High-Fidelity MRI Reconstruction with the Densely Connected Network Cascade and Feature Residual Data Consistency Priors

High-Fidelity MRI Reconstruction with the Densely Connected Network Cascade and Feature Residual Data Consistency Priors

Jingshuai Liu[1](✉), Chen Qin[1], and Mehrdad Yaghoobi[1]

IDCOM, School of Engineering, University of Edinburgh, Edinburgh, UK
{J.Liu,Chen.Qin,m.yaghoobi-vaighan}@ed.ac.uk

Abstract. Since its advent in the last century, magnetic resonance imaging (MRI) provides a radiation-free diagnosis tool and has revolutionized medical imaging. Compressed sensing (CS) methods leverage the sparsity prior of signals to reconstruct clean images from under-sampled measurements and accelerate the acquisition process. However, it is challenging to reduce strong aliasing artifacts caused by under-sampling and produce high-quality reconstructions with fine details. In this paper, we propose a novel GAN-based framework to recover the under-sampled images, which is characterized by a novel data consistency block and a densely connected network cascade used to improve the model performance in visual inspection and evaluation metrics. The role of each proposed block has been challenged in the ablation study, in terms of reconstruction quality metrics, using texture-rich FastMRI Knee image dataset.

Keywords: MRI reconstruction · GAN-based framework · dense network connections · data consistency.

1 Introduction

Magnetic resonance imaging (MRI) provides a non-invasive diagnosis tool in medical imaging. However, the long acquisition time hinders its growth and development in time-critic applications. The acquisition process can be accelerated by sampling fewer data. Under-sampling in k-space below the Nyquist-Shannon rate leads to aliasing artifacts in image domain, which restricts the acceleration factor in scanning. Many methods have been proposed to recover the under-sampled signals. Compressed sensing (CS) leverages the sparsity prior of signals to solve the ill-posed problems. Assuming sparsity representations in image domain [5] or in some transformed space [10, 14], CS methods retrieve reconstructions by solving sparsity regularized optimization. Nevertheless, the sparsity assumption can be difficult to hold in real-world scenarios and potentially fails to capture complicated features. Parallel imaging (PI) [8] unfolds aliasing artifacts in image domain and produce clean images by incorporating coil sensitivity priors. However, it is still difficult to remove strong artifacts and provide high-quality reconstructions under very low sampling rates.

In recent years, deep neural networks have achieved notable success in image tasks and show potential to benefit the development of modern MRI [22, 21, 9]. The work in [13] retrieves promising reconstructions using dual magnitude and phase networks. The method in [2] adopts a neural network to predict the missing k-space points. A primal-dual network is introduced in [25] to solve the conventional CS-MRI problem. VS-Net is proposed in [4] to find more accurate solutions. The method in [20] adopts a neural network to estimate the coil sensitivity maps used for parallel imaging. Despite their success, those methods still struggle to preserve sharp structures and local subtleties.

Generative adversarial networks (GAN) [7] model the data distribution via an adversarial competition between a generator and a discriminator. A GAN-based framework is introduced in [21] to yield more realistic images. The method in [18] incorporates a pre-trained generative network in the iterative reconstruction pipeline. A self-attention layer is used in [23] to improve the capacity of the generator and achieve better results. However, GAN-based methods can potentially produce unwanted artifacts and hallucinations and fail to capture diagnostic information. How to recover high-fidelity images from a few sampled data is still challenging, which is directly linked with the maximum MRI acceleration factor.

In this paper, we propose to reconstruct alias-free images from the under-sampled data in an end-to-end manner. We introduce a novel GAN-based reconstruction framework, which incorporates the data consistency prior and dense skip-connections to produce high-quality reconstructions with more accurate structures and local details. To evaluate the proposed framework, we compare it with other deep learning methods. It is demonstrated that the proposed method shows superiority in terms of visual quality and relevance metrics. The ablation studies show that the proposed model is able to deliver enhanced performance.

2 Method

2.1 Problem Formulation

The high-quality image s is under-sampled as follows,

$$y = A(s) = m \odot F(s), \tag{1}$$

where y is the measurement, A is the under-sampling operator, m is the sampling mask, F denotes the Fourier transform, and $\odot$ refers to the element-wise multiplication. CS methods recover the signal by solving the optimization below,

$$\min_x \|A(x) - y\|^2 + \lambda R(x), \tag{2}$$

where $R(x)$ denotes a regularization. Existing optimization solvers can be computationally expensive and hard to handle complex image features [21]. We instead provide an end-to-end solution using a trained neural network model.

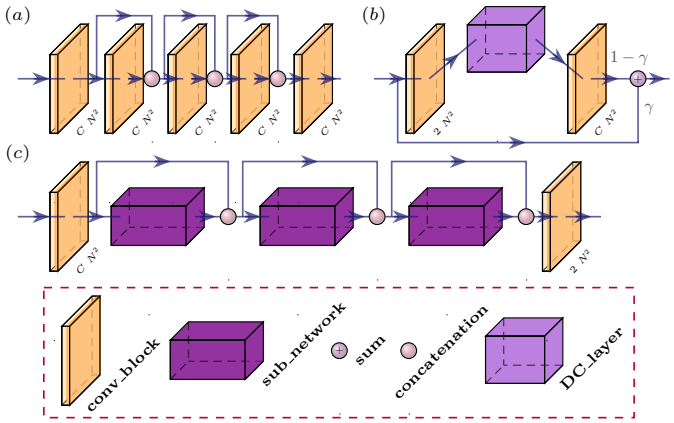


Fig. 1: Illustration of network features where the channel and spatial size of output features are denoted by C and N^2. a) Densely connected layers, b) FR-DCB, and c) densely connected network cascade.

2.2 Reconstruction Framework

In this section, we first introduce the data consistency block and network cascade, and then describe the model design.

Feature Residual Data Consistency Block (FR-DCB) We propose to leverage data consistency (DC) blocks to "correct" the intermediate predictions and provide more faithful results. However, conventional DC blocks inevitably require to collapse the feature channels to fit the complex-valued data, which can have a damaging effect on final performance due to the bottleneck design. To benefit feature propagation and alleviate the bottleneck problem, we take the advantage of residual learning at the feature level and implement the DC operation by,

$$h \leftarrow \gamma h + (1-\gamma) f^*(DC(f(h))), \tag{3}$$

where h denotes the output features, f and f^* are two convolutional layers used for channel collapse and expansion, and DC refers to the data consistency update given by,

$$DC(x) = F^{-1}(m \odot y + (1-m) \odot F(x)), \tag{4}$$

where F^{-1} denotes the inverse Fourier transform. The resultant block, dubbed feature residual data consistency block (FR-DCB), is illustrated in Fig. 1.

Densely Connected Network Cascade Deep cascade of neural networks are widely used for MRI reconstruction [13, 4, 19, 1], which potentially yield higher performance due to the powerful representations learned with a deep structure. Inspired by densely connected layers introduced in [11], we propose a densely connected reconstruction framework, see Fig. 1 and 2, to enable feature transmission and reuse by dense skip-connections between sub-networks. It is noteworthy

that the outputs of each sub-network are feature volumes rather than 2-channel images, e.g. in [1, 17], which potentially resists the bottleneck problem. The predictions from all preceding sub-networks are collected together with the current output via channel concatenation and fed to the following sub-network as input. The feature maps from all sub-networks are fused together to produce the final outcome, as illustrated in Fig. 1 and 2.

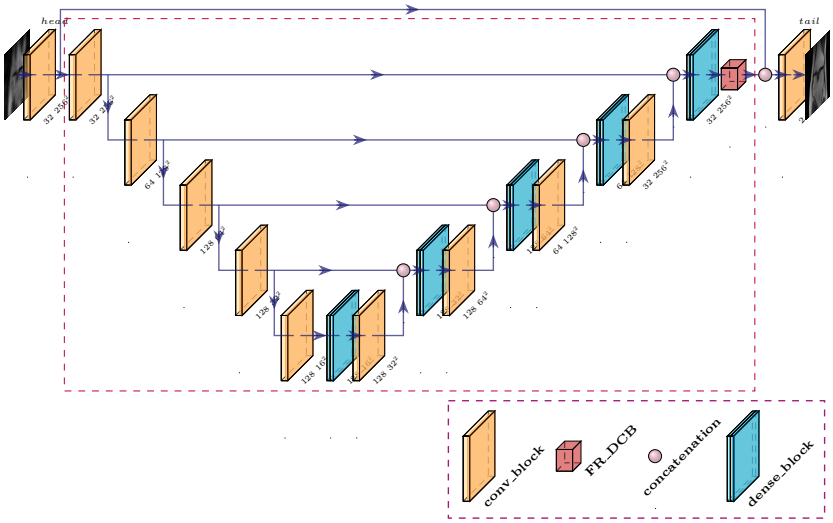


Fig. 2: Illustration of the model architecture where a single sub-network is displayed for brevity. The zero-filled as input is first mapped by a head layer. The output features from all preceding sub-networks are collected to form dense bypass connections. The final result is given via a tail layer.

Model Design We adopt the U-shaped structure, displayed in Fig. 2, as sub-networks to form the reconstruction framework where 5 sub-networks are deployed. Densely connected layers [11] are embedded in all decoding levels to refine feature representations and the FR-DCB blocks are appended to each sub-network. We use the zero-filled $z = F^{-1}(y)$ as input to the framework, which is first mapped via a convolutional layer and subsequently fed to sub-networks. The features from each sub-network are consecutively used to construct a densely connected cascade. The reconstruction G is given by fusing all collected features.

2.3 Objective Function

We use the L_1 metric and structural similarity index (SSIM) to measure the reconstruction errors. The loss is given by,

$$L_{rec} = (1 - \alpha)L_1(G, s) + \alpha L^{SSIM}(G, s), \tag{5}$$

where we set $\alpha = 0.4$. We adopt the Least Squares GAN (LSGAN) [16] to encourage realistic details. It can prevent the saturation issue and provide more stable and faster convergences [16], compared to vallina GAN. The adversarial loss is computed as follows,

$$\begin{aligned} L_{adv}^{D} &= E[\|D(s) - b\|_2^2] + E[\|D(G) - a\|_2^2] \\ L_{adv}^{G} &= E[\|D(G) - c\|_2^2], \end{aligned} \tag{6}$$

where D is a discriminator, E denotes the expectation, and the hyper-parameters are set to be $a = 0$ and $b = c = 1$. The perceptual loss, which is normally used in a GAN-based framework to resist hallucinations, is given as follows,

$$L_{vgg} = \sum_i (\|f_{vgg}^i(G) - f_{vgg}^i(s)\|_1 + \beta \|f_{gram}^i(G) - f_{gram}^i(s)\|_1), \tag{7}$$

where f_{vgg}^i is the pre-activations at the i-th layer of a pre-trained network, e.g. VGG [21], f_{gram}^i is the Gram matrix [6], and $\beta = 0.005$. It leverages the deep structure to learn more consistent representations with the human visual system. The total objective is given as below,

$$L = E_{\{(G,s)\}}[\lambda_{rec} L_{rec} + \lambda_{adv} L_{adv} + \lambda_{vgg} L_{vgg}], \tag{8}$$

where we set $\lambda_{rec} = 10$, $\lambda_{adv} = 0.05$, and $\lambda_{vgg} = 0.5$.

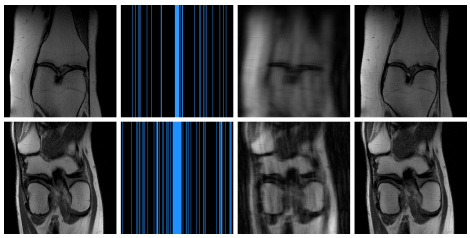
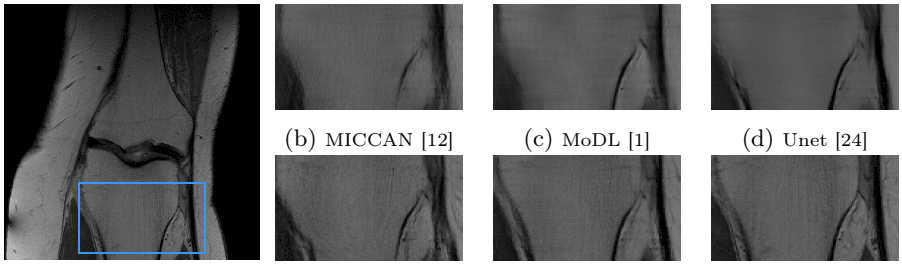


Fig. 3: Illustration of 8× and 4× under-sampling and reconstruction. First) fully sampled, second) sampling pattern, third) zero-filled, and last) reconstruction.

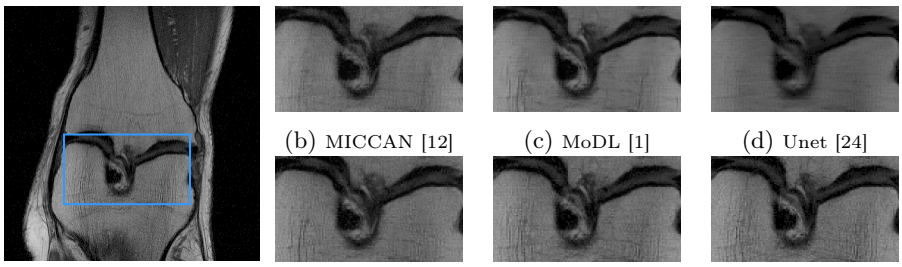
3 Experiment

We conduct experiments on single-coil knee MR images which contain rich structures and textures. We extract 2800 samples from the FastMRI Knee database [24] for training and 164 samples from different cases for test. A fixed random mask is adopted in the under-sampling operation, as presented in Fig. 3, where the total reduction factor is respectively set to 8 and 4 with 4% and 8% central lines preserved. The model is trained for 30 epochs with a batch size of 4. An Adam optimizer is used with β_1=0.5, β_2=0.999, and a learning rate of 1e-5. We use two channels to handle complex-valued data, e.g. inputs and outputs.



(a) ground truth (GT) (b) MICCAN [12] (c) MoDL [1] (d) Unet [24] (e) ASGAN [15] (f) proposed (g) GT patch

Fig. 4: Comparison results of 8× accelerated single-coil MRI reconstruction.



(a) ground truth (GT) (b) MICCAN [12] (c) MoDL [1] (d) Unet [24] (e) ASGAN [15] (f) proposed (g) GT patch

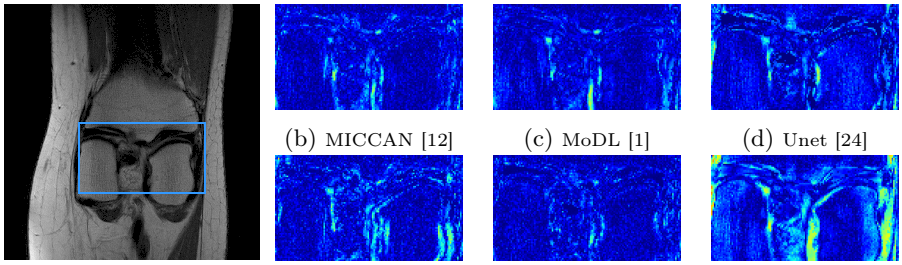
Fig. 5: Comparison results of 4× accelerated single-coil MRI reconstruction.

3.1 Comparison Results

We compare the proposed method with other state-of-the-art approaches: MICCAN [12], MoDL [1], FastMRI Unet [24], and ASGAN [15]. The reconstruction results are displayed in Fig. 4 and 5. From Fig. 4, we found that the proposed method produces superior reconstructions with rich textural and structural details, which leads to more realistic and visually promising results. We can observe that ASGAN generates fine local subtleties, whereas it can suffer from textural artifacts, see Fig. 5 (e), and disrupted structures, see Fig. 4 (e). By contrast, the proposed method provides more faithful reconstructions. From the residual maps shown in Fig 6, we can observe that our method produces fewer errors. We adopt PSNR and SSIM as evaluation metrics, where higher values are better, and use FID and KID [3] to measure the visual quality, which prefer lower scores. The quantitative results are shown in Table 1, which shows that the proposed method consistently surpasses other approaches in terms of relevant metrics.

3.2 Ablation Studies on Model Components

We conduct ablation studies to verify the effectiveness of the proposed model features. We remove the FR-DCB modules from the framework and compare it with the proposed model. For simplicity, the acceleration factor is selected to



(a) ground truth (GT) (b) MICCAN [12] (c) MoDL [1] (d) Unet [24] (e) ASGAN [15] (f) proposed (g) zero-filled

Fig. 6: Residual maps (2× amplified) of 8× accelerated single-coil MRI reconstruction.

Table 1: Quantitative Evaluation on Accelerated MRI Reconstruction.

method	PSNR↑	SSIM↑	FID↓	KID↓
8×				
proposed	**27.25**	**0.720**	**81.06**	**0.014**
ASGAN [15]	25.45	0.638	104.34	0.036
FastMRI Unet [24]	25.82	0.703	160.35	0.121
MoDL [1]	27.13	0.620	143.65	0.080
MICCAN [12]	26.61	0.642	180.66	0.146
zero-filled	20.54	0.388	423.32	0.533
4×				
proposed	**31.11**	**0.824**	**63.74**	**0.006**
ASGAN [15]	27.73	0.711	82.18	0.016
FastMRI Unet [24]	28.35	0.771	118.07	0.061
MoDL [1]	30.34	0.745	98.86	0.042
MICCAN [12]	30.11	0.711	99.44	0.040
zero-filled	23.94	0.486	255.06	0.239

8 for ablation studies. The ablation results are presented in Table 2. We found that the removal of FR-DCB leads to performance drop in all evaluation metrics by a large margin. To testify the efficacy of the densely connected cascade, we remove the dense bypasses between sub-networks and repeat the feature volumes to fit the input channel size. From Table 2, we observed that the densely connected cascade provides better reconstruction results. It is demonstrated that the proposed model features are all able to obtain improved performance.

3.3 Ablation Studies on Bottleneck Design in DC Blocks

To further verify the effectiveness of FR-DCB and show the influence of the bottleneck design in conventional DC blocks, we remove the feature residual

Table 2: Ablation Studies on Model Components Using 8× Acceleration.

method	PSNR↑	SSIM↑	FID↓	KID↓
proposed	**27.25**	**0.720**	**81.06**	**0.014**
w/o FR-DCB	25.72	0.687	94.12	0.028
w/o dense cascade	27.02	0.713	85.82	0.019

Table 3: Ablation on Feature Connection in FR-DCB Using 8× Acceleration.

method	PSNR↑	SSIM↑	FID↓	KID↓
proposed	**27.25**	**0.720**	**81.06**	**0.014**
w/o DC shortcut	26.91	0.713	88.81	0.023

shortcut in FR-DCB and implement the update rule as shown below,

$$h \leftarrow f^*(\gamma f(h) + (1-\gamma)DC(f(h))), \tag{9}$$

which is mathematically equivalent to those used in [1, 12, 19], where f^* and f can be omitted and absorbed into sub-networks. We present the results in Table 3. It shows that the update rule in (9) reduces PSNR and SSIM scores and concomitantly increases FID and KID, which indicates the adverse impact of the bottleneck design in conventional DC blocks and confirms the efficacy of the proposed FR-DCB.

4 Conclusions and Discussion

A novel GAN-based deep neural network framework is introduced in this paper to provide an end-to-end solution to the high-fidelity MRI reconstruction problem. The framework incorporates a novel data consistency block and a densely connected cascade structure to improve the model performance in recovering accelerated MR images with rich structural and textural details. In experiments, the proposed approach achieves superior high-quality reconstruction results with a high acceleration factor in a comparison with other deep learning-based methods both qualitatively and quantitatively, using FastMRI Knee dataset. The future researches include extending the method to parallel imaging, accelerating model deployments, and applying it to other texture-rich MRI imaging modalities.

References

1. Aggarwal, H., Mani, M., Jacob, M.: MoDL: model-based deep learning architecture for inverse problems. IEEE Transactions on Medical Imaging **38**(2), 394–405 (2019). https://doi.org/10.1109/TMI.2018.2865356

2. Anuroop, S., Jure, Z., Tullie, M., Lawrence, Z., Aaron, D., K.S., D.: GrappaNet: combining parallel imaging with deep learning for multi-coil MRI reconstruction. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) pp. 14303–14310 (2020). https://doi.org/10.1109/CVPR42600.2020.01432
3. Bińkowski, M., Sutherland, D., Arbel, M., Gretton, A.: Demystifying MMD GANs. International Conference on Learning Representations (2018)
4. Duan, J., Schlemper, J., Qin, C., Ouyang, C., Bai, W., Biffi, C., Bello, G., Statton, B., Regan, D., Rueckert, D.: VS-Net: variable splitting network for accelerated parallel MRI reconstruction. Springer International Publishing pp. 713–722 (2019)
5. Fair, M., Gatehouse, P., DiBella, E., Firmin, D.: A review of 3D first-pass, whole-heart, myocardial perfusion cardiovascular magnetic resonance. Journal of Cardiovascular Magnetic Resonance (2015)
6. Gatys, L., Ecker, A., Bethge, M.: Image style transfer using convolutional neural networks. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) pp. 2414–2423 (June 2016). https://doi.org/10.1109/CVPR.2016.265
7. Goodfellow, I., Pouget, A., Mirza, M., Xu, B., Warde, F., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial networks. Advances in Neural Information Processing Systems **27**, 2672–2680 (2014)
8. Griswold, M., Jakob, P., Heidemann, R.M., Nittka, M., Jellus, V., Wang, J., Kiefer, B., Haase, A.: Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magnetic resonance in medicine **47**(6), 1202—1210 (June 2002). https://doi.org/10.1002/mrm.10171
9. Hammernik, k., Klatzer, T., Kobler, E., Recht, M., Sodickson, D., Pock, T., Knoll, F.: Learning a variational network for reconstruction of accelerated MRI data. Magnetic Resonance in Medicine **79**(6), 3055–3071 (2018). https://doi.org/https://doi.org/10.1002/mrm.26977
10. Hong, M., Yu, Y., Wang, H., Liu, F., Crozier, S.: Compressed sensing MRI with singular value decomposition-based sparsity basis. Physics in Medicine and Biology pp. 6311–6325 (Sep 2021)
11. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) pp. 2261–2269 (2017)
12. Huang, Q., Yang, D., Wu, P., Qu, H., Yi, J., Metaxas, D.: MRI reconstruction via cascaded channel-wise attention network. 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) pp. 1622–1626 (2019). https://doi.org/10.1109/ISBI.2019.8759423
13. Lee, D., Yoo, J., Tak, S., Ye, J.: Deep residual learning for accelerated MRI using magnitude and phase networks. IEEE Trans. Biomed. Eng **65**(9), 1985–1995 (2018)
14. Lingala, S., Jacob, M.: Blind compressive sensing dynamic MRI. IEEE Transactions on Medical Imaging **32**(6), 1132–1145 (2013)
15. Liu, J., Yaghoobi, M.: Fine-grained MRI reconstruction using attentive selection generative adversarial networks. ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) pp. 1155–1159 (2021)
16. Mao, X., Li, Q., Xie, H., Lau, R., Wang, Z., Smolley, S.: Least squares generative adversarial networks. 2017 IEEE International Conference on Computer Vision (ICCV) pp. 2813–2821 (2017)
17. Mardani, M., Gong, E., Cheng, J.Y., Vasanawala, S.S., Zaharchuk, G., Xing, L., Pauly, J.M.: Deep generative adversarial neural networks for compressive sensing MRI. IEEE Transactions on Medical Imaging **38**(1), 167–179 (2019). https://doi.org/10.1109/TMI.2018.2858752

18. Narnhofer, D., Hammernik, K., Knoll, F., Pock, T.: Inverse GANs for accelerated MRI reconstruction. Wavelets and Sparsity XVIII **11138**, 111381A (Sep 2019). https://doi.org/10.1117/12.2527753
19. Schlemper, J., Caballero, J., Hajnal, J., Price, A., Rueckert, D.: A deep cascade of convolutional neural networks for dynamic MR image reconstruction. IEEE Transactions on Medical Imaging **37**(2), 491–503 (2018). https://doi.org/10.1109/TMI.2017.2760978
20. Sriram, A., Zbontar, J., Murrell, T., Defazio, A., Zitnick, C., Yakubova, N., Knoll, F., Johnson, P.: End-to-end variational networks for accelerated MRI reconstruction. Medical Image Computing and Computer Assisted Intervention - MICCAI **12262**, 64–73 (2020)
21. Yang, G., Yu, S., Dong, H., Slabaugh, G., Dragotti, P.L., Ye, X., Liu, F., Arridge, S., Keegan, J., Guo, Y., Firmin, D.: DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. IEEE Transactions on Medical Imaging **37**(6), 1310–1321 (2018)
22. Yang, Y., Sun, J., Li, H., Xu, Z.: Deep ADMM-Net for compressive sensing MRI. Advances in Neural Information Processing Systems **29** (2016)
23. Yuan, Z., Jiang, M., Wang, Y., Wei, B., Li, Y., Wang, P., Menpes-Smith, W., Niu, Z., Yang, G.: SARA-GAN: self-attention and relative average discriminator based generative adversarial networks for fast compressed sensing MRI reconstruction. Frontiers in Neuroinformatics **14**, 1–12 (2020). https://doi.org/10.3389/fninf.2020.611666
24. Zbontar, J., Knoll, F., Sriram, A., Muckley, M., Bruno, M., Defazio, A., Parente, M., Geras, K., Katsnelson, J., Chandarana, H., Zhang, Z., Drozdzal, M., Romero, A., Rabbat, M., Vincent, P., Pinkerton, J., Wang, D., Yakubova, N., Owens, E., Zitnick, C., Recht, M., Sodickson, D., Lui, Y.: FastMRI: an open dataset and benchmarks for accelerated MRI. CoRR **abs/1811.08839** (2018)
25. Zhang, C., Liu, Y., Shang, F., Li, Y., Liu, H.: A novel learned primal-dual network for image cmpressive sensing. IEEE Access **9**, 26041–26050 (2021). https://doi.org/10.1109/ACCESS.2021.3057621