

GROUND-TRUTH FREE MULTI-MASK SELF-SUPERVISED PHYSICS-GUIDED DEEP LEARNING IN HIGHLY ACCELERATED MRI

Burhaneddin Yaman^{†}, Seyed Amir Hossein Hosseini^{*†}, Steen Moeller[†],
Jutta Ellermann[†], Kâmil Uğurbil[†] 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

Deep learning based MRI reconstruction methods typically require databases of fully-sampled data as reference for training. However, fully-sampled acquisitions may be either challenging or impossible in numerous scenarios. Self-supervised learning enables training neural networks for MRI reconstruction without fully-sampled data by splitting available measurements into two disjoint sets. One of them is used in data consistency units in the network, and the other is used to define the loss. However, the performance of self-supervised learning degrades at high acceleration rates due to scarcity of acquired data. We propose a multi-mask self-supervised learning approach, which retrospectively splits available measurements into multiple 2-tuples of disjoint sets. Results on 3D knee and brain MRI shows that the proposed multi-mask self-supervised learning approach significantly improves upon single mask self-supervised learning at high acceleration rates.

Index Terms— Self-supervised learning, physics-guided deep learning, accelerated imaging, parallel imaging

1. INTRODUCTION

Data acquisition is lengthy in many MRI exams, necessitating the use of accelerated imaging methods. Parallel imaging [1, 2] and compressed sensing [3] are conventional accelerated MRI techniques, but their acceleration rate is limited due to either noise amplification or residual artifacts in the reconstruction. Recently, deep learning (DL) approaches have been proposed as an alternative to accelerate MRI [4–10]. Specifically, physics-guided DL (PG-DL) techniques, which unroll iterative optimization algorithms for solving a regularized inverse problem, have gained interest due to its robustness and improved reconstruction quality [6, 7, 11].

PG-DL approaches are typically trained in a supervised manner, requiring fully-sampled data as reference for training. However, acquisition of fully-sampled data may be challenging or impossible in many scenarios rendering such supervised PG-DL approaches inoperative [12]. Several

methods have been proposed to deal with this issue [12–17]. Among these, Self-Supervision via Data Undersampling (SSDU) is a self-supervised strategy that splits available k-space measurements into two disjoint sets by a masking operation [12, 13]. One of these sets is used for data consistency (DC) in the unrolled network and the other is used to define the loss in k-space. It was shown that SSDU performed similar to supervised learning approaches at moderately high acceleration rates. However, for higher acceleration rates, in which acquired data becomes increasingly scarce, further improvements that can efficiently augment the available data is essential for enhanced reconstruction quality.

In this study, we aim to improve the performance of SSDU with multiple masks. For each slice in the dataset, the proposed multi-mask SSDU splits available undersampled measurements into multiple pairs of disjoint sets, in which one of them is used in DC units and the other is used to define loss in k-space. Results on 3D knee and brain MRI show that the multi-mask SSDU approach significantly enhances the reconstruction quality of SSDU at high acceleration rates.

2. MATERIALS AND METHODS

2.1. Algorithm Unrolling for MRI Reconstruction

The inverse problem for accelerated MRI is given as

$$\arg \min_{\mathbf{x}} \|\mathbf{y}_{\Omega} - \mathbf{E}_{\Omega} \mathbf{x}\|_2^2 + \mathcal{R}(\mathbf{x}), \quad (1)$$

where $\mathbf{y}_{\Omega}$ is the acquired k-space data, Ω is the sub-sampling pattern, $\mathbf{E}_{\Omega} : \mathbb{C}^M \rightarrow \mathbb{C}^P$ is the multi-coil encoding operator, $\mathbf{x}$ is the image to be recovered and $\mathcal{R}(\cdot)$ is a regularizer. Eq. (1) is decoupled into regularizer and DC units by standard optimization methods, e.g. variable-splitting via quadratic penalty [18], as

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

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

where $\mathbf{z}^{(i)}$ is an intermediate variable and $\mathbf{x}^{(i)}$ is the desired image at iteration i . In PG-DL, the iterative algorithm is unrolled for fixed number of iterations. Neural networks are

used as regularizer for solving Eq. (2a) and DC sub-problem in Eq. (2b) is solved via conjugate gradient (CG) [7].

2.2. Supervised PG-DL Training

In supervised PG-DL, training is performed by using fully-sampled data reference. Let $\mathbf{y}_{\text{ref}}^i$ denotes the fully-sampled k-space data for subject i and $f(\mathbf{y}_{\Omega}^i, \mathbf{E}_{\Omega}^i; \theta)$ denotes the output of the unrolled network for sub-sampled k-space data $\mathbf{y}_{\Omega}^i$, in which the network is parameterized by θ . End-to-end training is performed by minimizing [12]

$$\min_{\theta} \frac{1}{N} \sum_{i=1}^N \mathcal{L}(\mathbf{y}_{\text{ref}}^i, \mathbf{E}_{\text{full}}^i f(\mathbf{y}_{\Omega}^i, \mathbf{E}_{\Omega}^i; \theta)), \quad (3)$$

where N is the number of samples in the training database, $\mathbf{E}_{\text{full}}^i$ is the fully-sampled encoding operator that transform network output to k-space and $\mathcal{L}(\cdot, \cdot)$ is a loss function.

2.3. SSDU PG-DL Training

Unlike supervised PG-DL approaches, SSDU does not require a fully-sampled data as reference for training. SSDU splits available undersampled measurements, Ω into two disjoint sets, Θ and Λ as $\Omega = \Theta \cup \Lambda$, where Θ denotes the k-space locations used in the DC units in the network during training and Λ denotes a set of k-space locations used in the loss function. SSDU performs training without fully-sampled data by minimizing difference between acquired k-space and multi-coil k-space of the network output at Λ [12]

$$\min_{\theta} \frac{1}{N} \sum_{i=1}^N \mathcal{L}(\mathbf{y}_{\Lambda}^i, \mathbf{E}_{\Lambda}^i (f(\mathbf{y}_{\Theta}^i, \mathbf{E}_{\Theta}^i; \theta))). \quad (4)$$

2.4. Proposed Multi-Mask SSDU PG-DL Training

Data scarcity becomes more pronounced at high acceleration rates, especially after further splitting them into two sets in SSDU. We propose to overcome this issue by employing a multi-mask SSDU approach, which retrospectively splits Ω into two disjoint sets Θ_j and Λ_j , for $j \in \{1, \dots, K\}$, as illustrated in Figure 1. Hence, available measurements for each slice in the dataset is partitioned K times such that

$$\Omega = \Theta_j \cup \Lambda_j, \quad j = 1, \dots, K. \quad (5)$$

with $\Lambda_j = \Omega / \Theta_j$. Hence the proposed multi-mask SSDU training objective is to minimize

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

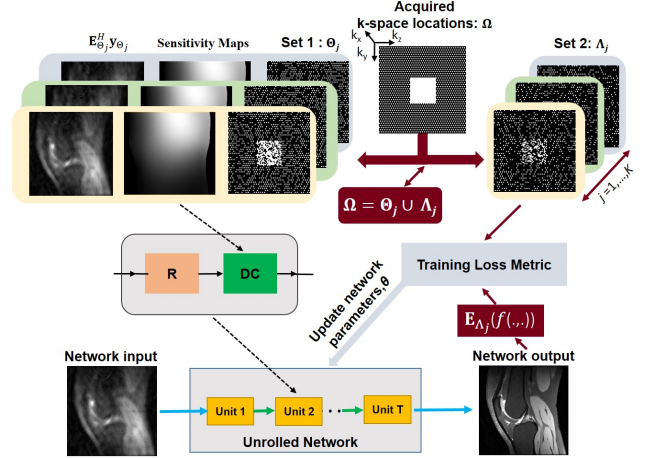


Fig. 1: The proposed multi-mask self-supervised learning for PG-DL MRI reconstruction. Acquired k-space locations for each scan, Ω , are retrospectively sub-sampled into two disjoint sets of Θ_j and Λ_j for $j \in \{1, \dots, K\}$. For each such partitioning, Θ_j is used for DC units and $\Lambda_j = \Omega / \Theta_j$ is used to define the loss function. Loss is performed in k-space by comparing acquired data with the multi-coil k-space of the network output at indices Λ_j . Subsequently, the network parameters are updated based on the training loss.

2.5. Network Training Details

Sub-problems (2a)-(2b) are unrolled for $T = 10$ iterations. DC units employs CG, and regularizers use the same ResNet structure as in [12]. $K = 7$ was empirically selected through hyperparameter tuning [19] and used throughout this study for the proposed multi-mask approach. Coil sensitivity maps were generated from central 24×24 ACS using ESPIRiT [20]. All PG-DL networks were trained using an Adam optimizer with a learning rate of $5 \cdot 10^{-4}$ over 100 epochs by minimizing a normalized $\ell_1 - \ell_2$ loss [12]. The network had a total of 592,129 trainable parameters.

2.6. Imaging Experiments

Fully-sampled 3D knee dataset were obtained from mri-data.org [21], for which the imaging protocols were approved by the local institutional review board. Data was acquired on a 3T GE Discovery MR 750 system with an 8-channel coil-array using a fast spin-echo sequence. Relevant imaging parameters were: FOV = $160 \times 160 \times 154$ mm³, resolution = $0.5 \times 0.5 \times 0.6$ mm³, matrix size = $320 \times 320 \times 256$.

Brain MRI was performed on a 3T Siemens Magnetom Prisma system with a 32-channel coil-array using a 3D-MPRAGE sequence [12]. The imaging protocols were approved by the local institutional review board, and written informed consent was obtained from all participants. Relevant imaging parameters were: FOV = $224 \times 224 \times 157$ mm³, resolution = $0.7 \times 0.7 \times 0.7$ mm³, matrix size = 320×320 , prospective acceleration R = 2, ACS lines = 32.

2D slices were processed after taking inverse Fourier

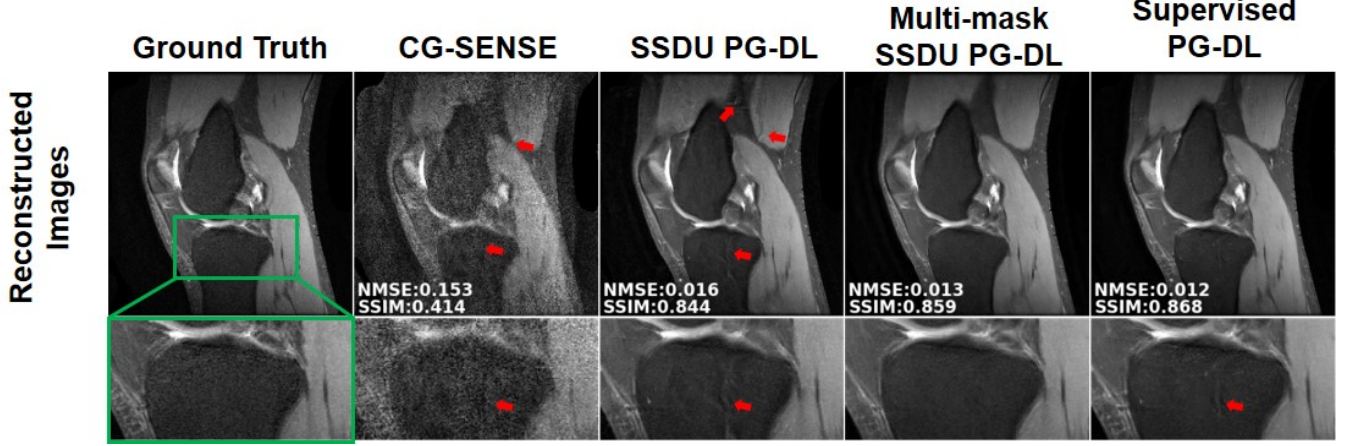


Fig. 2: A representative slice showing reconstruction results at $R = 8$ using CG-SENSE, supervised PG-DL, SSDU PG-DL and proposed multi-mask SSDU PG-DL. CG-SENSE suffers from significant noise and artifacts. At this high acceleration rate, SSDU PG-DL also shows residual artifacts (red arrows). Proposed multi-mask SSDU PG-DL suppresses these artifacts, which are still visible in supervised PG-DL.

transformed along the read-out direction for the 3D k-space dataset. Knee and brain datasets were further retrospectively subsampled to $R = 8$ by keeping 24×24 and 32×32 ACS regions in the $k_y - k_z$ plane using a sheared uniform undersampling pattern, respectively [12]. For both datasets, training was performed on 300 slices obtained from 10 subjects by taking 30 central slices. Testing was performed on 8 and 9 different subjects for knee and brain MRI, respectively.

The proposed multi-mask SSDU PG-DL approach was compared to CG-SENSE [22] and SSDU PG-DL, as well as supervised PG-DL when applicable. The same network structure was used for all PG-DL approaches. Normalized mean square error (NMSE) and structural similarity index (SSIM) were used for quantitative assessments. Furthermore, a reader study was performed for overall image quality on a 4-point ordinal scale (1: excellent, 2: good, 3: fair, 4: poor) [6], where the reader was blinded to the reconstruction method.

3. RESULTS

Figure 2 shows a representative knee MRI slice reconstructed using CG-SENSE, SSDU, proposed multi-mask SSDU and supervised PG-DL. CG-SENSE suffers from visible residual artifacts (red arrows). Proposed multi-mask SSDU visually outperforms SSDU, while also reducing residual aliasing artifacts compared to supervised PG-DL, which is not reflected in the quantitative metrics for this slice.

Figure 3 depicts reconstruction results of a slice from 3D-MPRAGE data for CG-SENSE at acquisition acceleration $R = 2$, as well as CG-SENSE, SSDU PG-DL and the multi-mask SSDU PG-DL methods at retrospective acceleration $R = 8$. CG-SENSE at $R = 8$ suffers from significant noise amplification, rendering it unusable. SSDU PG-DL at $R = 8$ shows similar reconstruction quality as the clinical baseline CG-SENSE at $R = 2$. Proposed multi-mask SSDU PG-DL further sup-

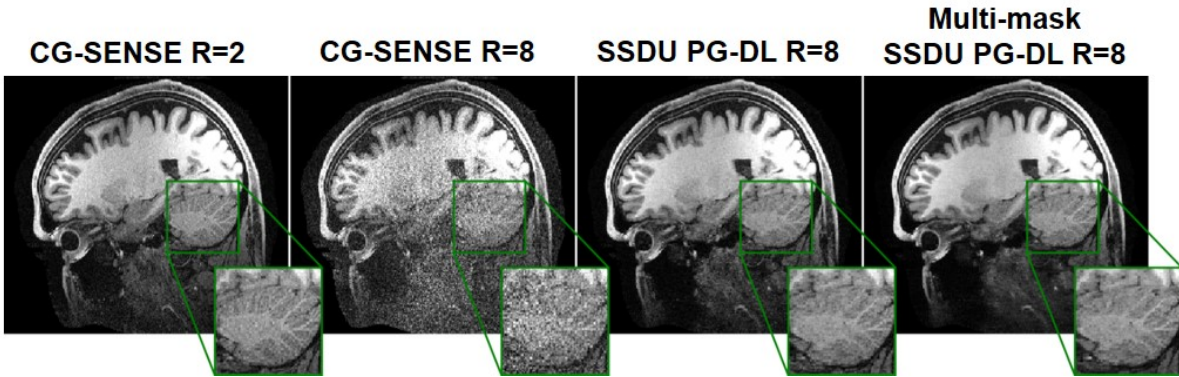


Fig. 3: Reconstruction results for a representative brain MRI test slice. CG-SENSE was applied at both the acquisition acceleration $R=2$ and further retrospective acceleration $R=8$, while SSDU PG-DL and the proposed multi-mask SSDU PG-DL approaches were applied at $R=8$. While CG-SENSE suffers from significant noise amplification at $R=8$, SSDU PG-DL at $R=8$ achieves similar reconstruction quality to CG-SENSE at $R=2$. The proposed multi-mask SSDU PG-DL further improves the reconstruction quality compared to SSDU PG-DL.

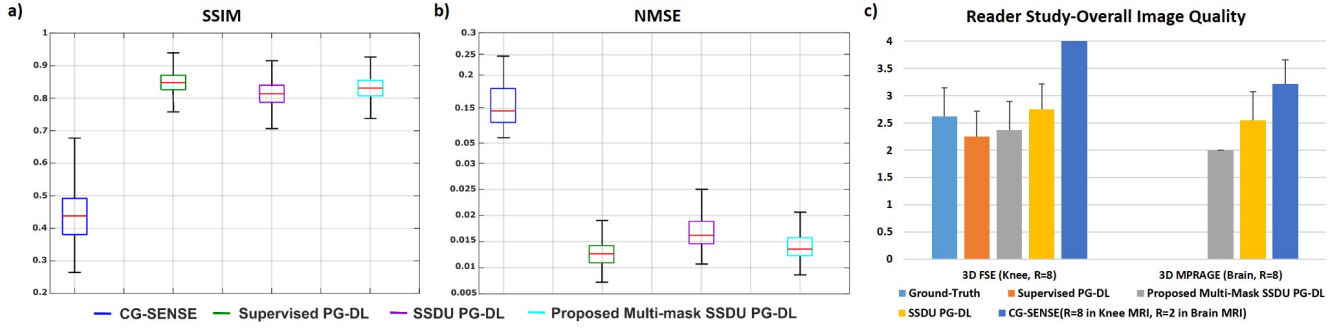


Fig. 4: Median and interquartile range (25th-75th percentile) of the a) SSIM and b) NMSE values on test dataset. All PG-DL approaches outperform CG-SENSE. Proposed multi-mask SSDU outperforms SSDU, while closely matching supervised DL-MRI in terms of both SSIM and NMSE. c) Average reader scores for overall image quality on an 4-point scale (1: excellent, 2: good, 3: fair, 4: poor). For knee MRI, CG-SENSE was rated lowest, whereas the proposed multi-mask SSDU PG-DL was rated higher than SSDU PG-DL and received similar scores with supervised PG-DL. Similarly in brain MRI, CG-SENSE at R=2, which is the clinical baseline, received the lowest scores, while the proposed multi-mask SSDU PG-DL at R=8 was rated highest.

presses the noise in SSDU PG-DL and achieves a superior reconstruction quality.

Figures 4a and 4b show the average SSIM and NMSE values for all reconstruction methods for the 3D knee data. The proposed multi-mask SSDU PG-DL outperforms SSDU PG-DL in terms of both SSIM and NMSE, while closely performing with supervised PG-DL. Figure 4c shows the average reader scores for overall image quality for both knee and brain MRI. CG-SENSE was rated worst in both anatomies. The proposed multi-mask SSDU PG-DL was rated higher than SSDU PG-DL for both knee and brain MRI, while ranking closely with supervised PG-DL in knee MRI. Furthermore for brain MRI, both SSDU PG-DL approaches at R=8 were rated higher than CG-SENSE at R=2, which corresponds to the current clinical baseline.

4. DISCUSSION AND CONCLUSION

In this study, we proposed a multi-mask SSDU approach for training PG-DL MRI reconstruction without fully-sampled data, where the acquired k-space measurements for each dataset in the training database were retrospectively split into two disjoint sets multiple times. Results showed that multi-mask SSDU approach improves the performance of SSDU at high acceleration rates, while performing closely with supervised PG-DL.

Although supervised PG-DL achieves higher quantitative metrics compared proposed multi-mask SSDU PG-DL approach in knee MRI, the qualitative results presented for knee MRI showed that the proposed multi-mask SSDU approach has the potential to perform better than supervised PG-DL in terms of handling residual artifacts. Moreover, the reader study on knee MRI showed that the proposed multi-mask SSDU improves upon SSDU PG-DL, considerably closing the gap with supervised PG-DL. These assessments align with recent self-supervised works that showed self-

supervised learning can indeed surpass supervised learning approaches in several scenarios [23, 24].

The use of multiple masks in the DC units of an unrolled network proposed here offers an alternative approach for data augmentation for medical imaging dataset as other approaches such as rotations may not fit well since it manipulates acquired undersampled data. This is especially important for cases where size of training dataset is small or at higher acceleration rates where acquired data is inherently scarce.

5. ACKNOWLEDGEMENTS

This work was partially supported by NIH R01HL153146, P41EB027061, U01EB025144; NSF CAREER CCF-1651825.

References

- [1] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: sensitivity encoding for fast MRI," *Magn Reson Med*, vol. 42, pp. 952–962, 1999.
- [2] M. A. Griswold, P. M. Jakob, et al., "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magn Reson Med*, vol. 47, pp. 1202–1210, 2002.
- [3] M. Lustig, D. Donoho, and J. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn Reson Med*, vol. 58, pp. 1182–1195, 2007.
- [4] S. Wang, Z. Su, et al., "Accelerating magnetic resonance imaging via deep learning," in *Proc IEEE ISBI*, April 2016, pp. 514–517.
- [5] M. Akçakaya, S. Moeller, S. Weingärtner, 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, pp. 439–453, 2019.

- [6] 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.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] 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.
- [11] 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 Sig Proc.*, vol. 14, no. 6, pp. 1280–1291, 2020.
- [12] 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.
- [13] B. Yaman, S. Hosseini, et al., “Self-supervised physics-based deep learning mri reconstruction without fully-sampled data,” in *Proc IEEE ISBI*, 2020, pp. 921–925.
- [14] J. I. Tamir, S. X. Yu, and M. Lustig, “Unsupervised deep basis pursuit: Learning inverse problems without ground-truth data,” *preprint arXiv:1910.13110*, 2019.
- [15] G. Oh, B. Sim, H. Chung, L. Sunwoo, and J. C. Ye, “Unpaired deep learning for accelerated MRI using optimal transport driven cycleGAN,” *IEEE Trans Comp Imaging*, vol. 6, pp. 1285–1296, 2020.
- [16] 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.
- [17] Z. Ke, J. Cheng, et al., “An unsupervised deep learning method for multi-coil cine MRI,” *Physics in Medicine & Biology*, 2020.
- [18] 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.
- [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. Ong, S. Amin, S. Vasanawala, and M. Lustig, “Mri-data.org: An open archive for sharing mri raw data,” in *Proc. Intl. Soc. Mag. Reson. Med*, 2018, vol. 26, p. 1.
- [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] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, “Momentum contrast for unsupervised visual representation learning,” in *Proc IEEE CVPR*, June 2020.
- [24] I. Misra and L. v. d. Maaten, “Self-supervised learning of pretext-invariant representations,” in *Proc IEEE CVPR*, June 2020.