Stochastic Deep Compressive Sensing for the Reconstruction of Diffusion Tensor Cardiac MRI

Jo Schlemper[1], Guang Yang[2,3(✉)], Pedro Ferreira[2,3], Andrew Scott[2,3], Laura-Ann McGill[2,3], Zohya Khalique[2,3], Margarita Gorodezky[2,3], Malte Roehl[2,3], Jennifer Keegan[2,3], Dudley Pennell[2,3], David Firmin[2,3], and Daniel Rueckert[1]

[1] Department of Computing, Imperial College London, London, UK
jo.schlemper11@imperial.ac.uk
[2] National Heart and Lung Institute, Imperial College London, London, UK
[3] Cardiovascular Research Centre, Royal Brompton Hospital, London, UK
g.yang@imperial.ac.uk

Abstract. Understanding the structure of the heart at the microscopic scale of cardiomyocytes and their aggregates provides new insights into the mechanisms of heart disease and enables the investigation of effective therapeutics. Diffusion Tensor Cardiac Magnetic Resonance (DT-CMR) is a unique non-invasive technique that can resolve the microscopic structure, organisation, and integrity of the myocardium without the need for exogenous contrast agents. However, this technique suffers from relatively low signal-to-noise ratio (SNR) and frequent signal loss due to respiratory and cardiac motion. Current DT-CMR techniques rely on acquiring and averaging multiple signal acquisitions to improve the SNR. Moreover, in order to mitigate the influence of respiratory movement, patients are required to perform many breath holds which results in prolonged acquisition durations (e.g., ∼30 min using the existing technology). In this study, we propose a novel cascaded Convolutional Neural Networks (CNN) based compressive sensing (CS) technique and explore its applicability to improve DT-CMR acquisitions. Our simulation based studies have achieved high reconstruction fidelity and good agreement between DT-CMR parameters obtained with the proposed reconstruction and fully sampled ground truth. When compared to other state-of-the-art methods, our proposed deep cascaded CNN method and its stochastic variation demonstrated significant improvements. To the best of our knowledge, this is the first study using deep CNN based CS for the DT-CMR reconstruction. In addition, with relatively straightforward modifications to the acquisition scheme, our method can easily be translated into a method for online, at-the-scanner reconstruction enabling the deployment of accelerated DT-CMR in various clinical applications.

1   Introduction

Diffusion Tensor Cardiovascular Magnetic Resonance (DT-CMR) is a unique non-invasive technique, which provides rich structural and functional information on the myocardium at a microscopic scale, including parameters relating to the alignment and integrity of cardiomyocytes and aggregates of cardiomyocytes, known as sheetlets. Despite a long history of *in vivo* DT-CMR and great efforts to drive the method towards a clinically usable technique [15], its clinical use remains limited to research studies in a few specialist centres. These limited studies have produced a number of interesting findings in detection and diagnosis of ischemic heart disease, hypertrophic and dilated cardiomyopathies [2,3,14].

In comparison to the well established application of diffusion tensor imaging (DTI) in the brain, DT-CMR faces a number of additional challenges, including: (1) the intrinsically low signal-to-noise ratio (SNR) of typical acquisition methods, which means that multiple signal acquisitions must be acquired and averaged to improve the net SNR; (2) signal loss caused by respiratory and cardiac motion during the application of diffusion sensitising gradients, means that these gradients must be short with respect to the motion and strong gradients must be used to provide sufficient sensitivity to the small diffusive movement of water molecules (on the order of microns) [1]; (3) the transverse relaxation time (T_2) is substantially shorter in the myocardium than in the brain ($\sim$40 ms for the myocardium vs. $\sim$80 ms for the white matter of the brain). In contrast to neurological DTI, this dramatically limits the possible echo time (TE) that can be used in DT-CMR [13]; (4) the increased B_0 inhomogeneity in the thorax may also result in more susceptibility-related distortions in the echo planar imaging (EPI) technique typically used for the DT-CMR [13], which also limits its spatial resolution. Some of these challenges have been partially addressed or have benefited from parallel imaging techniques for the in-plane acceleration, e.g., using SENSE (sensitivity encoding) and GRAPPA (generalized autocalibrating partially parallel acquisitions), or simultaneous multislice imaging [9]. However, typical acceleration factors are normally limited to 2–3 [13].

Compressive sensing (CS) is a promising technique for fast MRI [11] that circumvents the Nyquist-Shannon sampling criteria and can achieve a more aggressive acceleration. Comprehensive reviews of CS based fast MRI (CS-MRI) can be found elsewhere [6]. Essentially, CS-MRI can obtain a *perfect* reconstruction by using a *nonlinear optimisation* on *randomly undersampled* raw data, assuming the data or its transformation is *compressible*. Although CS-MRI have been widely investigated [6], most previous studies have focused on the acceleration of the structural MRI, and only very few research studies have been conducted on DT-CMR [8,12,22]. These CS based fast DT-CMR methods demonstrated promising reconstruction results; however, the iterative nonlinear optimisation used in these methods requires a lengthy reconstruction procedure that could prevent their widespread usage, where there is a clinical need to view images immediately at the scanner. Therefore, a CS-based on-the-fly reconstruction of DT-CMR data would be highly desirable.

More recently, deep learning approaches have shown intriguing results in solving various medical image segmentation, registration and reconstruction problems. In particular, there are several deep learning based architectures that have been proposed for reconstruction of MRI data. The most widely used architecture is named U-Net [18], which is used to perform an end-to-end reconstruction [5,10], and often combined with a residual learning [10,23] or generative adversarial networks [23]. Alternative approaches have also been proposed, which modified the deep network architectures to embed traditional optimisation algorithms. These include gradient descent [4], alternating direction method of multipliers (ADMM) [21] or optimisation algorithms inspired by variable splitting techniques [16]. In addition, various clinical applications have been explored including knee imaging [4], brain imaging [23], and dynamic cardiac imaging [19].

In this study, a novel cascaded Convolutional Neural Networks (CNN) based CS technique has been proposed to simulate an efficient reconstruction of highly undersampled DT-CMR data. The proposed architecture improves upon the previously proposed networks by introducing dilated convolution instead of pooling to efficiently increase the receptive field. In addition, we introduce a novel *stochastic* architecture, which is formulated by dropping the subnetworks at training. We show that this approach provides multifold benefits, including accelerated learning, improved robustness and an additional uncertainty estimate of the prediction. The following sections present the details of the proposed methodology, experimental set-up, achieved results and followed by discussions and conclusion.

2   Method

Let $\mathbf{x} \in \mathbb{C}^N$ 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 $\mathbf{y} \in \mathbb{C}^M$ ($M << N$) represent the undersampled k-space measurements. Our problem is to reconstruct $\mathbf{x}$ from $\mathbf{y}$, formulated as an unconstrained optimisation:

$$\operatorname*{argmin}_{\mathbf{x}} \quad \mathcal{R}(\mathbf{x}) + \lambda \|\mathbf{y} - \mathbf{F}_u \mathbf{x}\|_2^2 \tag{1}$$

Here $\mathbf{F}_u$ is an undersampling Fourier encoding matrix, $\mathcal{R}$ expresses regularisation terms on $\mathbf{x}$ and λ is a hyper-parameter often associated to the noise level.

Deep Cascaded CNN. In general, the regularisation terms $\mathcal{R}$ in Eq. 1 can be non-convex (such as ℓ_0 in the sparsifying domain). Therefore, traditionally, one introduces an auxiliary variable $\mathbf{z}$ as variable splitting technique and solves the following penalty functional:

$$\operatorname*{argmin}_{\mathbf{x},\mathbf{z}} \mathcal{R}(\mathbf{z}) + \lambda \|\mathbf{y} - \mathbf{F}_u \mathbf{x}\|_2^2 + \mu \|\mathbf{x} - \mathbf{z}\|_2^2 \tag{2}$$

where μ is a penalty parameter. By applying alternating minimisation over $\mathbf{x}$ and $\mathbf{z}$, Eq. 2 can be solved via the following iterative procedures:

$$\mathbf{z}^{(i)} = \underset{\mathbf{z}}{\operatorname{argmin}} \ \mathcal{R}(\mathbf{z}) + \mu \|\mathbf{x}^{(i)} - \mathbf{z}\|_2^2 \tag{3a}$$

$$\mathbf{x}^{(i+1)} = \underset{\mathbf{x}}{\operatorname{argmin}} \ \lambda \|\mathbf{y} - \mathbf{F}_u \mathbf{x}\|_2^2 + \mu \|\mathbf{x} - \mathbf{z}^{(i)}\|_2^2 \tag{3b}$$

where $\mathbf{x}^{(0)} = \mathbf{x}_u = \mathbf{F}_u^H y$ is the zero-filled (ZF) reconstruction taken as an initialisation and $\mathbf{z}$ can be seen as an intermediate state of the optimisation process. For MRI reconstruction, Eq. 3b is often regarded as a *data consistency* (DC) step where we could obtain the following closed-form solution [20]:

$$\mathbf{x}^{(i+1)} = \operatorname{DC}(\mathbf{z}^{(i)}; \mathbf{y}, \lambda_0, \Omega) = \mathbf{F}^H \mathbf{\Lambda} \mathbf{F} \mathbf{z}^{(i)} + \frac{\lambda_0}{1+\lambda_0} \mathbf{F}_u^H \mathbf{y}, \mathbf{\Lambda}_{kk} = \begin{cases} 1 & \text{if } k \notin \Omega \\ \frac{1}{1+\lambda_0} & \text{if } k \in \Omega \end{cases} \tag{4}$$

which $\mathbf{F}$ is the full Fourier encoding matrix (a discrete Fourier transform in this case), $\lambda_0 = \lambda/\mu$ is a ratio of regularization parameters from Eq. 4, Ω is an index set of the acquired k-space samples and $\mathbf{\Lambda}$ is a diagonal matrix. Equation 3a is the proximal operator of the prior $\mathcal{R}$, and instead of explicitly determining the form of the regularisation term, one can learn the proximal operator by using the CNN directly. In so doing, iterative reconstruction with a cascaded CNN and DC steps is performed. The whole framework can be optimized end-to-end, yielding one cascaded deep network. This network is referred to as *DC-CNN*.

Stochastic DC-CNN. Inspired by [7], we extend the DC-CNN framework into a *stochastic DC-CNN* (s-DC-CNN)—during the training, the i-th subnetwork is dropped with a probability of $p = (i-1)/2n_c$, where n_c is the total number of the cascaded CNN. During the testing (inference), we can use all of the subnetworks to perform the reconstruction, which is expected to provide the best performance as the most depths are used. Alternatively, we could sample the network configurations θ using the distribution P from the above strategy, and then reconstruct the image as an ensemble of the sampled model $\bar{x} = \mathbb{E}_{\theta \sim P}(f(x_u|\theta))$. In so doing, the variance of the predicted value can be used as an *uncertainty estimate* given the ensemble. This alternative approach provides several benefits: (1) it accelerates the learning because the expected depth of the network is much shorter; (2) this simultaneously helps the error terms to be backpropagated better as the depth is shortened, allowing us to train deeper networks if desired; (3) due to the stochastic connection, each subnetwork can see different levels of residual noise that help the network learn better and become more robust.

In addition, we used dilated convolution to efficiently increase the receptive field. In the original DC-CNN framework, each denoising subnetwork only consists of 5 layers of 3×3 2D convolution layers, which has a limited receptive field of size 11. In our new s-DC-CNN framework, we employed the dilated convolution except for the first and the last convolution that has a receptive field of size 23. It is of note that compared to pooling operation, the dilated convolution can avoid the needs of upsampling, and subsequently prevents information

loss and interpolation artefacts, and also keeps the network more compact. Furthermore, we employed batch-normalisation to improve the training and applied leaky rectified linear unit with $\alpha = 0.01$.

3   Results

Experimental Settings. In order to test the efficacy of our proposed DC-CNN method for the DT-CMR reconstruction, we performed the following simulation based studies. First, we simulated undersampled DT-CMR datasets using *in vivo* DT-CMR data acquired at peak-systole or in diastasis in healthy volunteers. The data were acquired using a breath hold stimulated echo acquisition mode EPI sequence [15], with diffusion encoded over 1 complete cardiac cycle (detailed scanning parameters in Supplementary Materials). Evaluation has been carried out using 178 DT-CMR scans.

The image data were firstly converted into k-space data using an inverse Fourier transform and undersampled by a pseudo-random 1D undersampling mask. Each line was sampled from a Gaussian distribution with mean centred at the origin of *k*-space and variance proportional to the extent of *k*-space. We applied various random undersampling patterns to the data acquired for different diffusion weighting directions and different signal averages in order to make the trained network generalise better on various aliasing artefacts. Note that in this work, the magnitude images were used. However, the effectiveness of the method to complex-valued data has been previously verified in [20].

The studied scans were split into independent training, validation and test sets containing 142, 17 and 19 cases, respectively. In the training stage, both information from fully-sampled data and undersampled data were used. In the testing (inference) stage, the zero-filled undersampled data were input to the trained network to yield a reconstruction.

Comparison Studies. To compare with other state-of-the-art conventional CS-MRI and recently proposed deep learning based methods, we re-implemented the following methods for the DT-CMR reconstruction, including a dictionary learning based CS-MRI (i.e., *DLMRI* [17]) and a U-Net based method (*UNET-CS*). For the proposed approach, we selected to use 6-layer subnetwork with $n_c = 15$ that yields a 105-layer network including the DC layers. Note that the hyperparameters were chosen from a grid search at a coarse scale but is by no means optimal. We performed an ablation study to test the benefits of the proposed improvements. The original network without and with dilated convolution are denoted as *DC-CNN* [19] and *DC-CNN-d*. The proposed stochastic version without and with dilated convolution are referred as *s-DC-CNN* and *s-DC-CNN-d* (the detailed network architectures are described in the supplementary materials). For the stochastic networks, we used all subnetworks for reconstruction except for when generating the uncertainty map.

Training Settings. Each network was trained for 100 epochs using mean squared error loss between the network reconstruction and the ground truth. In the first 50 epochs, the network was trained on various undersampling factors (UF) of $(0-12\times)$. For the last 50 epochs, the network was fine-tuned for the UF of $\{2\times, 5\times, 8\times\}$ individually for each experiment. We used the Adam optimiser with a learning rate of 10^{-4}, that was subsequently reduced by a factor of 10 for every 20 epochs during the fine-tuning.

Results. Table 1 tabulates mean and standard deviation (std) of the peak signal-to-noise ratio (PSNR) we obtained by various CS-MRI methods. For the three different UF we tested, the proposed s-DC-CNN-d obtained the best performance, and all our DC-CNN variations outperformed the DLMRI and the UNET-CS. This has been further confirmed by the root mean-squared-error (RMSE) calculated from the computed fractional anisotropy (FA), mean diffusivity (MD) and helix angle (HA) of the $5\times$ undersampling cases.

Table 1. Quantitative results [mean (std)] of the 19 independent testing subjects.

| UF	Model	s-DC-CNN-d	s-DC-CNN	DC-CNN-d	DC-CNN	UNET-CS	DLMRI	Zero-Filled
2×	**PSNR**	**37.747 (2.41)**	37.667 (2.35)	37.64 (2.33)	37.73 (2.36)	33.16 (2.14)	33.62 (2.26)	27.89 (2.08)
5×	**PSNR**	**30.96 (2.10)**	30.85 (2.13)	30.88 (2.08)	30.8 (2.1)	28.40 (2.44)	28.40 (2.44)	23.71 (1.72)
8×	**PSNR**	**28.81 (2.00)**	28.50 (2.01)	28.79 (1.97)	28.62 (2.00)	27.35 (2.00)	25.88 (2.36)	22.71 (1.64)
5×	**FA RMSE**	**0.08 (0.03)**	**0.08 (0.03)**	0.09 (0.03)	0.09 (0.03)	0.11 (0.05)	0.11 (0.04)	0.18 (0.04)
	MD RMSE	**0.11 (0.05)**	0.12 (0.05)	**0.11 (0.05)**	0.12 (0.05)	0.16 (0.08)	0.15 (0.07)	0.19 (0.10)
	HA RMSE	**13.34 (2.93)**	13.97 (3.11)	13.68 (2.91)	14.04 (3.04)	16.29 (3.27)	17.88 (2.9)	26.87 (3.31)
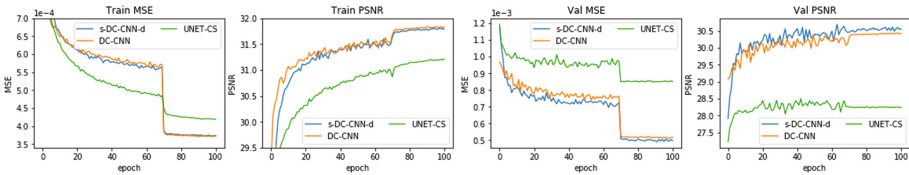
Fig. 1. Convergence analysis of the s-DC-CNN-d compared to the DC-CNN and the UNET-CS.

Convergence analysis (Fig. 1) shows that the proposed s-DC-CNN-d learned slower, but eventually generalised better. The fact that UNET-CS was quickly overfitted may be due to low SNR in the original images that hamper the network to learn a meaningful end-to-end mapping. Compared to the DC-CNN, the proposed s-DC-CNN-d obtained a better PSNR with less epochs during the validation that can be attributed to the fact that the proposed s-DC-CNN-d has an accelerated learning.

Qualitative visualisation (Fig. 2) demonstrates that perceptually both DLMRI and the UNET-CS are over-smoothed in their reconstructed results with clearly larger errors, and the textural details were better preserved in the results
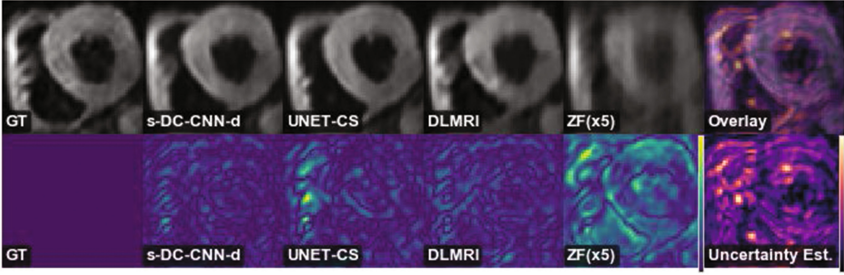
Fig. 2. Qualitative comparison of the reconstructions and error maps for the 5×
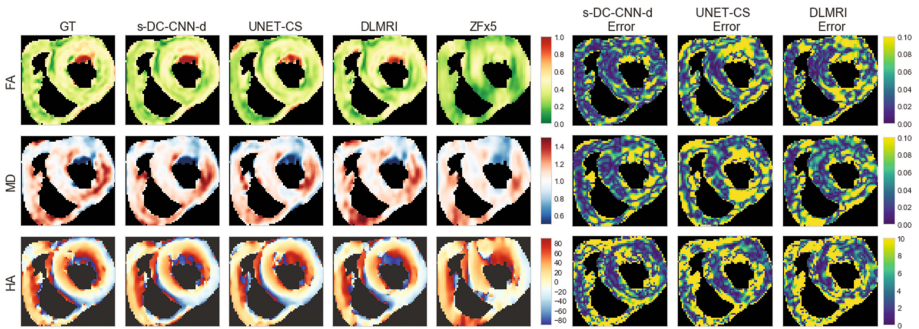undersampling.

Fig. 3. Comparison of the diffusion tensor parameters and error maps. From top to
bottom: fractional anisotropy (FA), mean diffusivity (MD) $(10^{-3}\text{mm}^2\text{s}^{-1})$ and Helix-
angle (HA) (degrees).

obtained by the proposed s-DC-CNN-d. From a qualitative analysis, we found
that the generated uncertainty estimate using our s-DC-CNN-d highlighted the
most challenging area for our algorithms to reconstruct, e.g., the edges of the
ventricle and the highly blurred areas in the undersampled input.

Figure 3 shows the computed FA, MD and HA for the 5× undersampling.
Overall, our proposed s-DC-CNN-d achieved better calculated diffusion tensor
parameters compared to the ones obtained by other methods. In particular,
results obtained by the s-DC-CNN-d have a more smooth transition of the HA
from epicardial to endocardial surface in the normal left ventricular free wall
that resembles the HA calculated by the fully sampled ground truth data.

Finally, we timed the reconstruction of each algorithm (on a GeForce GTX
1080 GPU) and obtained (1) s-DC-CNN-d: 0.065 ± 0.03 s per frame, (2)
s-DC-CNN: 0.04 ± 0.02 s per frame, (3) DC-CNN-d: 0.052 ± 0.02 s per frame,
(4) DC-CNN: 0.04 ± 0.02 s per frame and (5) UNET-CS: 0.003 ± 0.04 s per
frame. The DLMRI method has only a CPU implementation and takes ≈60 s
per iteration, and we used 400 iterations per image (>6 h).

4 Discussion and Conclusion

In this study, we proposed a novel deep cascaded CNN based CS-MRI and its stochastic variation for the DT-CMR reconstruction. To the best of our knowledge, it is the first work to consider such application to DT-CMR although the current study is nevertheless simulation based. In addition, we are the first to consider such a stochastic formulation to take the variance of the reconstruction models into account that can be visualised via an uncertainty map. Compared to other state-of-the-art methods using dictionary learning or U-Net based deep learning architecture, our proposed method incorporated dilated convolution and achieved significantly superior reconstruction fidelity with very efficient computation that can be translated into a real-time reconstruction scheme working directly on the scanner. As a future direction, we will carry on further development from on our current simulation based study to accelerate DT-CMR reconstruction and improve its spatial resolution and we can envisage its deployment in various clinical applications.

References

1. Axel, L.: Faster diffusion-weighted MR imaging of cardiac microstructure. Radiology **282**(3), 622–626 (2017)
2. von Deuster, C.: Studying dynamic myofiber aggregate reorientation in dilated cardiomyopathy using in vivo magnetic resonance diffusion tensor imaging. Circ. Cardiovasc. Imaging **9**(10), e005018 (2016)
3. Ferreira, P.F., et al.: In vivo cardiovascular magnetic resonance diffusion tensor imaging shows evidence of abnormal myocardial laminar orientations and mobility in hypertrophic cardiomyopathy. J. Cardiovasc. Magn. Reson. **16**, 87 (2014)
4. Hammernik, K., et al.: Learning a variational network for reconstruction of accelerated MRI data. arXiv preprint arXiv:1704.00447 (2017)
5. Han, Y.: Magn. Reson. Med. Deep learning with domain adaptation for accelerated projection-reconstruction MR **80**(3), 1189–1205 (2018)
6. Hollingsworth, G.: Phys. Med. Biol. Reducing acquisition time in clinical MRI by data undersampling and compressed sensing reconstruction **60**(21), 297–322 (2015)
7. Huang, G., et al.: Deep networks with stochastic depth. arXiv preprint arXiv:1603.09382 (2017)
8. Huang, J.: Cardiac diffusion tensor imaging based on compressed sensing using joint sparsity and low-rank approximation. Technol. Health Care **24**(s2), S593–S599 (2016)
9. Lau, A.Z., et al.: Accelerated human cardiac diffusion tensor imaging using simultaneous multislice imaging. Magn. Reson. Med. **73**(3), 995–1004 (2015)
10. Lee, D., et al.: Deep residual learning for compressed sensing MRI. In: International Symposium on Biomedical Imaging, pp. 15–18. IEEE (2017)
11. Lustig, M.: Sparse MRI: the application of compressed sensing for rapid MR imaging. Magn. Reson. Med. **58**(6), 1182–1195 (2007)
12. Ma, S., et al.: Accelerated cardiac diffusion tensor imaging using joint low-rank and sparsity constraints. IEEE Trans. Biomed. Eng. (2017)
13. Mekkaoui, C.: Diffusion MRI in the heart. NMR Biomed. **30**(3), e3426 (2017)

14. Nielles-Vallespin, S., et al.: Assessment of myocardial microstructural dynamics by in vivo diffusion tensor cardiac magnetic resonance. J. Am. Coll. Cardiol. **69**(6), 661–676 (2017)
15. Nielles-Vallespin, S., et al.: In vivo diffusion tensor MRI of the human heart: reproducibility of breath-hold and navigator-based approaches. Magn. Reson. Med. **70**(2), 454–65 (2013)
16. Qin, C., et al.: Convolutional recurrent neural networks for dynamic MR image reconstruction. arXiv preprint arXiv:1712.01751 (2017)
17. Ravishankar, S., Bresler, Y.: MR image reconstruction from highly undersampled k-space data by dictionary learning. IEEE Trans. Med. Imaging **30**(5), 1028–1041 (2011)
18. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
19. Schlemper, J.: A deep cascade of convolutional neural networks for dynamic MR image reconstruction. IEEE Trans. Med. Imaging **37**(2), 491–503 (2018)
20. Schlemper, J., Caballero, J., Hajnal, J.V., Price, A., Rueckert, D.: A deep cascade of convolutional neural networks for MR image reconstruction. In: Niethammer, M. (ed.) IPMI 2017. LNCS, vol. 10265, pp. 647–658. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59050-9_51
21. Sun, J., et al.: Deep ADMM-Net for compressive sensing MRI. In: NIPS, pp. 10–18 (2016)
22. Wu, Y., et al.: Accelerated MR diffusion tensor imaging using distributed compressed sensing. Magn. Reson. Med. **71**(2), 763–772 (2014)
23. Yang, G., et al.: DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. IEEE Trans. Med. Imaging, 1310–1321 (2018)