High-fidelity Database-free Deep Learning Reconstruction for Real-time Cine Cardiac MRI

Ömer Burak Demirel[1,2], Chi Zhang[1,2], Burhaneddin Yaman[1,2], Merve Gulle[1,2],
Chetan Shenoy[3], Tim Leiner[4], Peter Kellman[5] and Mehmet Akçakaya[1,2]

Abstract— **Real-time cine cardiac MRI provides an ECG-free free-breathing alternative to clinical gold-standard ECG-gated breath-hold segmented cine MRI for evaluation of heart function. Real-time cine MRI data acquisition during free breathing snapshot imaging enables imaging of patient cohorts that cannot be imaged with segmented or breath-hold acquisitions, but requires rapid imaging to achieve sufficient spatial-temporal resolutions. However, at high acceleration rates, conventional reconstruction techniques suffer from residual aliasing and temporal blurring, including advanced methods such as compressed sensing with radial trajectories. Recently, deep learning (DL) reconstruction has emerged as a powerful tool in MRI. However, its utility for free-breathing real-time cine MRI has been limited, as database-learning of spatio-temporal correlations with varying breathing and cardiac motion patterns across subjects has been challenging. Zero-shot self-supervised physics-guided deep learning (PG-DL) reconstruction has been proposed to overcome such challenges of database training by enabling subject-specific training. In this work, we adapt zero-shot PG-DL for real-time cine MRI with a spatio-temporal regularization. We compare our method to TGRAPPA, locally low-rank (LLR) regularized reconstruction and database-trained PG-DL reconstruction, both for retrospectively and prospectively accelerated datasets. Results on highly accelerated real-time Cartesian cine MRI show that the proposed method outperforms other reconstruction methods, both visibly in terms of noise and aliasing, and quantitatively.**

I. INTRODUCTION

Real-time cine cardiac MRI is a free-breathing and ECG-free alternative to gold-standard ECG-gated and breath-hold cine MRI for functional and volumetric assessment of the heart to diagnose cardiovascular diseases [1]. ECG-gated acquisitions require breath-hold imaging to minimize respiratory motion artifacts, but it may fail in patients with arrhythmias or difficulty breath-holding, and is not possible during exercise stress [2]. On the other hand, real-time cine MRI does not require ECG gating or breath-holding, which appeals to a more general patient cohort [3]. However, to achieve sufficient spatio-temporal resolutions, highly accelerated data acquisition is needed. This imposes significant challenges for subsequent image reconstruction.

Previous approaches to real-time cine MRI utilize parallel imaging in conjunction with Cartesian [4], [5] or non-

Cartesian trajectories, especially radial imaging [6]. Subsequently, compressed sensing (CS) with radial sampling was used to further improve spatio-temporal resolution [7]. Yet, at very high acceleration rates, CS reconstruction may suffer from residual aliasing artifacts and blurring [8]. Recently physics-guided deep learning (PG-DL) reconstruction has emerged as a powerful method for highly-accelerated MRI, improving on parallel imaging and CS [9]–[12]. The utility of DL reconstruction in real-time cine MRI has so far been limited to data-driven image enhancement approaches that learn a mapping between aliased and artifact-free images [13], [14]. However, learning a spatio-temporally regularized reconstruction for free-breathing real-time cine MRI has its own challenges, particularly due to the varying breathing and cardiac motion patterns between subjects, along with a lack of fully-sampled reference data for highly-accelerated scans.

In this study, we propose to use subject-specific zero-shot self-supervised PG-DL [15] without a training database or ground-truth data to improve highly accelerated real-time Cartesian cine MRI. The proposed PG-DL network operates on the subject of interest using spatio-temporal regularization. The proposed approach was compared with TGRAPPA [4], locally-low-rank (LLR) regularized reconstruction [16] and database-trained self-supervised PG-DL [12] with both retrospectively and prospectively accelerated datasets. Results show that the subject-specific zero-shot PG-DL reconstruction substantially improves upon other methods, showing excellent image quality compared to baseline images while preserving temporal fidelity. Results on retrospectively highly-accelerated real-time cine MRI show that the proposed method improves on existing methods in terms of PSNR (up to 35.98% gain) and SSIM (up to 39.29% gain) when compared to standard acceleration. Additionally, our approach establishes the feasibility of real-time Cartesian cine MRI with 8-fold prospective acceleration that matches the spatio-temporal resolution of breath-hold cine MRI.

II. METHODS

A. *Physics-Guided Deep Learning MRI Reconstruction*

Regularized MRI reconstruction is formulated as:

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

where $\mathbf{x}$ is the image of interest, $\mathbf{y}_\Omega$ is the acquired multi-channel k-space with Ω is the undersampling pattern, $\mathbf{E}_\Omega$ is the multi-coil encoding operator and $\mathcal{R}(\cdot)$ is the regularizer. PG-DL methods typically solve Eq. 1 by algorithm unrolling

[1]Department of Electrical and Computer Engineering, [2]Center for Magnetic Resonance Research, [3]Department of Medicine (Cardiology), University of Minnesota, Minneapolis, MN, USA, [4] Department of Radiology, Mayo Clinic, Rochester, MN, USA and [5]National Heart-Lung and Blood Institute, Bethesda, MD, USA. e-mails: {demir035,zhan490,yaman013,glle0001,cshenoy, akcakaya}@umn.edu, kellmanp@nhlbi.nih.gov and leiner.tim@mayo.edu

that alternates between conventional data-fidelity (DF) with acquired k-space and a proximal operator that is implicitly implemented via neural networks. Among various techniques to decouple the DF and regularizer terms [17], proximal gradient descent alternates between:

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

$$\mathbf{x}^{(i)} = \mathbf{x}^{(i-1)} + \mu \mathbf{E}_\Omega^H (\mathbf{y}_\Omega - \mathbf{E}_\Omega \mathbf{x}), \qquad (3)$$

where $\mathbf{z}$ is an auxiliary variable at the the ith iteration and μ is penalty term. These networks are typically trained end-to-end in a supervised manner using reference images [11].

B. Database-free Self-supervised PG-DL Reconstruction

Several unsupervised learning strategies have been proposed to tackle the challenges associated with the end-to-end supervised training of unrolled networks that require reference data [18]. Among these, self-supervised learning strategies have been utilized in various applications [12], [19]–[21]. Nonetheless, these require a database for training, which is hard to curate for real-time cine MRI due to breathing pattern differences among subjects. Recently, zero-shot self-supervised learning via data undersampling (ZS-SSDU) has been proposed to enable database-free training of PG-DL reconstructions [15]. ZS-SSDU splits the acquired k-space locations, Ω, into three disjoint sets. First two sets are similar to database-trained SSDU [12]: Ω used in the data fidelity units of the unrolled network, Λ used to define self-supervised loss. The third set Γ defines self-validation loss to determine an early stopping criteria and avoid overfitting. Using multi-mask data-augmentation strategy [22], the ZS-SSDU training loss is given as:

$$\min_{\boldsymbol{\theta}} \frac{1}{K} \sum_{k=1}^{K} \mathcal{L}(\mathbf{y}_{\Lambda_k}, \mathbf{E}_{\Lambda_k}(f(\mathbf{y}_{\Theta_k}, \mathbf{E}_{\Theta_k}; \boldsymbol{\theta})), \qquad (4)$$

where $\boldsymbol{\theta}$ denotes the network parameters, $f(\mathbf{y}_{\Theta_k}, \mathbf{E}_{\Theta_k}; \boldsymbol{\theta})$ is the network output for input $\mathbf{y}_{\Lambda_k}$ and corresponding forward operator $\mathbf{E}_{\Lambda_k}$, and $\mathcal{L}(\cdot, \cdot)$ is a loss function. This is supplemented with a self-validation loss on Γ, which is calculated at each epoch j from the current network weights specified by $\boldsymbol{\theta}^{(j)}$ as follows:

$$\mathcal{L}(\mathbf{y}_\Gamma, \mathbf{E}_\Gamma(f(\mathbf{y}_{\Omega\backslash\Gamma}, \mathbf{E}_{\Omega\backslash\Gamma}; \boldsymbol{\theta}^{(j)}))). \qquad (5)$$

For single dataset training, the loss in Eq. 4 keeps decreasing. The training is stopped once the self-validation loss in Eq. 5 starts increasing to avoid overfitting. A schematic of the ZS-SSDU is depicted in Fig. 1.

III. EXPERIMENTS AND IMPLEMENTATION DETAILS

A. Imaging Experiments

Imaging was performed at 1.5T, first at lower resolution with mild acceleration to investigate performance in a retrospective acceleration setting, and then at high resolution with high acceleration to establish translational feasibility.
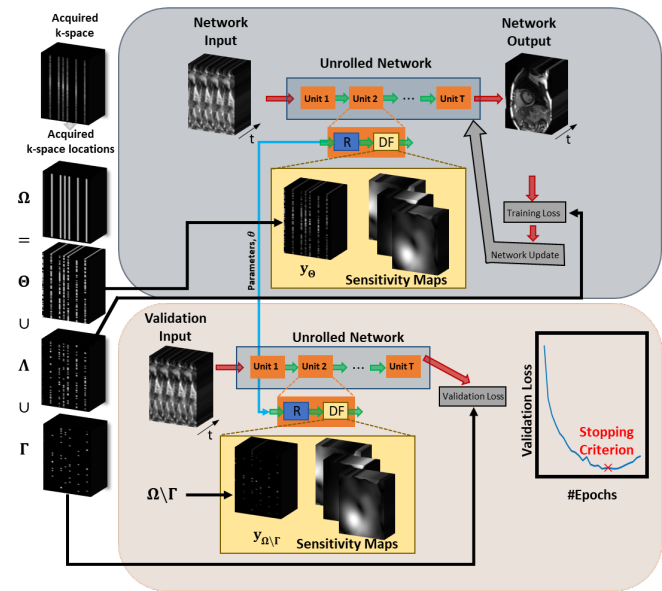

Fig. 1: A schematic of zero-shot PG-DL reconstruction. Acquired k-space locations (Ω) are split into three disjoint sets Θ, Λ and Γ. Θ is used in data-fidelity (DF) units of the unrolled network, Λ is used to define k-space self-supervised loss, and Γ is used for k-space self-validation loss to determine a stopping criterion for subject-specific learning. PG-DL network was unrolled for 10 iterations and a 3D ResNet was used as regularizer (R), with all 50 time-frames as input and 128 channels in hidden layers.

1) Retrospectively Accelerated Datasets: These were acquired on 11 subjects using a bSSFP sequence at acceleration R = 4 with a Cartesian TGRAPPA sampling pattern [4]. Relevant imaging parameters: FOV=360×270mm^2, resolution=2.25×2.93mm^2, and slice-thickness=8mm (11 slices). These datasets were further retrospectively undersampled to R = 7.2 by sampling every 8th line, while keeping the closest k_y line to central k-space for each time-frame.

2) Prospectively Accelerated High-Resolution Datasets: These were acquired on 2 subjects using a Cartesian bSSFP sequence with R = 8 acceleration using TGRAPPA sampling, along with corresponding breath-hold cine MRI. Relevant imaging parameters: slice-thickness=8mm (11 slices); real-time: FOV=360×340mm^2, resolution=1.5×1.5mm^2, temporal resolution=44ms; ECG-gated: FOV=360×360mm^2, resolution=1.4×1.4mm^2, temporal resolution=42ms.

B. Implementation Details

Subject-specific regularization across time-frames was performed using ZS-SSDU [15] using all cardiac cycles. A 3D PG-DL network was unrolled for 10 iterations with a 3D ResNet used for the regularizer where 50 time-frames in its inputs and 128 channels in its hidden layers. Training was performed with Adam optimizer with a learning rate of $4 \cdot 10^{-4}$ and a mixed normalized $\ell_1 - \ell_2$ loss in k-space [12]. R adjacent frames were used to generate calibration data in a sliding-window manner [4], [5], on which coil sensitivity maps were generated using ESPIRiT. 20% of Ω was uniformly randomly selected for Γ [15] and $K = 100$ was used for disjoint training pairs.

Comparisons were made to TGRAPPA [4], locally-low-rank (LLR) regularized reconstruction [16] and database-trained SSDU [12]. TGRAPPA was implemented using 5×4 kernels on calibration data generated from R adjacent frames [4]. For LLR regularized reconstruction, 8×8 patches were used with a thresholding value of 0.1 times the ℓ_∞ norm of the zero-filled image. Additionally, for the lower resolution datasets, database training with multi-mask SSDU [22] was used for comparisons. This was trained on 7 subjects using all available slices and 50 time-frames shifted by 20 frames with a total of 365 k_x-k_y-t k-spaces over 100 epochs. It used the same network architecture as in ZS-SSDU, and testing was performed on 2 different subjects not used in training. Database training could not be performed for the high-resolution datasets due to the small number of training samples. For lower resolution data, TGRAPPA at standard R = 4 was used as baseline images for comparisons, and PSNR and SSIM were calculated with respect to these images. These were assessed using paired t-test with $P < .05$ considered significant.

IV. RESULTS

A. Retrospectively Accelerated Datasets

Fig. 2a shows reconstructed time-frames for R = 7.2 accelerated real-time dataset. TGRAPPA reconstruction at R = 4 acquisition is depicted on the leftmost column for baseline image quality. At R = 7.2, TGRAPPA shows residual artifacts and noise amplification, whereas LLR-regularized reconstruction shows spatio-temporal blurring. Database-trained SSDU shows substantial residual aliasing artifacts due to varying spatio-temporal correlations among subjects in the database. The proposed database-free ZS-SSDU improves all methods, closely matching the baseline R = 4 image. Fig. 2b depicts the difference images to baseline image at R = 4, where proposed methods only show noise-like differences. Temporal signal intensity profiles are shown in Fig. 2c, where database-trained SSDU and TGRAPPA show visible aliasing, while LLR-regularized reconstruction shows substantial temporal blurring. The proposed method closely matches with the baseline profile.

Quantitative PSNR and SSIM metrics across all slices and cardiac cycles for two subjects are in concordance with the previous visual assessments: Database-trained SSDU has the lowest PSNR and SSIM (21.58 ± 3.32 (dB), 54.49 ± 13.80 %), followed by TGRAPPA (22.56 ± 1.75, 73.50 ± 4.06 %) and LLR-regularized reconstruction (28.14 ± 1.29, 81.44 ± 2.83 %). The proposed method outperforms all methods (all pairwise t-tests: $P < 10^{-4}$) with substantial gains in PSNR and SSIM (33.71 ± 1.94, 89.76 ± 2.07 %).

B. Prospectively Accelerated High-Resolution Datasets

Fig. 3a shows a representative reconstructed time-frame for a prospectively R = 8 accelerated real-time cine MRI acquisition. ECG-gated breath-hold segmented cine MRI is shown as baseline image quality on the leftmost column. As in Section IV-A, TGRAPPA suffers from aliasing, noise and signal void, and LLR-regularized reconstruction exhibits
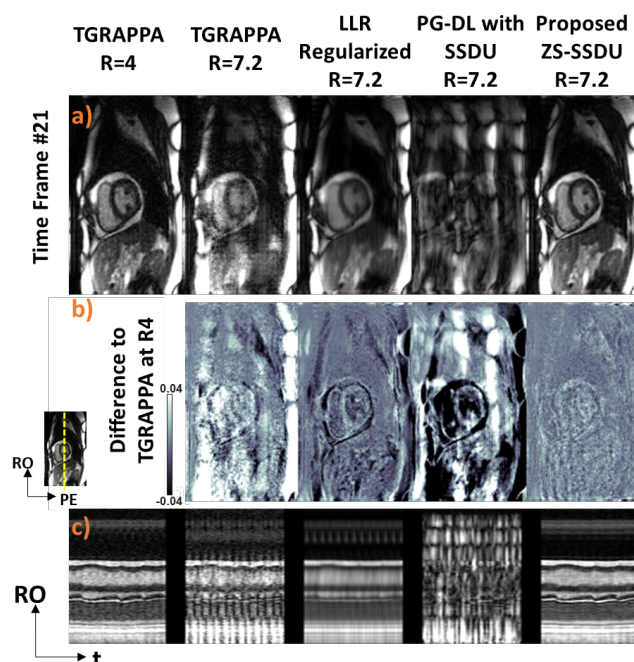

Fig. 2: Representative real-time cine MRI results with retrospective acceleration, along with TGRAPP at acquired R = 4 for visual baseline. **(a)** SSDU with database training suffers from severe aliasing artifacts, TGRAPPA shows residual aliasing and noise amplification, whereas LLR regularization exhibits blurring and signal void in the anterolateral segment. Proposed ZS-SSDU at R = 7.2 substantially improves upon all methods, closely matching acquisition acceleration R = 4 data. **(b)** Difference images, with R = 4 TGRAPPA, align with these observations, where our proposed ZS-SSDU only shows noise-like artifacts, while all other methods show structural differences. **(c)** Temporal signal intensity profiles across the blood–myocardium boundary (yellow dashed line) show that proposed ZS-SSDU has similar temporal profiles compared to R = 4 data, with similar cardiac and respiratory motion profiles.

blurring and signal voids. The proposed ZS-SSDU improves upon all methods and shows excellent image quality compared to breath-hold cine MRI. Fig. 3b shows temporal signal intensity profiles, where our proposed ZS-SSDU for R= 8 free-breathing real-time cine MRI closely matches with baseline profile of breath-hold segmented cine MRI. Note since these came from different acquisitions, exact alignment was not possible. Finally, we note database training of PG-DL reconstruction was not performed due to the small number of training subjects.

V. DISCUSSION

In this study, we proposed a subject-specific zero-shot PG-DL reconstruction with spatio-temporal regularization, which does not require a database to train, for highly-accelerated free-breathing real-time cine cardiac MRI. The main advantage of using ZS-SSDU is addressing robustness and generalizability issues of database-trained models [23] with regards to variations in subject-specific breathing and cardiac motions. The proposed approach improved upon a linear reconstruction, a conventional regularized reconstruction, and database-trained PG-DL reconstruction with better
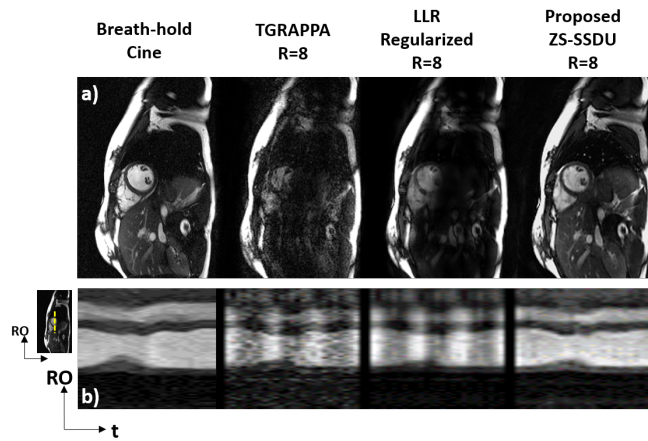
Fig. 3: Representative real-time cine MRI results for a prospectively R = 8 accelerated high-resolution dataset. Breath-hold ECG-gated segmented cine MRI from a separete scan is presented on the leftmost column for visual baseline. **(a)** TGRAPPA shows aliasing artifacts and LLR-regularized reconstruction shows signal void and spatial blurring, while proposed ZS-SSDU shows excellent image quality, closely matching to breath-hold segmented cine acquisition. **(b)** Temporal signal intensity profiles show excellent match between the proposed method and breath-hold cine, whereas other methods fail to preserve temporal information.

image quality and reduced aliasing.

The database-trained SSDU in this study is trained on 365 k-spaces with shifted time kernels for data augmentation. Even with the data augmentation, the results of database-trained SSDU show residual aliasing indicating database learning of spatiotemporal correlations is difficult. Moreover, the results show that when subject-specific breathing patterns and cardiac motion is presented, database training methods have poor generalizability. Further studies with improved spatio-temporal resolutions in a larger cohort are warranted.

VI. CONCLUSIONS

The proposed database-free PG-DL reconstruction enabled high-quality spatio-temporal regularization for highly-accelerated real-time Cartesian cine cardiac MRI, learning breathing and cardiac motions reliably without temporal blurring. and improved upon existing techniques, including database-trained PG-DL.

ACKNOWLEDGMENT

REFERENCES

[1] M. Francone, S. Dymarkowski, M. Kalantzi, and J. Bogaert, "Real-time cine MRI of ventricular septal motion: a novel approach to assess ventricular coupling," *J Magn Reson Imaging*, vol. 21, no. 3, pp. 305–309, 2005.

[2] S. Schalla, C. Klein, et al., "Real-time MR image acquisition during high-dose dobutamine hydrochloride stress for detecting left ventricular wall-motion abnormalities in patients with coronary arterial disease," *Radiology*, vol. 224, no. 3, pp. 845–851, 2002.

[3] Y. Y. Li, S. Rashid, et al., "Temporospatial characterization of ventricular wall motion with real-time cardiac MRI in health and disease," *Sci Rep*, vol. 12, no. 1, pp. 4070, 2022.

[4] F. A. Breuer, P. Kellman, M. A. Griswold, and P. M. Jakob, "Dynamic autocalibrated parallel imaging using temporal GRAPPA (TGRAPPA)," *Magn Reson Med*, vol. 53, no. 4, pp. 981–985, 2005.

[5] P. Kellman, F. H. Epstein, and E. R. McVeigh, "Adaptive sensitivity encoding incorporating temporal filtering (TSENSE)," *Magn Reson Med*, vol. 45, no. 5, pp. 846–852, 2001.

[6] N. Seiberlich, P. Ehses, J. Duerk, R. Gilkeson, and M. Griswold, "Improved radial GRAPPA calibration for real-time free-breathing cardiac imaging," *Magn Reson Med*, vol. 65, pp. 492–505, 2011.

[7] L. Feng, L. Axel, et al., "XD-GRASP: golden-angle radial mri with reconstruction of extra motion-state dimensions using compressed sensing," *Magn Reson Med*, vol. 75, no. 2, pp. 775–788, 2016.

[8] A. Hauptmann, S. Arridge, F. Lucka, V. Muthurangu, and J. A. Steeden, "Real-time cardiovascular MR with spatio-temporal artifact suppression using deep learning–proof of concept in congenital heart disease," *Magn Rson Med*, vol. 81, no. 2, pp. 1143–1156, 2019.

[9] 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.

[10] 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, 2018.

[11] 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.

[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, pp. 3172–3191, 2020.

[13] M. A. Morales, S. Assana, et al., "An inline deep learning based free-breathing ECG-free cine for exercise cardiovascular magnetic resonance," *J Cardiovasc Magn Reson*, vol. 24, no. 1, pp. 1–14, 2022.

[14] A. Rehman, P. Kellman, et al., "Convolutional neural network transformer (CNNT) for free-breathing real-time cine imaging," *Eur Heart J Cardiovasc Imaging*, vol. 23, pp. jeac141–001, 2022.

[15] B. Yaman, S. A. H. Hosseini, and M. Akcakaya, "Zero-shot self-supervised learning for MRI reconstruction," in *International Conference on Learning Representations*, 2022.

[16] X. Miao, S. G. Lingala, et al., "Accelerated cardiac cine MRI using locally low rank and finite difference constraints," *Magn Reson Imaging*, vol. 34, no. 6, pp. 707–714, 2016.

[17] 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.

[18] M. Akçakaya, B. Yaman, H. Chung, and J. C. Ye, "Unsupervised deep learning methods for biological image reconstruction and enhancement: an overview from a signal processing perspective," *IEEE Signal Processing Magazine*, vol. 39, no. 2, pp. 28–44, 2022.

[19] B. Yaman, C. Shenoy, et al., "Self-supervised physics-guided deep learning reconstruction for high-resolution 3D LGE CMR," in *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2021, pp. 100–104.

[20] O. B. Demirel, B. Yaman, et al., "Signal intensity informed multi-coil encoding operator for physics-guided deep learning reconstruction of highly accelerated myocardial perfusion CMR," *Magnetic resonance in medicine*, vol. 89, no. 1, pp. 308–321, 2023.

[21] O. B. Demirel, B. Yaman, et al., "20-fold accelerated 7T fMRI using referenceless self-supervised deep learning reconstruction," *arxiv preprint arXiv:2105.05827*, 2021.

[22] B. Yaman, H. Gu, et al., "Multi-mask self-supervised learning for physics-guided neural networks in highly accelerated magnetic resonance imaging," *NMR in Biomedicine*, vol. 35, pp. e4798, 2022.

[23] M. Akçakaya, S. Moeller, S. Weingärtner, and K. Uğurbil, "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, 2019.