Over-and-Under Complete Convolutional RNN for MRI Reconstruction

Pengfei Guo[1], Jeya Maria Jose Valanarasu[2], Puyang Wang[2], Jinyuan Zhou[3], Shanshan Jiang[3], Vishal M. Patel[1,2]

[1] Department of Computer Science, Johns Hopkins University, MD, USA
pguo4@jhu.edu
[2] Department of Electrical and Computer Engineering, Johns Hopkins University, MD, USA
[3] Department of Radiology, Johns Hopkins University, Baltimore, MD, USA

Abstract. Reconstructing magnetic resonance (MR) images from under-sampled data is a challenging problem due to various artifacts introduced by the under-sampling operation. Recent deep learning-based methods for MR image reconstruction usually leverage a generic auto-encoder architecture which captures low-level features at the initial layers and high-level features at the deeper layers. Such networks focus much on global features which may not be optimal to reconstruct the fully-sampled image. In this paper, we propose an **O**ver-and-**U**nder **C**omplete Convolutional **R**ecurrent Neural Network (OUCR), which consists of an overcomplete and an undercomplete Convolutional Recurrent Neural Network (CRNN). The overcomplete branch gives special attention in learning local structures by restraining the receptive field of the network. Combining it with the undercomplete branch leads to a network which focuses more on low-level features without losing out on the global structures. Extensive experiments on two datasets demonstrate that the proposed method achieves significant improvements over the compressed sensing and popular deep learning-based methods with less number of trainable parameters.

Keywords: Convolutional RNN · MRI reconstruction · Deep Learning.

1 Introduction

Magnetic resonance imaging (MRI) is a noninvasive medical imaging approach that provides various tissue contrast mechanisms for visualizing anatomical structures and functions. Due to the hardware constraint, one major limitation of MRI is relatively slow data acquisition process, which subsequently causes higher imaging cost and patients' discomfort in many clinical applications [3]. While the exploitation of advanced hardware and parallel imaging [20] can mitigate such issue, a common approach is to shorten the image acquisition time by under-sampling k-space (also known as Compressed Sensing (CS)) [9, 19]. However, reconstructing an image directly from partial k-space data results in a subop-timal image with aliasing artifacts. To deal with this issue, nonlinear recovery

algorithms based on ℓ_0, ℓ_1 or total variation minimization are often used to recover the image from incomplete k-space data. More advanced CS-based image reconstruction algorithms have been combined with parallel imaging [15], low-rank constraint terms [16], and dictionary learning [18]. Unfortunately, though CS image reconstruction algorithms are able to recover images, they lack noise-like textures and as a result many physicians find CS reconstructed images as "artificial". Moreover, when large errors are not reduced during optimization, high-frequency oscillatory artifacts cannot be properly removed [23]. Thus, the acceleration factors of CS-based algorithms are generally limited between 2.5 and 3 for typical MR images [23].

Recent advances in deep neural networks open a new possibility to solve the inverse problem of MR image reconstruction in an efficient manner [8]. Artificial neural network-based image reconstruction methods have been shown to provide much better MR image quality than conventional CS-based methods [1, 2, 4, 7, 12, 14, 21, 25, 31]. Most DL-based methods are convolutional which consist of a set of convolution and down/up sampling layers for efficient learning. The main intuition behind this kind of network architecture is that at the initial layers, the receptive field of the filters is smaller, so low-level features (e.g., edges) are captured. At deeper layers, the receptive field of the filters is larger, so high-level features (e.g., the interpretation of input) are captured. Using such generic architecture for MR image reconstruction might not be optimal, since low-level vision tasks are mainly concerned with extracting descriptions from input rather than the interpretation of input [5]. Prior to DL era, overcomplete representations were explored for dealing with noisy observations in the vision and image processing tasks [13, 30]. In an overcomplete architecture, the increase of receptive field is restricted through the network, which forces the filters to focus on low-level features [32]. Recently, overcomplete representations have been explored for image segmentation [29] and image restoration [32].

In this paper, a novel **O**ver-and-**U**nder **C**omplete Convolutional **R**ecurrent Neural Network (OUCR) is proposed that can recover a fully-sampled image from the under-sampled k-space data for accelerated MR image reconstruction. To summarize, the following are our key contributions: **1.** An over-complete convolutional RNN architecture is explored for MR image reconstruction. **2.** To recover finer details better, the proposed OUCR consists of two branches that can leverage the features of both undercomplete and overcomplete CRNN. **3.** Extensive experiments are conducted on two datasets and it is demonstrated that the proposed method achieves significant improvements over CS-based as well as popular DL-based methods and is more parameter efficient.

2   Methodology

Overcomplete Networks. Overcomplete representations were first explored for signal representation where overcomplete bases were used such that the number of basis functions are more than the input signal samples [13]. This enabled a high flexibility at expressing the structure of the data. In neural networks,
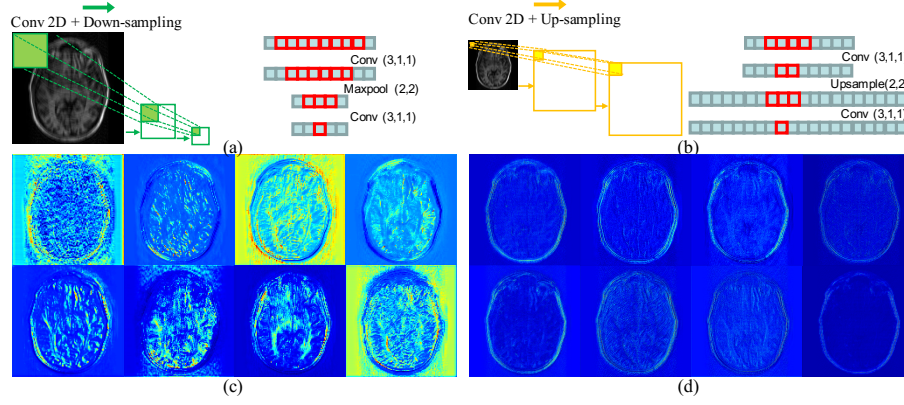
Fig. 1. *Top row*: Explanation for the receptive field change in (a) undercomplete and (b) overcomplete networks. Conv(3,1,1) represents a convolution layer with kernel size as 3, stride as 1 and padding as 1. Upsample (2,2) represents a nearest neighbor upsampling layer with upsampling coefficients as (2,2). Maxpool(2,2) represents maxpooling layer with kernel size as 2 and stride as 2. The red pixels denote receptive field. It can be noted that the receptive field is constrained in overcomplete network compared to undercomplete network. *Bottom row*: Visualization of filter responses (feature maps from ResBlock) for (c) undercomplete and (d) overcomplete CRNN. By restricting the size of receptive field, the overcomplete CRNN is able to focus on low-level features.

overcomplete fully connected networks were observed to be better feature extractors and hence perform well at denoising [30]. Recently, overcomplete convolutional networks are noted to be better at extracting local features because of restraining the receptive field when compared to the generic encoder-decoder networks [27–29, 32]. Rather than using max-pooling layers at the encoder like an undercomplete convolutional network, in an overcomplete convolutional network, one uses upsampling layers. Fig. 1(a) and (b) visually explains this concept in CNNs. As can be seen from Fig. 1(c) and (d), the overcomplete networks focus on fine structures and learning local features from an image as the receptive field is constrained even at deeper layers. More examples of comparison between over and under complete networks can be found in the supplementary material.

MR Image Reconstruction. Let $x \in \mathbb{C}^N$ denote the observed under-sampled k-space data, $y \in \mathbb{C}^M$ is the fully-sampled image that we want to reconstruct. To obtain a regularized solution, the optimization problem of MRI reconstruction can be formulated as follows [22, 25]:

$$\min_{y} R(y) + \lambda \|x - F_D y\|_2^2. \tag{1}$$

Here, $R(y)$ is the regularization term and λ controls the contribution of second term. F_D represents the undersampling Fourier encoding matrix that is defined as the multiplication of the Fourier transform matrix with a binary undersampling mask D. The ratio of the amount of k-space data required for a fully-sampled image to the amount collected in an accelerated acquisition is controlled by the acceleration factor (AF). The approximate fully-sampled image $\bar{y}$
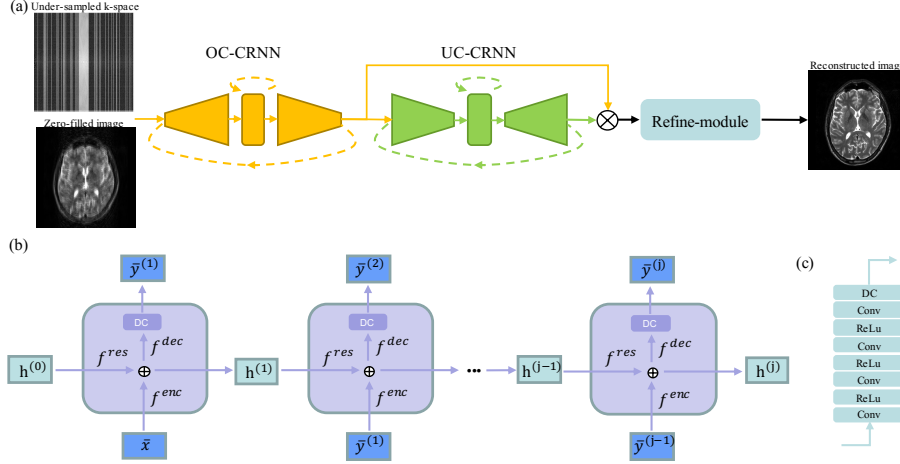
Fig. 2. (a) An overview of the proposed OUCR. Here, $\otimes$ denotes the channel-wise concatenation. We denote overcomplete and undercomplete CRNN as OC-CRNN and UC-CRNN, respectively. (b) A schematic of unrolled CRNN iterations. Here, $\oplus$ denotes the element-wise addition. (c) The network configuration of the refine-module. DC represents the data consistency layer.

can be measured from the observed under-sampled k-space data x via an optimization process. One can solve the objective function (Eq. 1) based on iterative optimization methods, such as gradient descent. A CRNN [22, 31] is capable of modeling the iterative optimization process in Eq. 1 as $\bar{y} = \text{CRNN}(\bar{x}, x, D, \Theta)$, where $\bar{y}$ is the reconstructed MR image from CRNN model, $\bar{x}$ is the zero-filled image and Θ denotes the trainable parameters of the CRNN model.

OUCR. To have special attention in learning low-level feature structures while not losing out on the global structures and inspired by previous CRNN methods [22, 31], we propose a novel OUCR network as shown in Fig. 2. OUCR consists of two CRNN modules with different receptive fields to reconstruct MR images (i.e. OC-CRNN and UC-CRNN in Fig.2(a)) and an refine-module (RM). In a CRNN module f_i, let f_i^{enc}, f_i^{dec}, and f_i^{res} denote the encoder, decoder, and ResBlock, respectively (Fig.2(b)). A data consistency (DC) layer is added at the end of each module to reinforce the data consistency in k-space. The iterations of a CRNN module can be unrolled as follows:

$$
\begin{aligned}
\bar{y}_i^{(j+1)} &= \text{DC}(f_i(\bar{y}_i^{(j)}, h_i^{(j)}), x, D), \\
&= F^{-1}[Dx + (1 - D)F[f_i(\bar{y}_i^{(j)}, h_i^{(j)})]], \\
&= F^{-1}[Dx + (1 - D)F[f_i^{\text{dec}}(f_i^{\text{res}}(h_i^{(j)}) + f_i^{\text{enc}}(\bar{y}_i^{(j)}))]],
\end{aligned}
\tag{2}
$$

where h_i^j is the hidden state after iteration j and F^{-1} denotes the inverse Fourier transform. After J iterations of the two CRNN modules, the final reconstructed MR image $\bar{y}$ is formulated as follows:

$$
\begin{aligned}
\bar{y}_{\text{oc}} &= \text{OC-CRNN}(\bar{x}, x, D, \Theta_{\text{oc}}), \\
\bar{y}_{\text{uc}} &= \text{UC-CRNN}(\bar{y}_{\text{oc}}, x, D, \Theta_{\text{uc}}), \\
\bar{y} &= \text{DC}(\text{RM}(\bar{y}_{\text{oc}} \otimes \bar{y}_{\text{uc}}, \Theta_{\text{rm}}), x, D),
\end{aligned}
\tag{3}
$$

where $\otimes$ denotes the channel-wise concatenation and $\Theta_{\mathrm{oc}}, \Theta_{\mathrm{uc}}, \Theta_{\mathrm{rm}}$ denote the parameters of the overcomplete, undercomplete CRNN and RM network, respectively. The intuition behind using both OC-CRNN and UC-CRNN is to make use of both local and global features. While we focus more on the local features using the OC-CRNN, the global features are not neglected altogether as they still have meaningful information for proper reconstruction. In each CRNN module, we have two convolutional blocks in both encoder and decoder. Each convolutional block in the encoder has a 2D convolutional layer followed by an upsampling layer in OC-CRNN or a max-pooling layer in UC-CRNN. In the decoder, each convolutional block has a 2D convolutional layer followed by max-pooling layer in OC-CRNN or upsampling layer in UC-CRNN. More details regarding the network configuration can be found in the supplementary material.
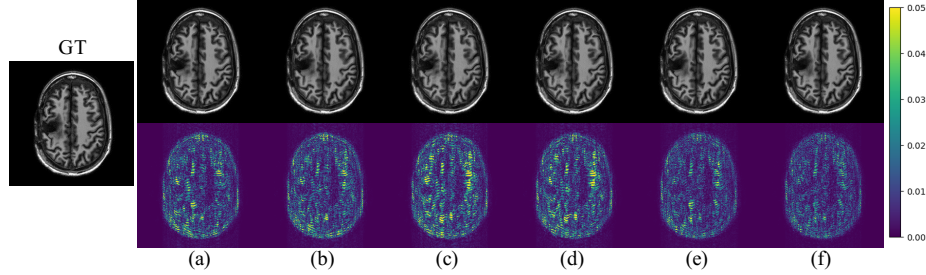
3   Experiments and Results

Evaluation and Implementation Details.   The following two datasets are used for conducting experiments – **fastMRI** [11] and **HPKS** [6, 10]. The **fastMRI** dataset consists of single-coil coronal proton density-weighted knee images corresponding to 1172 subjects. In particular, 973 subjects' data is used for training, and 199 subjects' data (fastMRI validation dataset) is used for testing. For each subject, there are approximately 35 knee images that contain tissues. The **HPKS** dataset is collected by an anonymous medical center from post-treatment patients with malignant glioma. T_1-weighted images from 144 subjects are used, where 102 subjects' data are used for training, 14 subjects' data set are used for validation, and 28 subjects' data are used for testing. For each subject, 15 axial cross-sectional images that contain brain tissues are provided in this dataset. We simulated k-space measurements using the same sampling mask function as the fastMRI challenge [11] with $4\times$ and $8\times$ accelerations. All models were trained using the ℓ_1 loss with Adam optimizer by the following hyperparameters: initial learning rate of 1.5×10^{-4} then reduced by a factor of 0.9 every 5 epochs; 50 maximum epochs; batch size of 4; the number of CRNN iteration J of 5. SSIM and PSNR are used as the evaluation metrics for comparison.

MR Image Reconstruction Results.   Table 1 shows the results corresponding to seven different methods evaluated on the HPKS and fastMRI datasets. The performance of the proposed model was compared with compressed sensing (CS) [26], UNet [24], KIKI-Net [4], Kiu-net [29], D5C5 [25], and PC-RNN [31]. For a fair comparison, UNet [24] and Kiu-net [29] are modified for data with real and imaginary channels and a DC layer is added at the end of the networks. KIKI-Net [4] which conducts interleaved convolution operation on image and k-space domains achieves better performance than UNet [24] on HPKS. Kiu-net [29] which is an overcomplete variant architecture of UNet [24] outperforms UNet [24] and KIKI-Net [4]. By leveraging the cascade of convolutional neural networks, D5C5 [25] outperforms the Kiu-net [29] in both HPKS and fastMRI datasets. PC-RNN [31] which learns the mapping in an iterative way by CRNN from three different scales achieves the second best performance. As it can be

Table 1. Quantitative results on the HPKS and fastMRI dataset. Param denotes the number of parameters.

Method	Param	HPKS				fastMRI			
		PSNR		SSIM		PSNR		SSIM	
		4X	8X	4X	8X	4X	8X	4X	8X
CS	-	29.94	24.96	0.8705	0.7125	29.54	26.99	0.5736	0.4870
UNet	8.634 M	34.47	29.47	0.9155	0.8249	31.88	29.78	0.7142	0.6424
KIKI-Net	1.790 M	35.35	29.86	0.9363	0.8436	31.87	29.27	0.7172	0.6355
Kiu-net	7.951 M	35.35	30.18	0.9335	0.8467	32.06	29.86	0.7228	0.6456
D5C5	2.237 M	37.51	30.40	0.9595	0.8623	32.25	29.65	0.7256	0.6457
PC-RNN	1.482 M	38.36	31.54	0.9696	0.8965	32.37	30.17	0.7281	0.6585
OUCR	**1.192 M**	**39.33**	**32.14**	**0.9747**	**0.9044**	**32.61**	**30.59**	**0.7354**	**0.6634**

seen from Table 1, the proposed OUCR outperforms other methods by leveraging the overcomplete architecture. Fig. 4 shows the qualitative results of two datasets with $4\times$ and $8\times$ accelerations. It can be observed that the proposed OUCR yields reconstructed images with remarkable visual similarity to the reference images compared to the others (see the last column of each sub-figure in Fig. 4) in two datasets with different modalities. The reported improvements achieved by OUCR are statistically significant ($p < 10^{-5}$). The computational details of different methods and statistical significance test are provided in Supplementary Table 1 and 2, respectively.

Fig. 3. Qualitative results and error maps of ablation study on HPKS with AF=4. (a) UC-CRNN. (b) UC-CRNN + RM. (c) OC-CRNN. (d) OC-CRNN + RM. (e) UC-CRNN + OC-CRNN. (f) UC-CRNN + OC-CRNN + RM (proposed OUCR).

In the k-space domain, the center frequencies determine the overall image contrast, brightness, and general shape. The peripheral area of k-space contains high spatial frequency information that controls edges, details, sharp transitions. [17]. To further analyze the performance of different methods in low and high spatial frequency, we carry out the k-space analysis in Fig. 5. We reconstruct an MR image from partially masked k-space and compare it with the reference image that is applied same mask on k-space, as shown in Fig. 5 top row. The reported PSNR and SSIM are presented by Boxplot in Fig. 5 bottom row. It can be seen that the proposed OUCR exhibits better reconstruction performance than other methods for both low and high frequency information.

Ablation Study. We conduct a detailed ablation study to separately evaluate the effectiveness of using OC-CRNN, UC-CRNN, and RM in the proposed
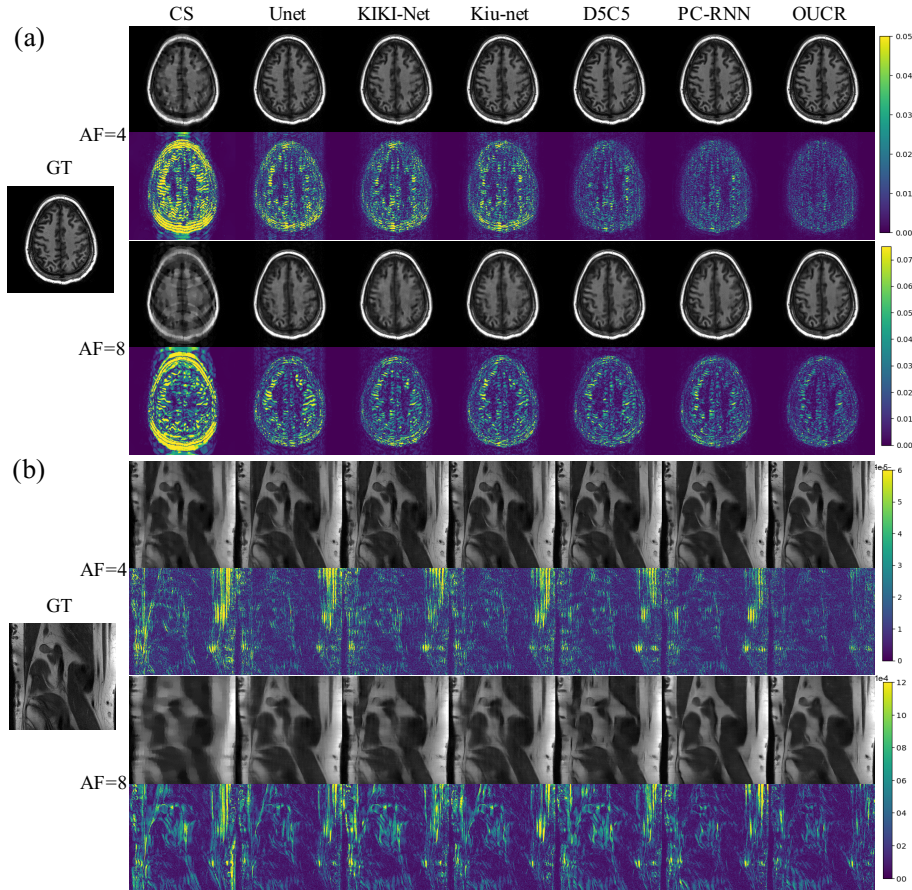
Fig. 4. Qualitative comparison of different methods on (a) HPKS and (b) fastMRI dataset. The second row of each subplot shows the corresponding error maps.

Table 2. Ablation study of designed modules in term of reconstruction quality on HPKS with AF=4.

Modules	PSNR	SSIM
UC-CRNN	37.87	0.9644
UC-CRNN + RM	38.02	0.9658
OC-CRNN	36.97	0.9564
OC-CRNN + RM	37.34	0.9600
UC-CRNN + OC-CRNN	38.98	0.9719
UC-CRNN + OC-CRNN + RM (OUCR)	**39.33**	**0.9747**

framework. The results are shown in Table 2. We start with only using OC-CRNN and UC-CRNN. It can be noted that the performance of OC-CRNN is lesser than UC-CRNN, since even though OC-CRNN captures the low-level features properly it does not capture most high-level features like UC-CRNN. Then, we show that adding RM with each individual module can improve the
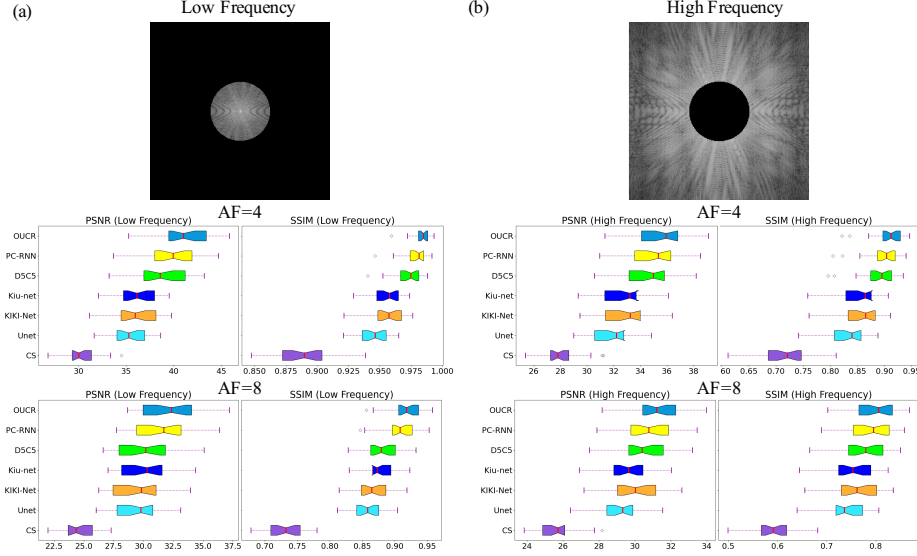
Fig. 5. K-space analysis on (a) low frequency and (b) high frequency. *Top row*: the examples of masked k-space image. *Bottom row*: the Boxplot of reconstruction performance by different methods on HPKS dataset.

reconstruction quality. Finally, combing both CRNN networks with RM (proposed OUCR) results in the best performance. Fig. 3 illustrates the qualitative improvements after adding each major block, which is consistent with the results reported in Table 2. Moreover, we observe that increasing the number of CRNN iterations can further improve the performance of the proposed OUCR, but consequently leads to lower computational efficiency. Due to space constraint, an ablation study regarding CRNN iterations, k-space analysis on fastMRI dataset, and more visualizations are provided in the supplementary material.

4   Discussion and Conclusion

We proposed a novel over-and-under complete convolutional RNN (OUCR) for MR image reconstruction. The purposed method leverages an overcomplete network to specifically capture low-level features, which are typically missed out in the other MR image reconstruction methods. Moreover, we incorporate an undercomplete CRNN, which results in an effective learning of low and high level information. The proposed method achieves better performance on two datasets and has less numbers of trainable parameters as compared to the CS and popular DL-based methods, including UNet [24], KIKI-Net [4], Kiu-net [29], D5C5 [25], and PC-RNN [31]. This study demonstrates the potential of using overcomplete networks in MR image reconstruction task.

References

1. Akçakaya, M., Moeller, S., Weingärtner, S., Uğurbil, K.: Scan-specific robust artificial-neural-networks for k-space interpolation (raki) reconstruction: Database-free deep learning for fast imaging. Magnetic resonance in medicine **81**(1), 439–453 (2019)

2. Chen, E.Z., Chen, T., Sun, S.: Mri image reconstruction via learning optimization using neural odes. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 83–93. Springer (2020)

3. Edmund, J.M., Nyholm, T.: A review of substitute ct generation for mri-only radiation therapy. Radiation Oncology **12**(1), 1–15 (2017)

4. Eo, T., Jun, Y., Kim, T., Jang, J., Lee, H.J., Hwang, D.: Kiki-net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images. Magnetic resonance in medicine **80**(5), 2188–2201 (2018)

5. Fisher, R.B.: Cvonline: The evolving, distributed, non-proprietary, on-line compendium of computer vision. Retrieved January 28, 2006 from http://homepages. inf. ed. ac. uk/rbf/CVonline (2008)

6. Guo, P., Wang, P., Yasarla, R., Zhou, J., Patel, V.M., Jiang, S.: Anatomic and molecular mr image synthesis using confidence guided cnns. IEEE Transactions on Medical Imaging pp. 1–1 (2020). https://doi.org/10.1109/TMI.2020.3046460

7. Guo, P., Wang, P., Zhou, J., Jiang, S., Patel, V.M.: Multi-institutional collaborations for improving deep learning-based magnetic resonance image reconstruction using federated learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2423–2432 (2021)

8. Guo, P., Wang, P., Zhou, J., Patel, V.M., Jiang, S.: Lesion mask-based simultaneous synthesis of anatomic and molecular mr images using a gan. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 104–113. Springer (2020)

9. Haldar, J.P., Hernando, D., Liang, Z.P.: Compressed-sensing mri with random encoding. IEEE transactions on Medical Imaging **30**(4), 893–903 (2010)

10. Jiang, S., et al.: Identifying recurrent malignant glioma after treatment using amide proton transfer-weighted mr imaging: a validation study with image-guided stereotactic biopsy. Clinical Cancer Research **25**(2), 552–561 (2019)

11. Knoll, F., et al.: fastmri: A publicly available raw k-space and dicom dataset of knee images for accelerated mr image reconstruction using machine learning. Radiology: Artificial Intelligence **2**(1), e190007 (2020)

12. Lee, D., Yoo, J., Tak, S., Ye, J.C.: Deep residual learning for accelerated mri using magnitude and phase networks. IEEE Transactions on Biomedical Engineering **65**(9), 1985–1995 (2018)

13. Lewicki, M.S., Sejnowski, T.J.: Learning overcomplete representations. Neural computation **12**(2), 337–365 (2000)

14. Liang, D., Cheng, J., Ke, Z., Ying, L.: Deep magnetic resonance image reconstruction: Inverse problems meet neural networks. IEEE Signal Processing Magazine **37**(1), 141–151 (2020)

15. Liang, D., Liu, B., Wang, J., Ying, L.: 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)

16. Majumdar, A.: Improving synthesis and analysis prior blind compressed sensing with low-rank constraints for dynamic mri reconstruction. Magnetic resonance imaging **33**(1), 174–179 (2015)

17. Mezrich, R.: A perspective on k-space. Radiology **195**(2), 297–315 (1995)
18. Patel, V.M., Chellappa, R.: Sparse representations, compressive sensing and dictionaries for pattern recognition. In: The first Asian conference on pattern recognition. pp. 325–329. IEEE (2011)
19. Patel, V.M., Maleh, R., Gilbert, A.C., Chellappa, R.: Gradient-based image recovery methods from incomplete fourier measurements. IEEE Transactions on Image Processing **21**(1), 94–105 (2011)
20. Pruessmann, K.P., Weiger, M., Scheidegger, M.B., Boesiger, P.: Sense: sensitivity encoding for fast mri. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **42**(5), 952–962 (1999)
21. Putzky, P., Welling, M.: Invert to learn to invert. arXiv preprint arXiv:1911.10914 (2019)
22. Qin, C., et al.: Convolutional recurrent neural networks for dynamic mr image reconstruction. IEEE Transactions on Medical Imaging **38**(1), 280–290 (2019). https://doi.org/10.1109/TMI.2018.2863670
23. Ravishankar, S., Bresler, Y.: Mr image reconstruction from highly undersampled k-space data by dictionary learning. IEEE transactions on medical imaging **30**(5), 1028–1041 (2010)
24. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241. Springer (2015)
25. Schlemper, J., et al.: A deep cascade of convolutional neural networks for dynamic mr image reconstruction. IEEE transactions on Medical Imaging **37**(2), 491–503 (2017)
26. Tamir, J.I., Ong, F., Cheng, J.Y., Uecker, M., Lustig, M.: Generalized magnetic resonance image reconstruction using the berkeley advanced reconstruction toolbox. In: ISMRM Workshop on Data Sampling & Image Reconstruction, Sedona, AZ (2016)
27. Valanarasu, J.M.J., Patel, V.M.: Overcomplete deep subspace clustering networks. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 746–755 (2021)
28. Valanarasu, J.M.J., Sindagi, V.A., Hacihaliloglu, I., Patel, V.M.: Kiu-net: Overcomplete convolutional architectures for biomedical image and volumetric segmentation. arXiv preprint arXiv:2010.01663 (2020)
29. Valanarasu, J.M.J., Sindagi, V.A., Hacihaliloglu, I., Patel, V.M.: Kiu-net: Towards accurate segmentation of biomedical images using over-complete representations. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 363–373. Springer (2020)
30. Vincent, P., Larochelle, H., Bengio, Y., Manzagol, P.A.: Extracting and composing robust features with denoising autoencoders. In: Proceedings of the 25th international conference on Machine learning. pp. 1096–1103 (2008)
31. Wang, P., Chen, E.Z., Chen, T., Patel, V.M., Sun, S.: Pyramid convolutional rnn for mri reconstruction. arXiv preprint arXiv:1912.00543 (2019)
32. Yasarla, R., Valanarasu, J.M.J., Patel, V.M.: Exploring overcomplete representations for single image deraining using cnns. IEEE Journal of Selected Topics in Signal Processing (2020)