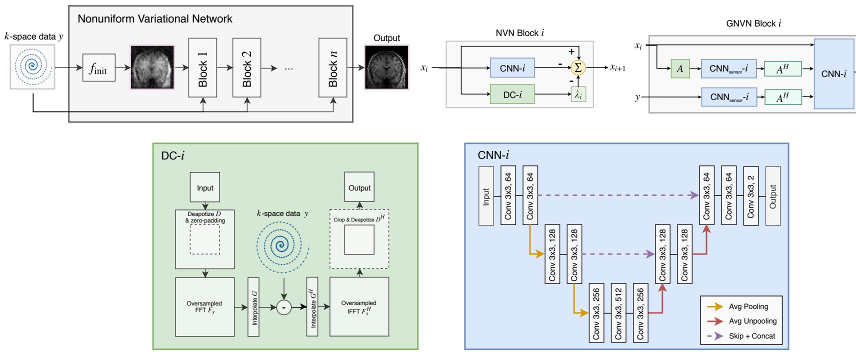

Fig. 1. The architecture of Nonuniform Variational Network.

Transform (NUFFT) [2, 4]. The idea is to approximate Eq. 2 by the following decomposition: $\boldsymbol{A} = \boldsymbol{G}\boldsymbol{F}_s\boldsymbol{D}$, where $\boldsymbol{G}$ is a gridding interpolation kernel, $\boldsymbol{F}_s$ is fast Fourier transform (FFT) with an oversampling factor s and $\boldsymbol{D}$ is a de-apodization weights. In contrast, the inversion of $\boldsymbol{A}$ is considerably more involving. For the (approximately) fully-sampled case, one can consider direct inversion ($\mathcal{O}(N^3)$) or a more computationally efficient *gridding* reconstruction, which has the form $\boldsymbol{x}_{\text{gridding}} = \boldsymbol{A}^H\boldsymbol{W}\boldsymbol{y}$, where $\boldsymbol{W}$ is a diagonal matrix used for the density compensation of non-uniformly spaced measurements. For the undersampled case, the inversion is ill-posed, and one requires solving Eq. 1 by iterative algorithms.

3 Proposed Approach

In this work, we propose a new deep learning algorithm to approximate the solution to the optimization problem in Eq. 1. Firstly, we consider a gradient descent algorithm which provides a locally optimal solution:

$$\boldsymbol{x}_0 = f_{\text{init}}(\boldsymbol{A}, \boldsymbol{y}) \tag{3}$$

$$\boldsymbol{x}_i = \boldsymbol{x}_{i-1} - \alpha_i \nabla_{\boldsymbol{x}} f(\boldsymbol{x})|_{\boldsymbol{x}=\boldsymbol{x}_{i-1}} \tag{4}$$

where f_{init} is an initializer, α_i is a step size and ∇f is the gradient of the objective function, which is given by:

$$\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) = \lambda \boldsymbol{A}^H(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}) + \nabla_{\boldsymbol{x}}\mathcal{R}(\boldsymbol{x}) \tag{5}$$

Common choices of the initializer are adjoint $f_{\text{init}}(\boldsymbol{A}, \boldsymbol{y}) = \boldsymbol{A}^H\boldsymbol{y}$ and gridding reconstruction $f_{\text{init}}(\boldsymbol{A}, \boldsymbol{y}) = \boldsymbol{A}^H\boldsymbol{W}\boldsymbol{y}$. The idea of variational network (VN) formalism is to first unroll the sequential updates into a feed-forward model, and approximate the gradient term $\nabla\mathcal{R}$ by a series of trainable convolution layers and non-linearities. Such generalization yields an end-to-end trainable network with N_{it} blocks:

$$\boldsymbol{x}_0 = f_{\text{init-cnn}}(\boldsymbol{A}, \boldsymbol{y}|\theta_0) \tag{6}$$

$$\boldsymbol{x}_i = \boldsymbol{x}_{i-1} - \lambda_i \underbrace{\boldsymbol{A}^H(\boldsymbol{A}\boldsymbol{x}_{i-1} - \boldsymbol{y})}_{\text{DC-}i} - \underbrace{f_{\text{cnn}}(\boldsymbol{x}_{i-1}|\theta_i)}_{\text{CNN-}i} \tag{7}$$

where the learnable parameters are $\{\theta_0, \ldots, \theta_{N_{\text{it}}}, \lambda_1, \ldots, \lambda_{N_{\text{it}}}\}$. Note that the step size α_i is absorbed in the learnable parameters. The key difference between the proposed network and the original formulation [5] is the following: firstly, in this work, a general non-convex regularization functional is used in place of the explicit convex regularization based on Fields-of-Experts model [16], which can be approximated by state-of-the-art CNN denoisers [13]. Secondly, the complexity of the methodology is in the implementation of $\boldsymbol{A} \in \mathbb{C}^{M\times N}$. In contrast to the Cartesian case, $\boldsymbol{A}$ is expressed as $\boldsymbol{G}\boldsymbol{F}_s\boldsymbol{D}$. For 2D cases, this can be a large matrix, which consumes a large portion of GPU memory.[3] To overcome this challenge, $\boldsymbol{G}$ is implemented as a sparse GPU matrix multiplication. $\boldsymbol{F}_s$ is a FFT, where an efficient GPU implementation is available. Finally, $\boldsymbol{D}$ is a diagonal matrix, which can be implemented as a Hadamard product of matrices. The adjoint can similarly be implemented as $\boldsymbol{A}^H = \boldsymbol{D}^H\boldsymbol{F}_s^H\boldsymbol{G}^H$, where $.^H$ is a conjugate transpose. The detailed steps involved in $\boldsymbol{A}$ and the proposed architecture, termed *nonuniform variational network* (NVN), is shown in Fig. 1, where the data consistency term is shown in *DC-i* block and CNN is shown in *CNN-i* block.

Generalized Nonuniform Variational Network Observe that network forces the bottleneck at the end of each iteration. However, an alternative view is that the network simply benefits from the image feature given by DC-i blocks. This motivates a generalized approach, where instead we feed each CNN-i block a concatenation of the following: the intermediate reconstruction $\boldsymbol{x}_i$, the self-adjoint $\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{x}_i$, and the adjoint of the input $\boldsymbol{A}^H\boldsymbol{y}$. Furthermore, one can also consider applying 1D-convolution in raw sensory domain using $f_{\text{sensor-cnn}}(.|\phi)$ to exploit the information along the sampling trajectory and remove unnecessary information (e.g. isolatable artifacts or noise). The resulting network is thus:

$$\boldsymbol{x}_0 = f_{\text{init-cnn}}\left(\boldsymbol{A}, f_{\text{sensor-cnn}}(\boldsymbol{y}|\phi_0)\middle|\theta_0\right) \tag{8}$$

$$\boldsymbol{x}_i = f_{\text{cnn}}\left(\boldsymbol{x}_{i-1}, \boldsymbol{A}^H f_{\text{sensor-cnn}}(\boldsymbol{A}\boldsymbol{x}_{i-1}|\phi_i), \boldsymbol{x}_0\middle|\theta_i\right) \tag{9}$$

where the learnable parameters are $\{\phi_0, \ldots, \phi_{N_{\text{it}}}, \theta_0, \ldots, \theta_{N_{\text{it}}}\}$. This variant is termed *Generalized Nonuniform Variational Network* (GNVN).

4 Results

Experimental Settings In order to test the efficacy of our proposed architecture, we performed the following simulation based studies. We used 640 randomly

[3] For $N = 192^2$ and $M = 10,000$ (i.e. $\approx 3\times$ acceleration), storing the complex-valued matrix alone already takes 3GB of memory.

selected T1-weighted and T2-weighted brain images from Human Connectome Project [4]. We used 600 for training and 40 for testing. To perform a realistic simulation, we performed the following pre-processing steps: firstly, we created complex-valued images by adding phase information to the magnitude data using two-dimensional Fourier bases with randomly sampled low order coefficients. Secondly, in order to simulate realistic measurements from receiver coils, we multiplied the images by spatially localized complex coil sensitivity profiles, which was derived from our in-house analytical coil model. Finally, we added a realistic amount of noise observable for parallel image acquisition. For the experiments, we resampled the images to a field of view (FOV) of $180 \times 180 \times 180\text{mm}^3$, with the isotrophic resolution of $3.4 \times 3.4 \times 3.4\text{mm}^3$, $1.7 \times 1.7 \times 1.7\text{mm}^3$ and $1.15 \times 1.15 \times 1.15\text{mm}^3$, resulting in the matrix size 64^3, 128^3 and 192^3 respectively.

For this work, we restrict our evaluation to single coil reconstruction in order to study the behaviour of non-uniform MR data reconstruction and leave the implementation as well as evaluation of parallel reconstruction as a future work. We undersampled the data using 2D nonuniform variable density, where the sampling density decay from the k-space centre at quadratic speed. For each matrix size, we generated the sampling trajectory with the target acceleration factor $R \in \{2, 4\}$. For evaluation, we measured mean squared error (MSE), structural similarity index measurement (SSIM) and peak signal-to-noise ratio (PSNR).

Implementation Details We re-implemented the following state-of-the-art methods which have been demonstrated on nonuniform MR data: AUTOMAP [18], image domain U-net and k-space domain U-net [7]. The input to AUTOMAP is a vector of k-space measurements, whereas the input to the U-net models were the gridding reconstruction in their respective domains. Note that we did not compare with models based on generative adversarial networks (GAN) because, if the capacity allows, one can always add GAN component to all models to observe similar effect. Due to our GPU memory limitation, we could only train AUTOMAP for the matrix size 64×64. For the proposed approach, termed NVN, we used U-net with 3 levels of downsampling (see Fig. 1) for each subnetwork. We used $N_{it} = 5$ for the number of blocks. We used adjoint for f_{init}. We also trained GNVN, where for $f_{\text{sensor-cnn}}$, we used a 5-layer CNN with a residual connection. We initialized the forward and adjoint operator based on [2], with oversampling factor 2. All deep learning methods were trained using MSE. Each network was trained for 8,000 epochs using Adam optimizer with $\alpha = 10^{-4}, \beta_1 = 0.9, \beta_2 = 0.999$. All methods were implemented in Tensorflow.

Results The quantitative result is summarized in Table 1. The proposed approaches consistently outperformed the baseline approaches for both acceleration factors. AUTOMAP and k-space U-net both underperformed compared to other methods. We speculate that the benefit of k-space convolution can only be appreciated for multicoil case or when combined with image-domain convolution.

[4] Available: `https://www.humanconnectome.org/study/hcp-young-adult/document/1200-subjects-data-release`

Table 1. Quantitative result for acceleration factor (R) 2 and 4. For each metric, mean and standard deviation is computed. MSE's are scaled by 10^3.

		$R = 2$			$R = 4$		
	Methods	MSE	SSIM	PSNR	MSE	SSIM	PSNR
64×64	AUTOMAP	2.40 (42.14)	0.87 (0.14)	29.87 (3.73)	2.59 (8.09)	0.84 (0.14)	28.36 (3.51)
	U-net	1.53 (18.13)	0.92 (0.11)	31.44 (3.86)	2.25 (21.87)	0.90 (0.10)	29.81 (3.74)
	U-net (k)	1.91 (7.40)	0.86 (0.13)	30.07 (3.57)	2.51 (6.58)	0.81 (0.13)	28.48 (3.34)
	NVN	**1.22 (12.51)**	0.93 (0.11)	32.33 (3.92)	**1.38 (4.04)**	0.92 (0.09)	30.95 (3.62)
	GNVN	1.22 (16.88)	**0.93 (0.09)**	**32.54 (4.00)**	1.37 (4.58)	**0.92 (0.08)**	**31.08 (3.66)**
128×128	U-net	0.75 (3.73)	0.94 (0.09)	34.06 (3.68)	0.91 (4.10)	0.94 (0.07)	32.76 (3.50)
	U-net (k)	1.02 (1.26)	0.89 (0.10)	32.51 (3.58)	1.54 (13.77)	0.87 (0.11)	31.32 (3.48)
	NVN	**0.57 (0.86)**	0.95 (0.06)	34.68 (3.57)	0.82 (1.07)	0.93 (0.07)	32.95 (3.54)
	GNVN	0.58 (1.99)	**0.95 (0.07)**	**34.83 (3.64)**	**0.67 (0.79)**	**0.95 (0.03)**	**33.65 (3.47)**
192×192	U-net	0.47 (1.55)	0.96 (0.05)	35.68 (3.67)	0.67 (1.13)	0.94 (0.07)	33.71 (3.23)
	U-net (k)	0.77 (0.81)	0.89 (0.10)	33.83 (3.62)	1.31 (7.53)	0.87 (0.11)	31.84 (3.35)
	NVN	**0.40 (0.60)**	0.96 (0.06)	36.11 (3.60)	0.66 (1.40)	0.91 (0.12)	34.01 (3.43)
	GNVN	0.40 (0.77)	**0.96 (0.05)**	**36.15 (3.57)**	**0.52 (0.44)**	**0.96 (0.03)**	**34.36 (3.07)**

Comparing the two of the proposed methods, while NVN showed higher data fidelity (lower MSE), GNVN offered better values for PSNR and SSIM. The sample reconstructions of T1-weighted image for $R = 2$ and T2-weighted image for $R = 4$ is shown in Fig. 2 and Fig. 3 respectively. While the overall differences between U-net, NVN and GNVN were small, the reconstructions from NVN and GNVN resulted in lower error, owing to the data consistency block. GNVN resulted in the lowest overall errors and preserved more of the fine details. Nevertheless, a certain level of blurriness can be observed in all images, due to the added noise. Again, U-net (k-space) for single coil resulted in a suboptimal reconstruction qualitatively. In Fig 4, we visualize the output of NVN and GNVN at each block. Interestingly, unlike CS-methods, the intermediate image can diverge from the final image. This is unsurprising as there was no constraint to enforce such behaviour. For NVN, most output of each block seems closer to the ground truth, presumably because the output of DC-i and CNN-i are explicitly combined. In comparison, GNVN showed more interesting features for all the intermediate stages, mainly highlighting the high frequency information. This observation was consistent across all images.

The number of parameters were 128.1M, 22.0M, 6.6M and 7.3M for AUTOMAP (64×64), U-net, NVN and GNVN respectively. The reconstruction speed were 5.928 ± 0.020 ms, 19.145 ± 0.072 ms, 19.459 ± 0.077 ms, 44.934 ± 0.088 ms, and 65.520 ± 0.100 ms for AUTOMAP (for the image size 64^2), U-net, U-net (k-space), NVN and GNVN respectively for the image size 192^2.

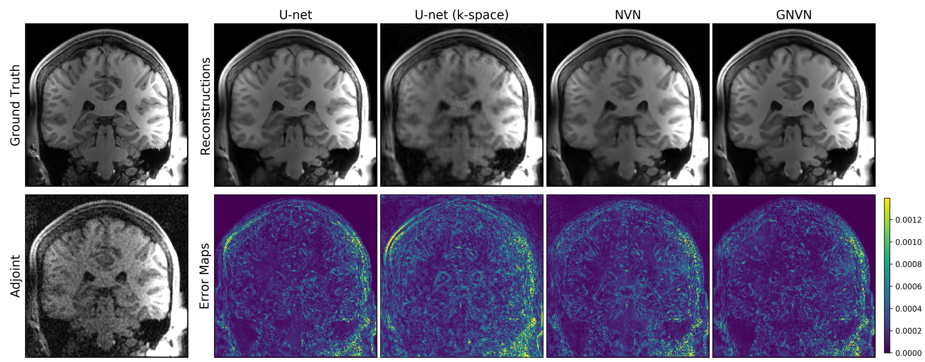

Fig. 2. The reconstructions for 192×192 T1 weighted image, $R = 2$. (top) Reconstructions (bottom) error maps.

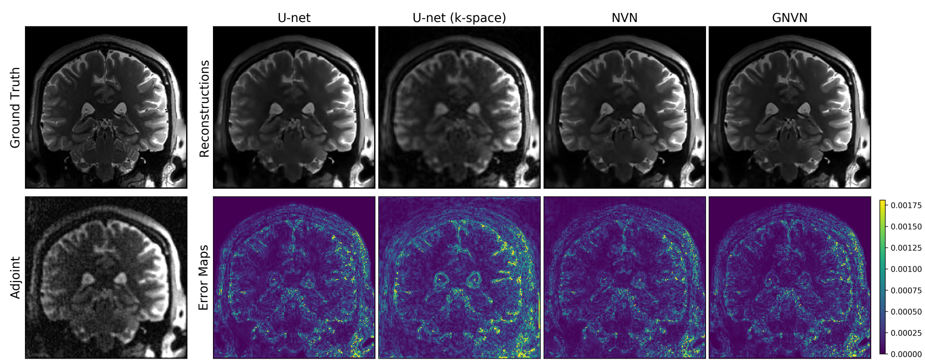

Fig. 3. The reconstructions for 192×192 T2 weighted image, $R = 4$. (top) Reconstructions (bottom) error maps.

5 Discussion and Conclusion

In this work, we proposed nonuniform variational network (NVN), a nonuniform extension of variational network, which performed the state-of-the-art deep learning models for single-coil scenario. The method is closely related to traditional optimization algorithms, however, we have shown that the generalized formulation (GNVN) also worked well in practice. In future, the extension to multi-coil data is considered. Another interesting direction is nonuniform sampling optimization for the deep learning models. Finally, we believe that the model is widely applicable to different image protocol and hence readily applied in a clinical scenario.

References

[1] Delattre, B.M., et al.: Spiral demystified. JMRI 28(6), 862–881 (2010)

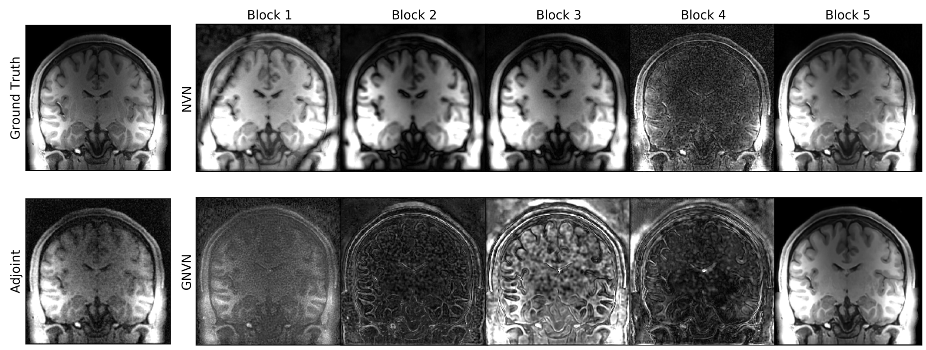

Fig. 4. Visualization of the iterative process of the network for NVN and GNVN. The intermediate results are used by the network to extract useful image features.

[2] Fessler, J.A., Sutton, B.P.: Nonuniform fast fourier transforms using min-max interpolation. IEEE TSP 51(2), 560–574 (2003)

[3] Forbes, K.P., et al.: Propeller mri: clinical testing of a novel technique for quantification and compensation of head motion. MRM 14(3), 215–222 (2001)

[4] Greengard, L., Lee, J.Y.: Accelerating the nonuniform fast fourier transform. SIAM Review 46(3), 443–454 (2004)

[5] Hammernik, K., et al.: Learning a variational network for reconstruction of accelerated mri data. MRM 79(6), 3055–3071 (2018)

[6] Han, Y., Ye, J.C.: k-space deep learning for accelerated MRI. arXiv preprint arXiv:1805.03779 (2018)

[7] Han, Y., et al.: Deep learning with domain adaptation for accelerated projection-reconstruction MR. MRM 80(3), 1189–1205 (2018)

[8] Knoll, F., et al.: Second order total generalized variation (TGV) for MRI. MRM 65(2), 480–491 (2011)

[9] Knoll, F., et al.: Adapted random sampling patterns for accelerated mri. Magnetic resonance materials in physics, biology and medicine 24(1), 43–50 (2011)

[10] Lazarus, C., et al.: Sparkling: variable-density k-space filling curves for accelerated t2*-weighted mri. MRM (2019)

[11] Lustig, M., et al.: Compressed sensing MRI. IEEE signal processing magazine 25(2), 72–82 (2008)

[12] Mardani, M., et al.: Deep generative adversarial networks for compressed sensing automates MRI. arXiv preprint arXiv:1706.00051 (2017)

[13] Meinhardt, T., et al.: Learning proximal operators: Using denoising networks for regularizing inverse imaging problems. In: ICCV. pp. 1781–1790 (2017)

[14] Pipe, J.G.: Motion correction with propeller mri: application to head motion and free-breathing cardiac imaging. MRM 42(5), 963–969 (1999)

[15] Qin, C., et al.: Convolutional recurrent neural networks for dynamic MR image reconstruction. arXiv preprint arXiv:1712.01751 (2017)

[16] Roth, S., Black, M.J.: Fields of experts. IJCV 82(2), 205 (2009)

[17] Yang, G., et al.: Dagan: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. IEEE TMI 37(6), 1310–1321 (2018)

[18] Zhu, B., et al.: Image reconstruction by domain-transform manifold learning. Nature 555(7697), 487 (2018)