

IMPROVED SUPERVISED TRAINING OF PHYSICS-GUIDED DEEP LEARNING IMAGE RECONSTRUCTION WITH MULTI-MASKING

Burhaneddin Yaman^{†}, Seyed Amir Hossein Hosseini^{*†}, Steen Moeller[†] and Mehmet Akçakaya^{*†}*

^{*} Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN, USA

[†] Center for Magnetic Resonance Research, University of Minnesota, Minneapolis, MN, USA

ABSTRACT

Physics-guided deep learning (PG-DL) via algorithm unrolling has received significant interest for improved image reconstruction, including MRI applications. These methods unroll an iterative optimization algorithm into a series of regularizer and data consistency units. The unrolled networks are typically trained end-to-end using a supervised approach. Current supervised PG-DL approaches use all of the available sub-sampled measurements in their data consistency units. Thus, the network learns to fit the rest of the measurements. In this study, we propose to improve the performance and robustness of supervised training by utilizing randomness by retrospectively selecting only a subset of all the available measurements for data consistency units. The process is repeated multiple times using different random masks during training for further enhancement. Results on knee MRI show that the proposed multi-mask supervised PG-DL enhances reconstruction performance compared to conventional supervised PG-DL approaches.

Index Terms— Supervised learning, algorithm unrolling, physics-guided deep learning, accelerated imaging, magnetic resonance imaging, data augmentation.

1. INTRODUCTION

Lengthy acquisition times remains a challenge in MRI often causing patient discomfort or artifacts in the reconstruction. Hence, accelerating MRI remains an ongoing research topic. MRI reconstruction is based on an inverse problem that incorporates the physics of the data acquisition via the encoding operator. Direct recovery from acquired sub-sampled measurements is often ill-posed, and thus regularizers are often used for solving this inverse problem. Such regularized reconstruction is conventionally solved using iterative algorithms, such as gradient descent (GD), or proximal gradient descent (PGD) and variable splitting methods that alternate between data consistency (DC) and a proximal operator for the regularizer [1].

Recently, deep learning has emerged as an alternative approach for solving such inverse problems in MRI [2–13]. Among deep learning methods, physics-guided deep learning

(PG-DL) methods, also known as algorithm unrolling [14], have gained interest, as it incorporates the forward model for acquired measurements to the network architecture. PG-DL unrolls iterative optimization algorithms consisting of DC and regularizer units for a fixed number of iterations [4, 5, 15–18]. The conventional fixed sparsifying regularizers in these algorithms are replaced with neural networks, while gradient descent or conjugate gradient (CG) methods are employed in DC units [4, 5]. These network are then typically trained end-to-end in a supervised manner using a ground-truth reference. While PG-DL approaches may differ based on the optimization algorithm they unroll, all existing approaches use all of the available sub-sampled measurements in their DC units. Thus, the unrolled network learns to fit the rest of the unacquired measurements while the DC units ensure consistency with the acquired ones.

In this study, we propose to improve the performance and robustness of supervised training of PG-DL methods by using a multi-masking operation on the available measurements by retrospectively selecting a random subset of measurements from the original sub-sampled data for use in DC units multiple times. DC units in the unrolled network only use these subset of the measurements as opposed to conventional supervised training, where all of them are used. We hypothesize that including such random masking may inherently improve the robustness of the trained PG-DL algorithm, which will learn to fit to larger sets of different measurements for different masks, reducing residual aliasing artifacts, while not requiring any other modifications to the training process. Results on knee MRI show that the proposed multi-mask supervised PG-DL approach enhances the reconstruction performance compared to conventional supervised PG-DL by further removing residual artifacts, and improving quantitative metrics, such as SSIM and PSNR.

2. METHODS

2.1. Algorithm Unrolling for MRI Reconstruction

Let \mathbf{y}_Ω be the acquired k-space measurements with Ω denoting the undersampling pattern, and \mathbf{x} be the image to recover.

The forward model for acquired measurements is given as

$$\mathbf{y}_\Omega = \mathbf{E}_\Omega \mathbf{x} + \mathbf{n}, \quad (1)$$

where $\mathbf{E}_\Omega : \mathbb{C}^M \rightarrow \mathbb{C}^P$ is the encoding operator containing the partial Fourier sampling, coil sensitivities and the undersampling pattern, and $\mathbf{n} \in \mathbb{C}^P$ is the measurement noise. Recovery of \mathbf{x} from \mathbf{y}_Ω is formulated as

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

where the first term enforces data consistency, and $\mathcal{R}(\cdot)$ is a regularizer. Optimization techniques [1] such as variable splitting with quadratic penalty [5, 16, 19] can be employed to cast Eq. (2) into an alternating minimization problem as

$$\mathbf{z}^{(i-1)} = \arg \min_{\mathbf{z}} \mu \|\mathbf{x}^{(i-1)} - \mathbf{z}\|_2^2 + \mathcal{R}(\mathbf{z}) \quad (3a)$$

$$\mathbf{x}^{(i)} = \arg \min_{\mathbf{x}} \|\mathbf{y}_\Omega - \mathbf{E}_\Omega \mathbf{x}\|_2^2 + \mu \|\mathbf{x} - \mathbf{z}^{(i-1)}\|_2^2 \quad (3b)$$

where $\mathbf{z}^{(i)}$ is an auxiliary intermediate variable and $\mathbf{x}^{(i)}$ is the desired image at iteration i . In PG-DL methods, this iterative algorithm is unrolled for a fixed number of iterations [14]. Sub-problem Eq. (3a) is solved implicitly with a neural network, while the DC unit in sub-problem (3b) is solved using conventional linear methods [4, 5, 16].

2.2. Supervised PG-DL Training

Supervised PG-DL algorithms aims to map network input (sub-sampled k-space/distorted image) to ground-truth reference (fully-sampled k-space/artifact free image) by training neural networks on a database of acquired slices. Let $\mathbf{y}_{\text{ref}}^i$ represents the fully-sampled data for subject i and $f(\mathbf{y}_\Omega^i, \mathbf{E}_\Omega^i; \boldsymbol{\theta})$ be the the unrolled network output for input sub-sampled k-space data \mathbf{y}_Ω^i , in which the network is parameterized by $\boldsymbol{\theta}$. The objective function for supervised PG-DL training in k-space is formulated as

$$\min_{\boldsymbol{\theta}} \frac{1}{N} \sum_{i=1}^N \mathcal{L}(\mathbf{y}_{\text{ref}}^i, \mathbf{E}_{\text{full}}^i f(\mathbf{y}_\Omega^i, \mathbf{E}_\Omega^i; \boldsymbol{\theta})), \quad (4)$$

where N is the number of slices in the database, $\mathbf{E}_{\text{full}}^i$ is the multi-coil encoding operator that transform network output to k-space, $\mathbf{y}_{\text{ref}}^i$ is the fully-sampled ground-truth k-space and $\mathcal{L}(\cdot, \cdot)$ denotes the loss function. Learnt parameters for the unrolled network during training are subsequently used to reconstruct unseen undersampled test data.

2.3. Proposed Multi-Mask Supervised PG-DL Training

With supervised training as in Eq. (4), PG-DL approaches learn to fit the unacquired measurements during training, while their DC units ensure consistency with all the acquired sub-sampled measurements. Thus, we hypothesize that the

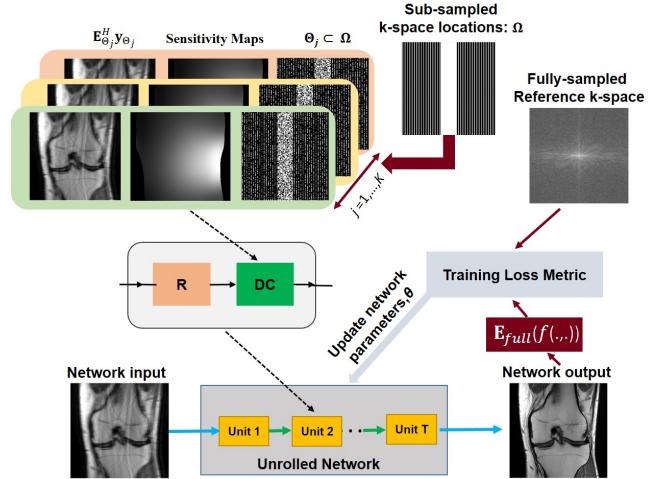


Fig. 1: The proposed multi-mask supervised PG-DL MRI reconstruction. The sub-sampled acquired measurements for each scan, Ω , are retrospectively further masked into subsets $\Theta_j \subset \Omega$, $j \in \{1, \dots, K\}$, multiple times for use in DC units. Loss is performed in k-space by comparing the fully-sampled reference k-space with the corresponding reconstructed k-space of the network output. Network parameters are subsequently updated based on the training loss.

quality and robustness of reconstruction can be further improved by masking the entries available to the DC units during supervised training [19]. To this end, instead of using all available sub-sampled measurements Ω during data consistency, we propose to use a random subset of available measurements in the DC units by retrospectively masking Ω multiple times as illustrated in Figure 1. Formally, the acquired sub-sampled measurements for each slice in the dataset is retrospectively masked K times as

$$\Theta_j \subset \Omega, \quad j \in \{1, \dots, K\}. \quad (5)$$

Hence, the loss function in Eq. (4) is reformulated as

$$\min_{\boldsymbol{\theta}} \frac{1}{N \cdot K} \sum_{i=1}^N \sum_{j=1}^K \mathcal{L}(\mathbf{y}_{\text{ref}}^i, \mathbf{E}_{\text{full}}^i(f(\mathbf{y}_{\Theta_j}^i, \mathbf{E}_{\Theta_j}^i; \boldsymbol{\theta}))). \quad (6)$$

2.4. Network Training Details

The proposed multi-mask supervised and conventional supervised PG-DL approaches are implemented by unrolling iterative sub-problems in (3a)-(3b) for 10 iterations. Each iteration contains DC and regularizer units, which are respectively implemented with CG [5] and the ResNet structure used in [9]. Coil sensitivity maps are generated from central 24×24 ACS using ESPIRiT [20]. Training is performed by using Adam optimizer with a learning rate of $5 \cdot 10^{-4}$. Network is trained to minimize a normalized $\ell_1 - \ell_2$ loss over 100 epochs with a batch size of 1 [9]. All experiments for PG-DL approaches are performed using Tensorflow in Python.

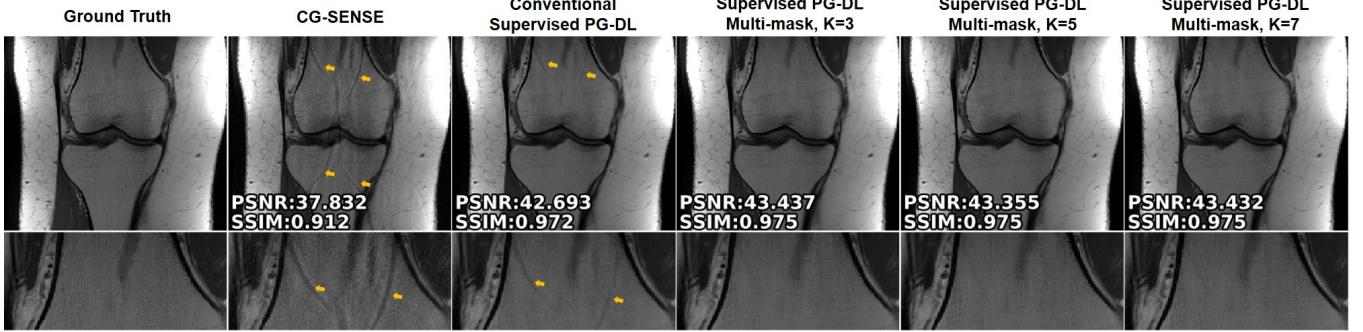


Fig. 2: Reconstruction results at $R = 4$ with uniform undersampling using CG-SENSE, proposed multi-mask and conventional supervised PG-DL approaches. CG-SENSE suffers from significant noise amplification and residual artifacts shown with yellow arrows. Conventional supervised PG-DL approach also exhibits residual artifacts for this slice. Proposed multi-mask supervised PG-DL approach outperforms the conventional PG-DL approach by successfully removing the residual artifacts, while $K=3$ masks achieve the highest metrics.

2.5. Imaging Experiments

Fully-sampled coronal proton density weighted knee MRI dataset were obtained from the New York University (NYU) fastMRI initiative database [21]. Imaging was performed on a clinical 3T system (Magnetom Skyra; Siemens, Erlangen, Germany) with a 15-channel coil-array using 2D turbo spin-echo sequences. Relevant imaging parameters were [4]: matrix size = 320×368 , in-plane resolution = $0.49 \times 0.44 \text{ mm}^2$, slice thickness = 3 mm. Training was performed on 300 slices from 15 subjects using 20 central slices from each. Testing was performed on all slices of 10 different subjects, leading to total of 380 slices.

Fully-sampled raw data was retrospectively subsampled to rate, $R = 4$ using a uniform undersampling pattern with 24 ACS lines. For the proposed method, Θ_j of each partition was selected based on a uniformly random distribution. Furthermore, $|\Theta_j|/|\Omega|$, where $|\cdot|$ denotes the cardinality of the index set, was chosen as 0.6 based on a previous self-supervised learning study [19]. The number of partitions for each slice was empirically investigated for $K \in \{3, 5, 7\}$. Proposed multi-mask supervised PG-DL approach was compared with conventional supervised PG-DL and conjugate gradient SENSE (CG-SENSE) [22]. Testing was performed both for uniform and random undersampling patterns. PSNR and SSIM were used for quantitative evaluation on the test dataset.

3. RESULTS

Figure 2 shows reconstruction results from a representative test slice using conventional supervised PG-DL, i.e. $K = 1$, and the proposed multi-mask supervised PG-DL for $K \in \{3, 5, 7\}$, as well as conventional CG-SENSE. CG-SENSE suffers from significant residual artifacts and noise amplification. Conventional supervised PG-DL approach also displays residual artifacts on this slice, marked with yellow arrows. Proposed multi-mask supervised PG-DL achieves a better re-

construction performance by further suppressing these residual artifacts for all choices of K .

Figure 3a and 3b display reconstruction results on two different subjects using uniform and random undersampling, respectively. CG-SENSE suffers from residual artifacts in both cases, with random undersampling showing fewer artifacts, due to the incoherent nature of aliasing. For both of these slices, proposed multi-mask supervised PG-DL and conventional supervised PG-DL do not suffer from any visible artifacts, while closely performing with each other.

Table 1 summarizes the median and interquartile ranges (25th-75th percentile) of SSIM and PSNR values of all reconstruction methods for uniform and random undersampling patterns. For both uniform and random undersampling scenarios, CG-SENSE performs worse than PG-DL approaches, while the proposed multi-mask supervised PG-DL approaches achieves improved metrics compared to conventional supervised PG-DL approach. Proposed multi-mask supervised PG-DL achieves the best metrics for $K = 3$, which is consistent with the observations in Figures 2 and 3.

4. DISCUSSION AND CONCLUSION

In this study, we proposed a multi-mask supervised PG-DL approach, which retrospectively masks the undersampled measurements in the DC units multiple times to enhance the reconstruction quality. Results on knee MRI showed that the proposed method successfully removes residual artifacts that may occur in some test datasets using conventional supervised PG-DL. Moreover, it preserves the image quality and further improves the quantitative metrics of conventional supervised PG-DL if there are no artifacts in the reconstruction.

Supervised PG-DL provides improved reconstruction quality compared to conventional clinical approaches [4, 5, 13, 17]. However, it may still exhibit residual artifacts on some test slices. Hence, we hypothesized that bringing a degree of randomness into supervised learning can further

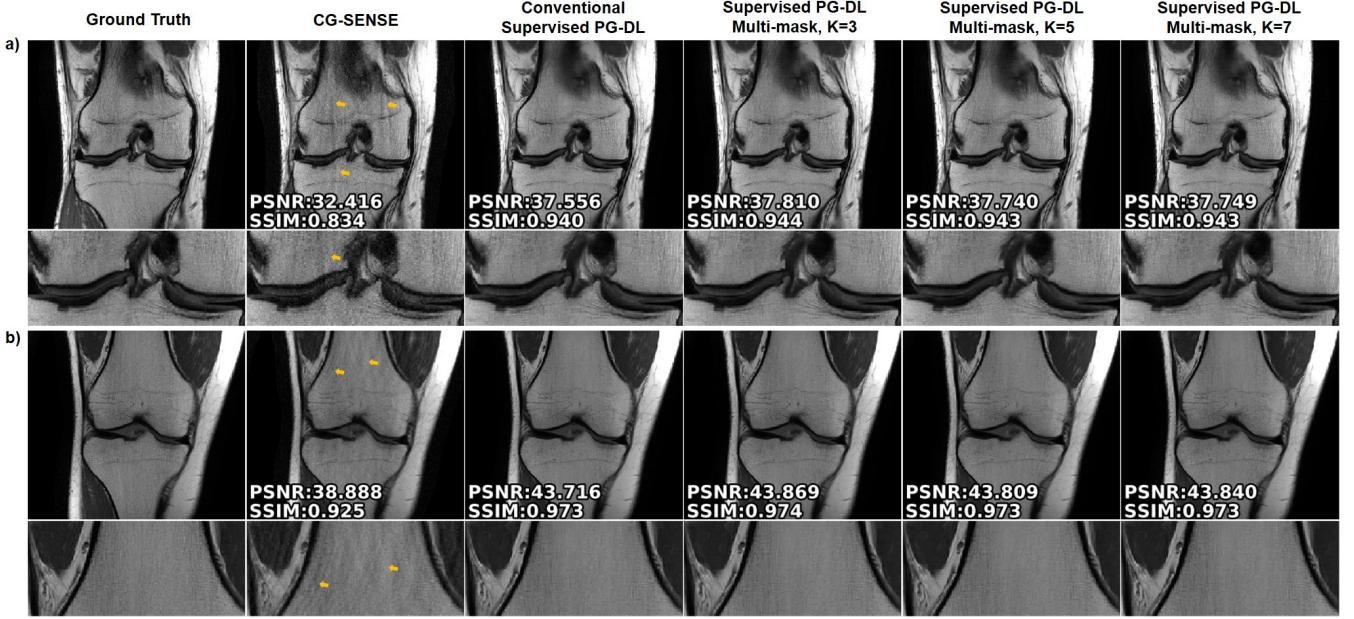


Fig. 3: Reconstruction results for two different slices at $R = 4$ using CG-SENSE, proposed multi-mask and conventional supervised PG-DL approaches for a) uniform and b) random undersampling scenarios. For both sampling patterns, CG-SENSE suffers from noise amplification and residual artifacts shown with yellow arrows, while conventional supervised and proposed multi-mask supervised PG-DL approaches perform closely without depicting any residual artifacts. For both random and uniform undersampling, proposed multi-mask supervised at $K=3$ achieves the highest PSNR and SSIM.

improve its robustness and reconstruction performance. The proposed multi-masking approach achieves improved robustness, as the additional randomization in the DC units enforce trained networks to learn different sets of measurements. Importantly, this is achieved without requiring any major change in the training process.

The proposed multi-mask supervised PG-DL was shown to suppress residual artifacts that were observed in some test slices using conventional supervised PG-DL. Additionally, for all K values, there was also a quantitative improvement in metrics compared to conventional supervised PG-DL. These improvements come only at the expense of K -fold increased training time, which does not have any impact on the testing time. While proposed multi-masking ($K > 1$) improves performance over conventional supervised PG-DL with $K=3$ achieving the best results quantitatively, increasing K further does not necessarily lead to more improvements. These

observations on selection of K are consistent with the literature on data augmentation, where the optimal size of the post-augmented data remains an open research problem [23].

5. ACKNOWLEDGEMENTS

This work was partially supported by NIH R01HL153146, P41EB027061, U01EB025144; NSF CAREER CCF-1651825. Knee MRI data were obtained from the NYU fastMRI initiative database [21]. NYU fastMRI database was acquired with the relevant institutional review board approvals as detailed in [21]. 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. There is no conflict of interest for the authors.

| | Quantitative Metric | CG-SENSE | Conventional Supervised PG-DL | Supervised PG-DL Multi-mask, $K=3$ | Supervised PG-DL Multi-mask, $K=5$ | Supervised PG-DL Multi-mask, $K=7$ |
|------------------|---------------------|-------------------------|-------------------------------|------------------------------------|------------------------------------|------------------------------------|
| Uniform Sampling | SSIM | 0.881 [0.849, 0.919] | 0.965 [0.955, 0.977] | 0.968 [0.957, 0.979] | 0.967 [0.956, 0.977] | 0.966 [0.955, 0.977] |
| | PSNR | 35.329 [33.771, 37.413] | 40.822 [39.308, 42.290] | 41.199 [39.417, 42.702] | 41.031 [39.372, 42.727] | 41.069 [39.341, 42.634] |
| Random Sampling | SSIM | 0.901 [0.874, 0.930] | 0.968 [0.955, 0.978] | 0.970 [0.958, 0.979] | 0.969 [0.957, 0.978] | 0.968 [0.957, 0.978] |
| | PSNR | 36.122 [34.672, 38.135] | 40.987 [39.435, 43.068] | 41.216 [39.620, 43.315] | 41.099 [39.529, 43.338] | 41.087 [39.534, 43.298] |

Table 1: The median and interquartile range (25th-75th percentile) of SSIM and PSNR metrics on test dataset for uniform and random undersampling cases. For both sampling scenarios, CG-SENSE is outperformed by PG-DL approaches in terms of both SSIM and PSNR. The proposed multi-mask supervised PG-DL achieves improved metrics compared to conventional supervised PG-DL approach for all $K > 1$. For multi-mask supervised PG-DL, $K = 3$ leads to the highest PSNR and SSIM values for both uniform and random undersampling scenarios.

References

- [1] J. A. Fessler, “Optimization methods for magnetic resonance image reconstruction: Key models and optimization algorithms,” *IEEE Sig Proc Mag*, vol. 37, no. 1, pp. 33–40, 2020.
- [2] S. Wang, Z. Su, et al., “Accelerating magnetic resonance imaging via deep learning,” in *Proc. IEEE Int. Symp. Biomed. Imag. (ISBI)*, April 2016, pp. 514–517.
- [3] M. Akçakaya, S. Moeller, S. Weingärtner, and K. Ügurbil, “Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: Database-free deep learning for fast imaging,” *Magn Reson Med*, vol. 81, pp. 439–453, 2019.
- [4] K. Hammernik, T. Klatzer, et al., “Learning a variational network for reconstruction of accelerated MRI data,” *Magn Reson Med*, vol. 79, pp. 3055–3071, 2018.
- [5] H. K. Aggarwal, M. P. Mani, and M. Jacob, “MoDL: Model-Based Deep Learning Architecture for Inverse Problems,” *IEEE Trans Med Imaging*, vol. 38, pp. 394–405, 2019.
- [6] D. Lee, J. Yoo, S. Tak, and J. C. Ye, “Deep residual learning for accelerated MRI using magnitude and phase networks,” *IEEE Trans Biomed Eng*, vol. 65, pp. 1985–1995, 2018.
- [7] M. Mardani, E. Gong, et al., “Deep Generative Adversarial Neural Networks for Compressive Sensing MRI,” *IEEE Trans Med Imaging*, vol. 38, pp. 167–179, 2019.
- [8] F. Knoll, K. Hammernik, et al., “Deep-learning methods for parallel magnetic resonance imaging reconstruction: A survey of the current approaches, trends, and issues,” *IEEE Sig Proc Mag*, vol. 37, no. 1, pp. 128–140, 2020.
- [9] B. Yaman, S. A. H. Hosseini, et al., “Self-Supervised Learning of Physics-Guided Reconstruction Neural Networks without Fully-Sampled Reference Data,” *Magn Reson Med*, vol. 84, no. 6, pp. 3172–3191, Dec 2020.
- [10] S. A. Hossein Hosseini, B. Yaman, S. Moeller, and M. Akçakaya, “High-fidelity accelerated mri reconstruction by scan-specific fine-tuning of physics-based neural networks,” in *IEEE Engineering in Medicine Biology Society (EMBC)*, 2020, pp. 1481–1484.
- [11] J. Liu, Y. Sun, et al., “Rare: Image reconstruction using deep priors learned without groundtruth,” *IEEE J. Sel. Top. Signal Process.*, vol. 14, pp. 1088–1099, 2020.
- [12] C. D. Bahadir, A. Q. Wang, A. V. Dalca, and M. R. Sabuncu, “Deep-learning-based optimization of the under-sampling pattern in mri,” *IEEE Transactions on Computational Imaging*, vol. 6, pp. 1139–1152, 2020.
- [13] D. Liang, J. Cheng, Z. Ke, and L. Ying, “Deep magnetic resonance image reconstruction: Inverse problems meet neural networks,” *IEEE Sig Proc Mag*, vol. 37, no. 1, pp. 141–151, 2020.
- [14] V. Monga, Y. Li, and Y. Eldar, “Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing,” *arXiv preprint arXiv:1912.10557*, 2019.
- [15] Y. Yang, J. Sun, H. Li, and Z. Xu, “Deep ADMM-Net for compressive sensing MRI,” in *Advances in neural information processing systems*, 2016, pp. 10–18.
- [16] C. Qin, J. Schlemper, et al., “Convolutional Recurrent Neural Networks for Dynamic MR Image Reconstruction,” *IEEE Trans Med Imaging*, vol. 38, pp. 280–290, 2019.
- [17] S. A. H. Hosseini, B. Yaman, S. Moeller, M. Hong, and M. Akçakaya, “Dense recurrent neural networks for accelerated mri: History-cognizant unrolling of optimization algorithms,” *IEEE J. Sel. Top. Signal Process.*, vol. 14, no. 6, pp. 1280–1291, 2020.
- [18] S. Wang, H. Cheng, et al., “Deepcomplexmri: Exploiting deep residual network for fast parallel mr imaging with complex convolution,” *Magnetic Resonance Imaging*, vol. 68, pp. 136–147, 2020.
- [19] B. Yaman, S. Hosseini, et al., “Multi-mask self-supervised learning for physics-guided neural networks in highly accelerated MRI,” *arXiv preprint arXiv:2008.06029*, 2020.
- [20] M. Uecker, P. Lai, et al., “ESPIRiT—an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA,” *Magn Reson Med*, vol. 71, no. 3, pp. 990–1001, Mar 2014.
- [21] F. Knoll, J. Zbontar, et al., “fastMRI: A Publicly Available Raw k-Space and DICOM Dataset of Knee Images for Accelerated MR Image Reconstruction Using Machine Learning,” *Radiol Artif Intell*, vol. 2, no. 1, pp. e190007, Jan 2020.
- [22] K. P. Pruessmann, M. Weiger, P. Bornert, and P. Boesiger, “Advances in sensitivity encoding with arbitrary k-space trajectories,” *Magn Reson Med*, vol. 46, pp. 638–651, 2001.
- [23] C. Shorten and T. Khoshgoftaar, “A survey on image data augmentation for deep learning,” *Journal of Big Data*, vol. 6, no. 1, pp. 60, 2019.