Comparison of Neural Network Architectures for Physics-Driven Deep Learning MRI Reconstruction

Burhaneddin Yaman[1,2], Seyed Amir Hossein Hosseini[1,2], Steen Moeller[2] and Mehmet Akçakaya[1,2]
[1]Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN
[2]Center for Magnetic Resonance Research, University of Minnesota, Minneapolis, MN
Emails: {yaman013, hosse049, moell018, akcakaya}@umn.edu

***Abstract*—Machine learning techniques have recently received interest as a means of improving MRI reconstruction. In physics-driven machine learning approaches, the known forward encoding model is used for enforcing data consistency in an unrolled iterative regularized least squares reconstruction. A neural network, which may or may not share weights across different unrolled iterations, is used as the regularizer prior. In this study, we aim to compare several neural network architectures, namely U-Net, ResNet and DenseNet for such physics-driven reconstruction. The performance of these architectures are evaluated on the publicly available fastMRI knee database. Comparisons are made for uniform and random undersampling masks. The results indicate that a DenseNet regularization unit performs as well as the other strategies for both uniform and random undersampling patterns, even though it has considerably fewer trainable parameters.**

***Index Terms*—recurrent neural networks, MRI reconstruction, deep learning, parallel imaging**

I. Introduction

Magnetic resonance imaging (MRI) provides a noninvasive and radiation-free diagnostic tool in medical imaging. Due to the inherently lengthy process of data acquisition in MRI, accelerated MRI reconstruction has been an active area of research for many years. Parallel imaging [1, 2] and compressed sensing [3, 4] are widely used techniques for accelerated MRI reconstruction. During the image reconstruction from sparse measurements, various regularizers such as wavelets or total variation based priors are generally applied [5, 6].

Recently, deep learning (DL) has emerged as an alternative method for accelerated MRI reconstruction [7–9]. Several approaches use end-to-end training for mapping the undersampled image data with artifacts to artifact-free reconstruction [10]. However, these approaches do not exploit the information from the MRI encoding model. Alternative strategies either use reconstruction in the sensor domain [8, 11] or ensures data consistency with acquired measurements for solving an inverse problem to recover the desired image [9].

A principled strategy for solving imaging inverse problems using priors that utilize deep learning is based on the principle of unrolling, which turns an iterative algorithm into a deep neural network [12]. Such unrolled networks are referred to as "physics-driven" in the context of DL-MRI reconstruction [9, 13, 14]. The unrolled network consists of several steps each having a regularization and data consistency unit. Various different implementation strategies have been used for physics-driven DL-MRI reconstruction. In [9], a variational network (VN) is proposed to learn the entire reconstruction model and parameters. In this technique, the unrolled network combines a regularizer which employs a variational neural network and data consistency that uses proximal gradient descent at each step. The network parameters are not shared among the blocks in the VN algorithm, thus number of trainable parameters increases with the number of unrolled steps. Model based deep learning method (MoDL) is another method to learn the unrolled network reconstruction problem [14]. Unlike the VN algorithm, parameters are shared across the iterations of the unrolled networks in MoDL, thus the number of parameters does not increase as number of unrolled step increases. This feature is particularly important for medical imaging applications, where large datasets are not always readily available. Similar to VN algorithm, in MoDL technique each unrolled step in the network contains a regularizer block which employs a residual network (ResNet) [15] and a data consistency block that employs a conjugate gradient method, which is another difference from [9] that uses proximal gradient method.

In this work, we studied different neural network architectures for unrolled networks in physics-driven DL-MRI reconstruction. We utilized the known forward encoding model in data consistency blocks, which are implemented with an unrolled conjugate gradient approach [14]. Three different neural networks, namely ResNet [15], U-Net [16] and DenseNet [17], which have substantially different number of trainable network parameters were used as the regularizer prior. Network parameters were shared among the unrolled iterations to reduce the number of unknown parameters within the network. These methods were trained and tested on the publicly available fastMRI knee datasets [18].

II. Materials and Methods

A. Data Modelling

Let $\mathbf{x}$ denote the image to be recovered, and $\mathbf{y}$ be the acquired undersampled noisy data. The forward model for this system is given as

$$\mathbf{y} = \mathbf{E}\mathbf{x} + \mathbf{n}, \tag{1}$$

This work was partially supported by NIH, Grant numbers: P41EB015894, P41EB027061; NSF, Grant number: CAREER CCF-1651825. The first two authors contributed equally to this work.

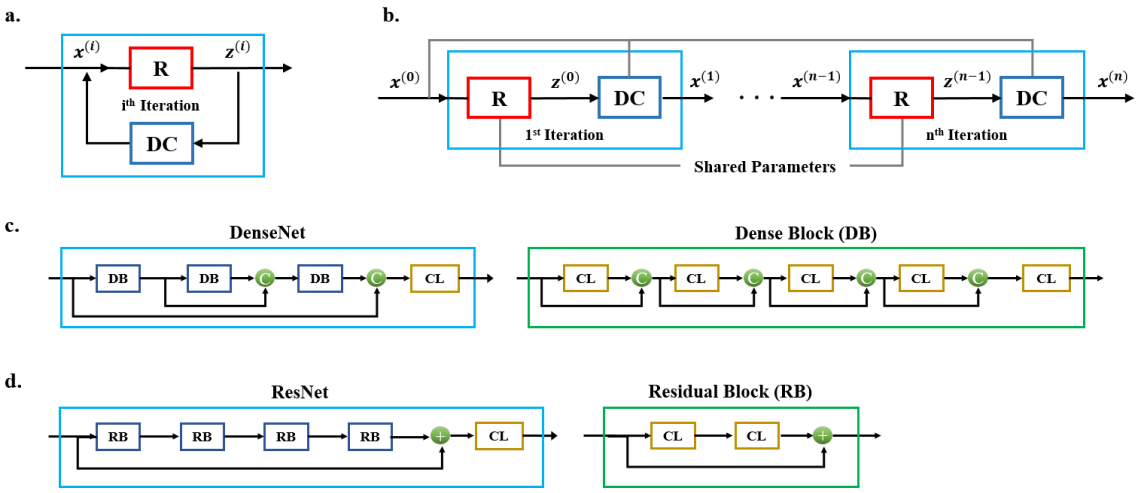

Fig. 1: a) The iterative scheme of a reconstruction problem. b) The recurrent neural network architecture with each step consisting of a regularization (R) and a data consistency (DC) unit. c) The DenseNet architecture with dense blocks (DB) containing convolutional layers (CL), and skip connections to improve the flow of information across blocks. d) The ResNet architecture consisting of residual blocks (RB) that contain convolutional layers (CL), and skip connections.

where $\mathbf{E} : \mathbb{C}^{M\times N} \rightarrow \mathbb{C}^{P}$ is the encoding operator, including a partial Fourier matrix and the sensitivities of the receiver coil array [2], and $\mathbf{n} \in \mathbb{C}^{P}$ is the measurement noise. When $P > M \cdot N$, $\mathbf{x}$ can be estimated from $\mathbf{y}$ using a least squares approach [19]. However, in general, the system may be ill-conditioned and a least squares approach may lead to noise amplification. Thus, regularizers are generally applied during the recovery of $\mathbf{x}$, as in the following objective function:

$$\arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{E}\mathbf{x}\|_2^2 + \mathcal{R}(\mathbf{x}), \tag{2}$$

where $\mathcal{R}(\mathbf{x})$ is a regularization term, and the first term ensures data consistency with measurements.

A variety of functions have been conventionally used for the regularization term such as sparsity or low-rank based priors depending on the problem of interest [3, 20, 21]. More recently, neural networks have emerged as an alternative for the regularization term that can learn a fully data-driven prior. The objective function (2) can be optimized via a variable splitting approach with quadratic penalty as follows:

$$\arg\min_{\mathbf{x},\mathbf{z}} \|\mathbf{y} - \mathbf{E}\mathbf{x}\|_2^2 + \mu\|\mathbf{x} - \mathbf{z}\|_2^2 + \mathcal{R}(\mathbf{z}), \tag{3}$$

where μ is the parameter of the quadratic penalty. By alternating the minimization over variables $\mathbf{x}$ and $\mathbf{z}$, the optimization problem of Equation (3) can be iteratively solved as:

$$\mathbf{z}^{(i-1)} = \arg\min_{\mathbf{z}} \mu\|\mathbf{x}^{(i-1)} - \mathbf{z}\|_2^2 + \mathcal{R}(\mathbf{z}) \tag{4a}$$

$$\mathbf{x}^{(i)} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{E}\mathbf{x}\|_2^2 + \mu\|\mathbf{x} - \mathbf{z}^{(i-1)}\|_2^2 \tag{4b}$$

where $\mathbf{x}^{(0)}$ is the initial image, often corresponding to the zero-filled undersampled k-space data, $\mathbf{z}^{(i)}$ is an intermediate variable and $\mathbf{x}^{(i)}$ is the output image at iteration i. Equation (4a) can be implicitly solved using a neural network architecture with convolutional kernels to find the solution at each iteration. Equation (4b) has a closed form solution as:

$$\mathbf{x}^{(i)} = (\mathbf{E}^H\mathbf{E} + \mu\mathbf{I})^{-1}(\mathbf{E}^H\mathbf{y} + \mu\mathbf{z}^{(i-1)}), \tag{5}$$

where $\mathbf{I}$ is the identity matrix and $(\cdot)^H$ is the conjugate transpose operator. Given the size of multi-coil MRI data, it is computationally more efficient to find the solution of Equation (4b) using iterative algorithms such as conjugate gradient [14], which eliminate the need for a matrix inversion. Conventionally, Equations (4a) and (4b), often referred to as regularization and data consistency units, are alternately solved until a stopping criterion was met, as shown in Figure 1a. However, with neural networks substituting the regularization unit, the recursive scheme of Figure 1a is unrolled into a recurrent neural network architecture with a fixed number of iterations [12], as in Figure 1b. The regularization units in the recurrent neural network architecture can generally be trained with different parameters, which can be redundant in practice. Therefore, for improved efficiency, the trainable parameters can be shared across the regularization units [13, 14].

B. Implementation Details

The recurrent neural network architecture was unrolled for 5 iterations to perform an end-to-end image reconstruction as shown in Figure 1b. In each iteration, the input is refined by a regularization unit, and then a data consistency (DC) operator is applied to ensure the iteration output image remains consistent with the acquired data. The data consistency unit employs the conjugate gradient method with 10 iterations [14]. For the regularization unit, three different neural network architectures, U-Net, ResNet and DenseNet, were utilized, .

For the DenseNet architecture, 3 blocks of densely connected CNNs were employed with outer skip connections to further facilitate the flow of information through the network as shown in Figure 1c. Each dense block (DB) consisted of 4 convolutional layers with a growth factor of 12 channels. The ResNet regularization unit was designed to have 4 residual blocks with the skip connections as shown in Figure 1d. Each residual block consisted of 2 convolutional layers each having 64 channels. For the U-Net architecture, we have employed

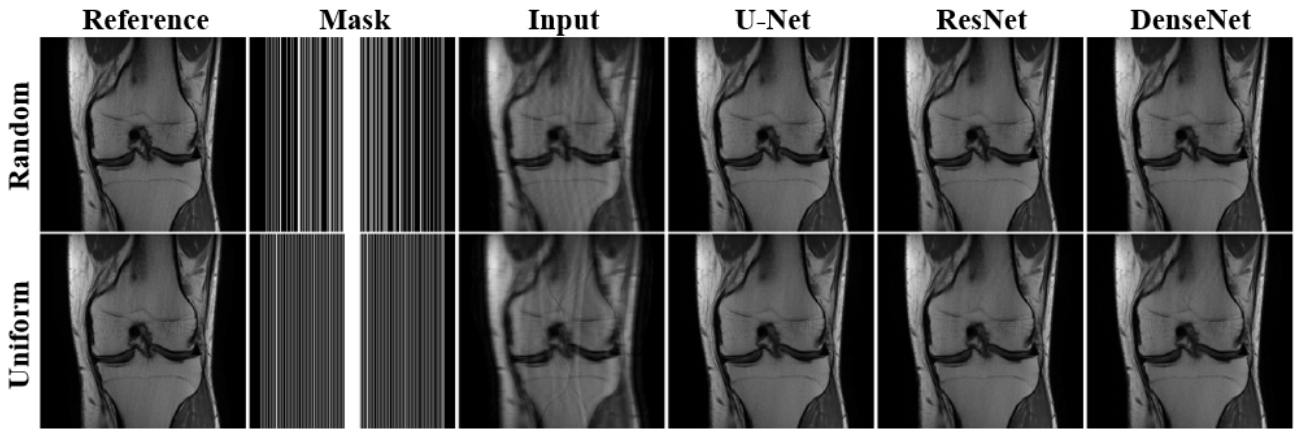

Fig. 2: A representative test slice reconstructed using the physics-driven DL-MRI method with the three neural network architectures, U-Net, ResNet and DenseNet, as regularization prior. Similar visual improvements were observed across the networks for both uniform and random undersampling masks.

a similar structure proposed in [18]. The total number of trainable parameterts for U-Net, ResNet and DenseNet were 3,540,065; 296,193; and 41,213, respectively.

All layers in the three architectures had a kernel size of 3×3, except the last layer that used a kernel size of 1×1 and was used to map the number of output channels to 2 as real and imaginary components of the reconstructed image. A zero-padding strategy was used at each convolutional layer to maintain input size. In addition, all layers were followed by rectifier linear units (ReLU) as activation functions except the last layer.

The magnitude of all k-space data were normalized between 0 and 1 as a preprocessing step, to ensure that the network parameters are learned for a standard range of data. The network was trained using an Adam optimizer with a learning rate of 10^{-3} by minimizing a normalized mean square error (NMSE) loss with a batch size of 2 for ResNet and DenseNet and a batch size of 1 for U-Net over 100 epochs. The batch sizes were kept small due to GPU memory limitations.

C. In Vivo MRI Dataset

Knee MRI data were obtained from the New York University (NYU) fastMRI initiative database [18]. Coronal proton density weighted data acquired from 10 subjects were utilized for training and testing. The fully sampled raw data were acquired with a 15-channel knee coil. The following imaging parameters were reported for coronal proton-weighted data; TR = 2750 ms, TE = 27 ms, echo-train length = 4, matrix size = 320×288, in-plane resolution = 0.49×0.44 mm^2, slice thickness = 3 mm. In the training, fully-sampled data from 10 subjects, corresponding to 381 slices were utilized. Each raw k-space data was of size $640 \times 368 \times 15$ where the first two dimensions are the oversampled matrix sizes and the last dimension denotes the number of coils. The testing was performed on 100 slices collected from 10 new subjects. The fully sampled raw data were undersampled retrospectively using random and uniform sampling patterns provided by NYU with an acceleration rate of 4. The center of kspace was fully sampled with 24 lines of autocalibrated signal (ACS). The coil sensititivity maps were estimated using ESPIRiT [22] using a 24×24 ACS region of the k-space. The reference and undersampled input images were obtained through SENSE-1 combination [2]. All the training and testing results were compared with the fully-sampled reference.

Experimental results were quantitatively evaluated using normalized mean square error (NMSE) and structural similarity index (SSIM). Statistical group analysis for these metrics over the 100 slices was performed using analysis of variance (ANOVA) method, comparing the three networks under uniform and random sampling scenarios. All experiments were performed using Tensorflow in Python, and processed on a workstation with an Intel E5-2640V3 CPU (2.6GHz and 256GB memory), and an NVIDIA Tesla V100 GPU with 32GB memory.

III. Results

Figure 2 shows a representative test slice reconstructed using the physics-driven DL-MRI method with the three neural network architectures, U-Net, ResNet and DenseNet, as regularization prior. All three networks exhibit similar visual improvement over the input image for both random and uniform undersampling cases. Stronger aliasing artifacts are present for the uniform undersampling, which is also the more clinically utilized protocol [2]. Even though DenseNet has substantially fewer parameters, it still performs as well as the other neural networks for regularizing the reconstruction. The quantitative evaluation of the results for the test slice in Figure 2 is tabulated in Table I. All three networks used in this study have shown comparable performances in both random and uniform undersampling scenarios.

Figure 3 summarizes the NMSE and SSIM metrics over the 100 test slices using the three neural network architectures. Experiments are presented for both random and uniform undersampling masks. The bar plots indicate similar performance among the networks, with the random sampling results slightly outperforming the uniform ones. This observation is confirmed with the ANOVA tests, which indicate no statistically

significant differences among the three networks for both undersampling scenarios and both metrics ($P > 0.05$).

IV. Discussion

In this study, we used unrolled recursive neural networks for solving a physics-driven DL-MRI reconstruction problem. Each unrolled network consisted of a data consistency unit and regularizer unit, which was implemented using a CNN. We compared three neural network architectures, U-Net, ResNet and DenseNet, as the regularization module of this recurrent neural network design. The comparisons were made for both random and uniform undersampling masks on knee MRI data [18]. The results indicate that all three network architectures successfully remove the aliasing artifacts for both random and uniform undersampling masks. The reconstruction of uniformly undersampled data is typically more difficult since the aliasing artifacts are stronger, and may remain in the reconstructed images. Nonetheless, all three networks were able to remove these artifacts in this study.

The DenseNet architecture had the fewest number of trainable parameters among the three architectures (approximately 85 and 7 times fewer compared to the U-Net and ResNet, respectively). Despite fewer trainable parameters, DenseNet results were similar to the other techniques in terms of both visual and quantitative assessment. The fewer number of parameters is advantageous, since it can reduce the risks of overfitting by reducing the degrees of freedom in training. Furthermore, a neural network with fewer number of trainable parameters requires smaller amount of data for training, which may be critical in medical imaging applications, where raw k-space data for training may be limited.

V. Conclusion

A DenseNet architecture is as effective as ResNet and U-Net architectures as the regularization prior for a physics-driven DL-MRI reconstruction approach. The DenseNet architecture has considerably fewer trainable parameters, which in turn reduces the risks of overfitting, as well as the required amount of data for training, which may be crucial in many MRI applications with scarce training data.

	Random			Uniform		
	U-Net	ResNet	DenseNet	U-Net	ResNet	DenseNet
NMSE	0.048	0.046	0.047	0.047	0.052	0.052
SSIM	0.955	0.956	0.956	0.956	0.951	0.951

TABLE I: The NMSE and SSIM metric values for the test slice in Figure 2 undersampled with random and uniform patterns and reconstructed with the three neural network architectures, namely U-Net, ResNet and DenseNet.

Acknowledgment

Knee MRI data were obtained from the NYU fastMRI initiative database [18]. NYU fastMRI investigators provided data but did not participate in analysis or writing of this report. A listing of NYU fastMRI investigators, subject to updates, can be found at fastmri.med.nyu.edu.

References

[1] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase, "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magn Reson Med*, vol. 47, no. 6, pp. 1202–1210, Jun 2002.

[2] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: sensitivity encoding for fast MRI," *Magn Reson Med*, vol. 42, no. 5, pp. 952–962, Nov 1999.

[3] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn Reson Med*, vol. 58, no. 6, pp. 1182–1195, Dec 2007.

[4] M. Akcakaya, T. A. Basha, B. Goddu, L. A. Goepfert, K. Kissinger, V. Tarokh, W. J. Manning, and R. Nezafat, "Low-dimensional-structure self-learning and thresholding: regularization beyond compressed sensing for MRI reconstruction," *Magn Reson Med*, vol. 66, no. 3, pp. 756–767, Sep 2011.

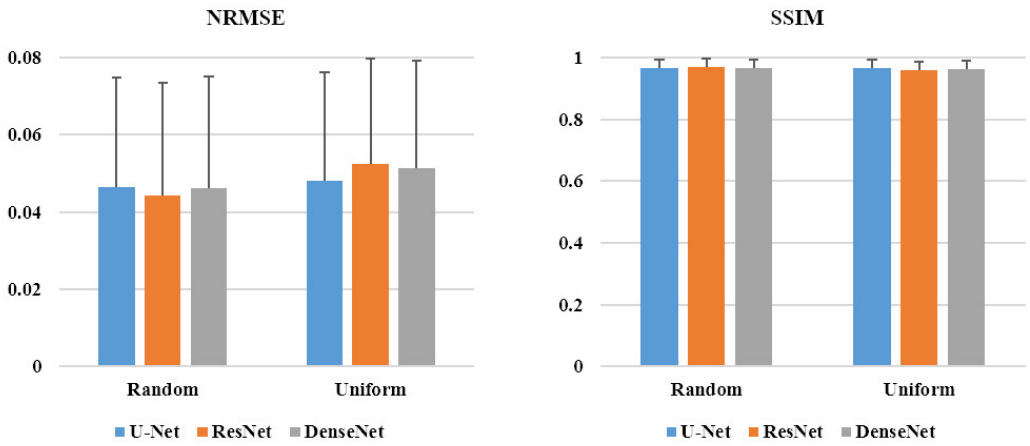

Fig. 3: The average NMSE (left panel) and SSIM (right panel) metric values over 100 test slices using the three neural network architecture, namely U-Net, ResNet and DenseNet. The experiments were performed for both random and uniform undersampling masks.

[5] K. T. Block, M. Uecker, and J. Frahm, "Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint," *Magn Reson Med*, vol. 57, no. 6, pp. 1086–1098, Jun 2007.

[6] M. Akcakaya, S. Nam, P. Hu, M. H. Moghari, L. H. Ngo, V. Tarokh, W. J. Manning, and R. Nezafat, "Compressed sensing with wavelet domain dependencies for coronary MRI: a retrospective study," *IEEE Trans Med Imaging*, vol. 30, no. 5, pp. 1090–1099, May 2011.

[7] S. Wang, Z. Su, L. Ying, X. Peng, S. Zhu, F. Liang, D. Feng, and D. Liang, "Accelerating magnetic resonance imaging via deep learning," in *Proc. IEEE Int. Symp. Biomed. Imag. (ISBI)*. IEEE, 2016, pp. 514–517.

[8] M. Akcakaya, S. Moeller, S. Weingartner, and K. Ugurbil, "Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging," *Magn Reson Med*, vol. 81, no. 1, pp. 439–453, 01 2019.

[9] K. Hammernik, T. Klatzer, E. Kobler, M. P. Recht, D. K. Sodickson, T. Pock, and F. Knoll, "Learning a variational network for reconstruction of accelerated MRI data," *Magn Reson Med*, vol. 79, no. 6, pp. 3055–3071, 06 2018.

[10] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, and G. Wang, "Low-Dose CT With a Residual Encoder-Decoder Convolutional Neural Network," *IEEE Trans Med Imaging*, vol. 36, no. 12, pp. 2524–2535, 12 2017.

[11] S. A. H. Hosseini, S. Moeller, S. Weingärtner, K. Uğurbil, and M. Akçakaya, "Accelerated coronary MRI using 3D SPIRiT-RAKI with sparsity regularization," in *Proc. ISBI*. IEEE, 2019, pp. 1692–1695.

[12] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proc. Int. Conf. Mach. Learn*. Omnipress, 2010, pp. 399–406.

[13] C. Qin, J. Schlemper, J. Caballero, A. N. Price, J. V. Hajnal, and D. Rueckert, "Convolutional Recurrent Neural Networks for Dynamic MR Image Reconstruction," *IEEE Trans Med Imaging*, vol. 38, no. 1, pp. 280–290, 01 2019.

[14] H. K. Aggarwal, M. P. Mani, and M. Jacob, "MoDL: Model-Based Deep Learning Architecture for Inverse Problems," *IEEE Trans Med Imaging*, vol. 38, no. 2, pp. 394–405, 02 2019.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2016, pp. 770–778.

[16] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Medical Image Comput. Comput.-Assisted Intervention*. Springer, 2015, pp. 234–241.

[17] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 4700–4708.

[18] J. Zbontar, F. Knoll, A. Sriram, et al., "fastMRI: An open dataset and benchmarks for accelerated MRI," *arXiv preprint arXiv:1811.08839*, 2018.

[19] K. P. Pruessmann, M. Weiger, P. Bornert, and P. Boesiger, "Advances in sensitivity encoding with arbitrary k-space trajectories," *Magn Reson Med*, vol. 46, no. 4, pp. 638–651, Oct 2001.

[20] S. G. Lingala, Y. Hu, E. DiBella, and M. Jacob, "Accelerated dynamic MRI exploiting sparsity and low-rank structure: k-t SLR," *IEEE Trans Med Imaging*, vol. 30, no. 5, pp. 1042–1054, May 2011.

[21] B. Yaman, S. Weingärtner, N. Kargas, N. D. Sidiropoulos, and M. Akcakaya, "Locally low-rank tensor regularization for high-resolution quantitative dynamic MRI," in *Proc. CAMSAP*. IEEE, 2017, pp. 1–5.

[22] M. Uecker, P. Lai, M. J. Murphy, P. Virtue, M. Elad, J. M. Pauly, S. S. Vasanawala, and M. Lustig, "ESPIRiT–an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA," *Magn Reson Med*, vol. 71, no. 3, pp. 990–1001, Mar 2014.