Dynamic-Aware Spatio-temporal Representation Learning for Dynamic MRI Reconstruction

Dayoung Baik[1][0009-0001-1393-2195] and Jaejun Yoo[1][0000-0001-5252-9668]*

Ulsan National Institute of Science and Technology, Ulsan, Republic of Korea
da981116@unist.ac.kr, jaejun.yoo@unist.ac.kr

Abstract. Dynamic MRI reconstruction, one of inverse problems, has seen a surge by the use of deep learning techniques. Especially, the practical difficulty of obtaining ground truth data has led to the emergence of unsupervised learning approaches. A recent promising method among them is implicit neural representation (INR), which defines the data as a continuous function that maps coordinate values to the corresponding signal values. This allows for filling in missing information only with incomplete measurements and solving the inverse problem effectively. Nevertheless, previous works incorporating this method have faced drawbacks such as long optimization time and the need for extensive hyperparameter tuning. To address these issues, we propose Dynamic-Aware INR (DA-INR), an INR-based model for dynamic MRI reconstruction that captures the spatial and temporal continuity of dynamic MRI data in the image domain and explicitly incorporates the temporal redundancy of the data into the model structure. As a result, DA-INR outperforms other models in reconstruction quality even at extreme undersampling ratios while significantly reducing optimization time and requiring minimal hyperparameter tuning. Our code is available at here.

Keywords: Dynamic MRI reconstruction, · Deep learning, · Unsupervised learning, · Implicit Neural Representation.

1 Introduction

Dynamic Magnetic Resonance Imaging (MRI) captures sequential images of moving organs, such as the heart, while Dynamic Contrast Enhanced (DCE) MRI monitors temporal changes in in-vivo drug effects on vasculature. Due to the slow acquisition speed of MRI, only partial data can be collected per frame, leading to a trade-off between spatial and temporal resolution. Recent approaches have accelerated data acquisition while maintaining image quality by exploiting sparsity in the spatial and temporal domains [4, 8, 13]. Early deep learning methods [6, 10, 11, 18] applied supervised learning, but required large amounts of paired undersampled and fully sampled data, limiting their practicality.

To address this, unsupervised learning methods have emerged, leveraging inherent priors in Convolutional Neural Networks (CNNs) [19] and Implicit Neural Representations (INRs) [7, 9]. CNN-based approaches, such as [19], exploit

* Corresponding author

the structural prior of randomly initialized CNNs to capture low-level image statistics, which serves as an implicit regularization across all frames. However, the use of discrete grid representations in CNN constrains their ability to fully capture the continuous nature of dynamic MRI data. In contrast, INR-based methods [7, 9] represent dynamic MRI data as a continuous neural function in both spatial and temporal dimensions. By optimizing this function using spatio-temporal coordinates as inputs and predicting the corresponding values based on the available measurements, INR models effectively infer missing information during the reconstruction process, enabling full data recovery.

Specifically, Neural Implicit k-space (NIK) [7] introduced learning neural representation in the frequency domain to avoid regridding loss, but aliasing artifacts remained evident in the reconstructed images. The Fourier feature MLP (FMLP) [9] used a Fourier feature encoder [17] without requiring explicit regularization terms and showed superior performance over previous methods. However, a common drawback across all these approaches is the lengthy optimization process, which can take several hours to an entire day for networks to converge. More recent work [3] replaces the Fourier feature encoder with a hash encoder [15] to achieve faster convergence. However, this approach remains time-intensive due to complexity of tuning hyperparameters for the hash encoders and regularization terms required for spatial and temporal consistency. Moreover, the results are highly sensitive to the weighting of these regularization terms.

To address these challenges, we propose Dynamic-Aware INR (DA-INR), which explicitly models the temporal redundancy inherent in dynamic MRI data, inspired by D-NeRF [16]. It circumvents the need for manual weighting of regularization terms by making canonical space play as a regularization role to the other frames during optimization. Thus, it enables more stable convergence than a hash encoder alone. As a result, DA-INR not only simplifies the training process, but also enhances adaptability to diverse undersampling conditions and data complexities, offering an efficient solution for dynamic MRI reconstruction.

2 Method

2.1 Dynamic-aware INR

We propose Dynamic-aware INR (DA-INR), which accelerates the optimization process in dynamic MRI reconstruction through a novel dynamic hash encoding scheme. In this section, we provide an overview of (1) the overall workflow, (2) the core components of the framework, and (3) the optimization strategy.

Overall workflow DA-INR consists of three learning stages (Fig.1). The framework operates within the canonical space, which serves as a reference coordinate system that captures the static structure of the dynamic MRI data. The input coordinate (x, y, t) is encoded by frequency encoding [14] and passed into the deformation network Ψ_t which outputs the deformation field $(\Delta x, \Delta y)$ based on the canonical space. The pretrained feature extractor extracts image features

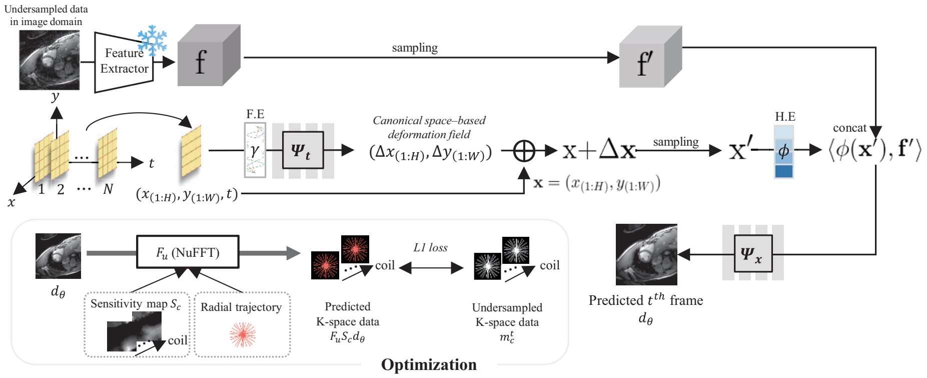

Undersampled data in image domain
Feature Extractor
f
sampling
f′
y
x
1 2 ⋯ N
t
$(x_{(1:H)}, y_{(1:W)}, t)$
F.E
γ
Ψ_t
Canonical space–based deformation field
$(\Delta x_{(1:H)}, \Delta y_{(1:W)})$
$\mathbf{x} + \Delta \mathbf{x}$
sampling
$\mathbf{x}'$
H.E
ϕ
concat
$\langle \phi(\mathbf{x}'), \mathbf{f}' \rangle$
$\mathbf{x} = (x_{(1:H)}, y_{(1:W)})$
d_θ
F_u (NuFFT)
Sensitivity map S_c
coil
Radial trajectory
coil
Predicted K-space data $F_u S_c d_\theta$
L1 loss
coil
Undersampled K-space data m_c^t
Optimization
Ψ_x
Predicted t^{th} frame d_θ

Fig. 1: DA-INR model architecture. It maps coordinates to the corresponding value with image features and displacement vector based on canonical space.

from the undersampled data in the image domain. Then, the canonical network Ψ_x takes the deformed coordinate $(x + \Delta x, y + \Delta y)$ and the image features as input and predicts the corresponding value within the canonical space.

Deformation network The deformation network Ψ_t estimates the deformation field between cells at a specific time t and cells in the canonical space. More specifically, given the input coordinate $\mathbf{x} = (x, y)$ at time t, Ψ_t predicts the deformation field $\Delta\mathbf{x}$ to transform the cell position (x, y) to the cell position $(x + \Delta x, y + \Delta y)$ in the canonical space. Before going into Ψ_t, $\mathbf{x}$ and t is encoded by the frequency encoding, $\gamma(p) = \left\langle (sin(2^i \pi p), cos(2^i \pi p)) \right\rangle_0^I$ [14]. It is applied to each component of the input coordinate with $I = 10$ and the time component with $I = 6$.

Feature extraction In medical imaging or inverse problems with limited data, it is common to use models pre-trained on large-scale natural images for feature extraction [2, 12]. We use the pretrained image feature extractor, MDSR [5], to provide auxiliary features for dynamic MRI reconstruction[1]. The frozen feature extractor takes the undersampled data in the image domain at time t (spatial interpolation, reconstruction) or of two neighboring frames of time t (temporal interpolation) as input and outputs the image features $\mathbf{f} \in \mathbb{R}^{c' \times H \times W}$ whose size is the same as the input frame. c' is the size of the channel dimension. We upscale $\mathbf{f}$ to $\mathbf{f}' \in \mathbb{R}^{c' \times rH \times rW}$ by bilinear interpolation, and $\mathbf{x} + \Delta\mathbf{x}$ to $\mathbf{x}' \in \mathbb{R}^{2 \times rH \times rW}$ by nearest-neighborhood interpolation based on the scale ratio r. During optimization, r is fixed as 1, $r = 1$. During inference of spatial interpolation, r is bigger than 1, $r > 1$.

[1] We compared various encoders, but omitted results due to space limits. Please refer to our project page.

Canonical network The canonical network Ψ_x predicts the corresponding image intensity value in the canonical space, given the resampled deformed coordinate $\mathbf{x}'$ and the image features $\mathbf{f}'$. The input $\mathbf{x}'$ is first encoded by a hash encoder ϕ [15] and is concatenated with $\mathbf{f}'$ in the channel dimension. Then, they are fed into Ψ_x to output the corresponding image intensity value in the cell position of the canonical space.

Optimization We only use L1 loss as data-consistency for optimization. The final loss $\mathcal{L}$ is defined as follows:

$$\mathcal{L} = \sum_{c=1}^{C} ||F_u S_c d_\theta - m_c^t||_1^1, \tag{1}$$

where d_θ is the reconstructed image by DA-INR defined , m_c^t is a c^{th} coil golden-angle radial undersampled k-space data at time t. F_u is NuFFT operator with a given radial trajectory and multi-coil sensitivity map.

2.2 Difference between Feng et al. and DA-INR for encoding temporal redundancy

Feng et al. [3] learns to map (x, y, t) directly to the corresponding value in the image domain and its optimization is operated on under-sampled k-space data at each frame individually. Hence, it leads to over-fitting to each under-sampled frame when the explicit regularization terms are not used. DA-INR enforces temporal consistency via a shared canonical space jointly optimized with Ψ_t across the sequence. Though described separately, Ψ_t and Ψ_x act as a unified module: Ψ_x models canonical signal across time, while Ψ_t learns to deform it per frame. The canonical space is updated using spokes from all frames, aggregating structural details and forming a complete high-frequency representation. Ψ_t adjusts this to each frame, accounting for dynamic deviations. This acts like multi-view regularization that emerges from the framework itself rather than handcrafted priors, removing manual hyperparameter tuning.

3 Experiments

3.1 Baseline methods

We compare our method against Non-uniform Fast Fourier Transform (NuFFT), GRASP [4], TD-DIP [19], and the method proposed by Feng et al. [3]. Feng et al. [3] is an INR-based method with hash encoding that incorporates explicit temporal TV and low-rank regularization terms for optimization and set different values on the weights of the regularization terms according to data types and acceleration factors. However, using the official code of [3], we found the results highly sensitive to regularization weights—small change (e.g., ±0.05) all led to noisy or black images. Despite extensive tuning, stable reproduction was infeasible[2]. To ensure reproducibility and isolate the method's inherent behavior, we ran it without regularization. This adaptation is referred to as *HashINR* in our paper.

3.2 Datasets

Retrospective cardiac cine data was obtained using a 3T whole-body MRI scanner (Siemens Tim Trio) equipped with a 32-element cardiac coil array. The full-sampled k-space data is used as ground truth (GT). To simulate a retrospective undersampling pattern, we adopt a 2D golden-angle radial acquisition scheme, where the spokes repeatedly traverse the center of k-space, rotating with a step of 111.25°. It is applied to ground truth with multi-coil NuFFT to obtain the undersampled radial trajectories of Fibonacci numbers [1]. **Dynamic Contrast-Enhanced (DCE) liver data** scan was conducted on a healthy volunteer using axial orientation and breath-holding techniques by a whole-body 3 Tesla MRI system (MAGNETOM Verio/Avanto, Siemens AG, Erlangen, Germany), employing a combination of body-matrix and 12-element spine coil array. For data acquisition, a radial stack-of-stars 3D Fast Low Angle Shot (FLASH) pulse sequence with golden-angle ordering was utilized.

3.3 Performance evaluation

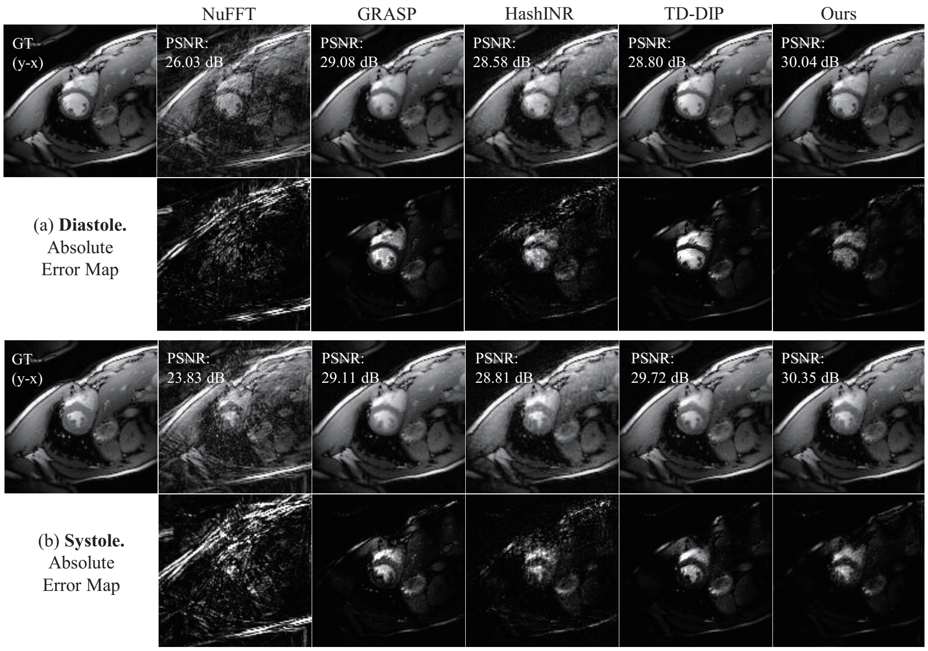
For cardiac cine data, we use Peak Signal-to-Noise Ratio (PSNR) and structural similarity index (SSIM) as evaluation metrics, both calculated frame-by-frame. For DCE liver data, we conduct a visual comparison and assess temporal fidelity based on the signal intensity in the region of interest (ROI), as no ground truth images are available. The ROIs for the aorta (AO) and portal vein (PV) are manually drawn for each signal intensity flowmap. We use NuFFT as a reference because the contrast changes can be preserved due to the average signal intensity over a large ROI. We test the performance of each method with 21, 13, and 5 spokes per frame ($AF = 6.1, 9.8, 25.6$) on cardiac cine data, and with 34 spokes per frame ($AF = 11.3$) on DCE liver data.

3.4 Implementation details

Ψ_t and Ψ_x are 5-layer MLPs (64 hidden units, ReLU); Ψ_t and Ψ_x predicts deformation vectors and complex-values each. The non-Cartesian Fourier undersampling operation is executed using the NuFFT package[3], facilitating rapid computation and gradient backpropagation on a GPU. Feng et al. [3] tunes the hyperparameters L, T, F, N_{min}, and b of its hash encoder according to each data type and AF, while we use consistent values across all data types and AFs as $L = 16, T = 2^{19}, F = 2, N_{min} = 16$, and $b = 2$.

[2] To reproduce these experiments under identical setups including the instability we saw, please refer to our project page

[3] https://github.com/dfm/python-nufft



(a) **Diastole.** Absolute Error Map

(b) **Systole.** Absolute Error Map

Fig. 2: Visual comparisons between results of $AF = 9.8$ in cardiac cine data reconstruction at diastole and systole. The upper row is the reconstruction output in the $(y - x)$ domain for each method and the below row is the absolute error map between ground truth and the reconstructed output of each method. PSNR values are specific to each frame.

4 Results

4.1 Retrospective cardiac cine data

Fig. 2 presents the visual comparisons between our method and the existing methods for cardiac cine data reconstruction at $AF = 9.8$. We evaluate the reconstructed frames for each method during the diastolic and systolic phases. NuFFT struggles to accurately capture the cardiac structure. HashINR shows noise in its results. While GRASP sometimes achieves high PSNR, it struggles to reconstruct fine structural details, such as the shape of the papillary muscle, under both undersampling ratios. In contrast, our method closely approximates the ground truth, achieving high fidelity in both phases. The quantitative error—computed as the sum of the squared differences between the reconstructed and ground truth images—is the smallest among all methods.

Tab. 1 reports the quantitative results analyzed on cardiac cine data. The reconstruction quality of NuFFT and HashINR is highly dependent on the number of spokes, resulting in substantial gaps in PSNR and SSIM values for $AF = 25.6$ and $AF = 9.8$, with differences ranging from 2.66 to 3.5 dB in PSNR and 0.1418

Table 1: Quantitative results of cardiac cine data reconstruction. We compare ours to NuFFT, GRASP, TD-DIP, and HashINR at $AF = 25.6$ and $AF = 9.8$.

Method	Undersampling ratio	PSNR (dB)	SSIM
NuFFT	$AF = 25.6$	21.58 ± 0.0	0.3809 ± 0.0
	$AF = 9.8$	25.08 ± 0.0	0.5250 ± 0.0
GRASP	$AF = 25.6$	27.95 ± 0.59	0.8410 ± 0.005
	$AF = 9.8$	28.77 ± 0.48	0.8558 ± 0.003
TD-DIP	$AF = 25.6$	28.73 ± 0.55	0.8608 ± 0.003
	$AF = 9.8$	28.86 ± 0.46	0.8714 ± 0.002
HashINR	$AF = 25.6$	24.01 ± 0.54	0.6226 ± 0.003
	$AF = 9.8$	26.68 ± 0.44	0.7643 ± 0.001
Ours	$AF = 25.6$	29.48 ± 0.51	0.8702 ± 0.002
	$AF = 9.8$	30.09 ± 0.47	0.8805 ± 0.001

to 0.1441 in SSIM. In contrast, our results show relatively consistent reconstruction quality in both conditions by learning to reconstruct the canonical space in every iteration. Reconstructed images at specific time points are then obtained by warping the canonical space with the deformation field estimated based on temporal differences, leading to stable convergence in any condition. Our results achieve the best PSNR and SSIM values, 29.59 dB/0.8712 and 30.13 dB/0.8835 in $AF = 25.6$ and $AF = 9.8$, respectively.

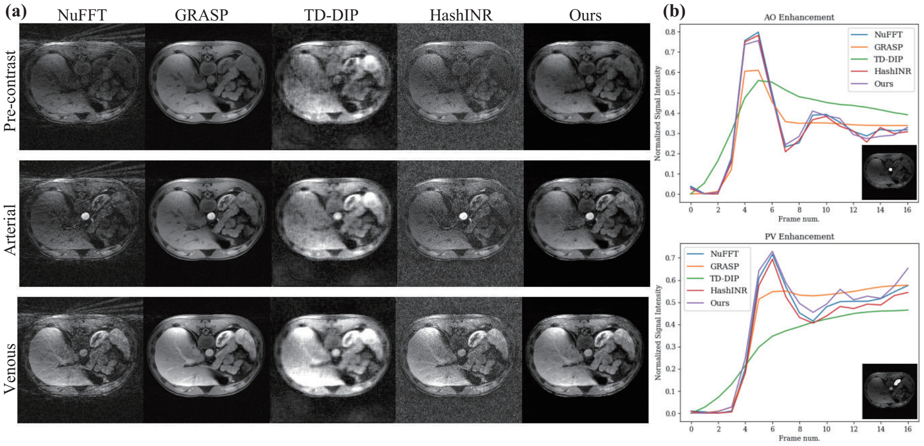


Fig. 3: Qualitative results of DCE liver data reconstruction with an undersampling ratio of 34 spokes per frame ($AF = 11.3$). (a) We visualize reconstruction at different contrast phases, and (b) compare signal intensity flow for aorta (AO) and portal vein (PV) ROI.

4.2 Dynamic Contrast-Enhanced (DCE) liver data

Fig. 3 presents the qualitative reconstruction results and the corresponding signal intensity flowmap for DCE liver data reconstruction performed with an undersampling ratio of 34 spokes per frame ($AF = 11.3$). While GRASP appears clean in the $y - x$ domain, the temporal signal curves in Fig. 3 (b) (orange line) show it fails to capture dynamic contrast changes, overfitting to a few frames. This highlights its limitation in modeling temporal variation, which is essential in dynamic MRI. On the other hand, HashINR exhibits good temporal fidelity in the flowmap, but its outputs display noticeable noise in the $(y - x)$ domain, overfitting to the undersampled frames. TD-DIP produces the reconstructions characterized by overly smooth appearances. This smoothness results in a failure to accurately delineate fine structural and contrast changes (Fig. 3 (a)). Consequently, it achieves the lowest temporal fidelity among the compared approaches, as reflected in the signal intensity flowmap in Fig. 3 (b). In contrast, our proposed method achieves a prominent performance, delivering both high temporal fidelity and phase-specific contrast changes. The reconstructions in Fig. 3 (a) are well-defined, showing clear structural details and accurate phase enhancements. Furthermore, our signal intensity flowmap in Fig. 3 (b) shows a strong capacity of our model to preserve temporal dynamics, capturing the changes in signal intensity over time with high accuracy.

4.3 Time consumption and GPU memory usage

Table 2 presents the comparison of runtime and GPU memory usage of every method for dynamic MRI reconstruction at $AF = 9.8$ with GeForce RTX 4090. GRASP requires 2.7 GB of GPU memory for cardiac cine data and 11.6 GB for DCE liver data because its cost on GPU memory depends on the image sequence size. TD-DIP utilizes the least GPU memory, but has the longest reconstruction time for cardiac cine data. HashINR takes 6.97 to 7.25 times longer runtime than ours, along with significantly higher memory consumption for both datasets. In contrast, our method achieves the shortest optimization time among learning-based methods for both data reconstruction with comparatively low GPU memory usage.

5 Conclusion

In this paper, we propose a novel framework for spatio-temporal representation learning tailored to dynamic MRI reconstruction without requiring ground truth data. Our method, Dynamic-aware INR (DA-INR), combines the efficiency of hash encoding for rapid optimization with an explicit design inspired by D-NeRF to effectively capture continuous temporal redundancy. By leveraging a canonical network, DA-INR incorporates temporal consistency into its structure, reducing dependency on explicit regularization terms while ensuring fast convergence. Comprehensive experiments demonstrate that DA-INR achieves superior reconstruction quality and efficiency, making it a robust solution for dynamic MRI reconstruction even under extreme undersampling conditions.

Table 2: Runtime and GPU memory usage of different methods on each data, in the unit of seconds and gigabyte. It is examined on GeForce RTX 4090.

Data type	cardiac cine data 128 × 128, 23 frames, 32 coils		DCE liver data 384 × 384, 17 frames, 12 coils	
Method	Runtime (sec)	GPU memory (GB)	Runtime (sec)	GPU memory (GB)
NuFFT	13.36	0.7	18.96	0.7
GRASP	146.40	2.7	423.24	11.6
TD-DIP	14164.31	1.6	38581	2.4
HashINR	10484.32	10.1	58088	7.6
Ours	1445.50	3.5	8329.55	5.4

Acknowledgments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT)(No.2022R1C1C100849612), Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.RS-2020-II201336, Artificial Intelligence Graduate School Program (UNIST), No.2022-0-00959, No.RS-2022-II220959 (Part 2) Few-Shot Learning of Causal Inference in Vision and Language for Decision Making, No.RS-2022-II220264, Comprehensive Video Understanding and Generation with Knowledge-based Deep Logic Neural Network). This work was also supported by the 2025 Research Fund (1.250006.01) of UNIST(Ulsan National Institute of Science & Technology).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Chandarana, H., Feng, L., Block, T., Rosenkrantz, A., Lim, R., Babb, J., Sodickson, D., Otazo, R.: Free-breathing contrast-enhanced multiphase mri of the liver using a combination of compressed sensing, parallel imaging, and golden-angle radial sampling. Investigative radiology **48** (11 2012). https://doi.org/10.1097/RLI.0b013e318271869c
2. Fang, W., Tang, Y., Guo, H., Yuan, M., Mok, T.C., Yan, K., Yao, J., Chen, X., Liu, Z., Lu, L., et al.: Cycleinr: Cycle implicit neural representation for arbitrary-scale volumetric super-resolution of medical data. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11631–11641 (2024)
3. Feng, J., Feng, R., Wu, Q., Zhang, Z., Zhang, Y., Wei, H.: Spatiotemporal implicit neural representation for unsupervised dynamic mri reconstruction (2023)
4. Feng, L., Grimm, R., Block, K.T., Chandarana, H., Kim, S., Xu, J., Axel, L., Sodickson, D.K., Otazo, R.: Golden-angle radial sparse parallel mri: Combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric mri. Magnetic Resonance in Medicine **72**(3), 707–717 (2014). https://doi.org/https://doi.org/10.1002/mrm.24980, https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.24980
5. Gao, S., Zhuang, X.: Multi-scale deep neural networks for real image super-resolution. CoRR **abs/1904.10698** (2019), http://arxiv.org/abs/1904.10698

6. Han, Y., Yoo, J.J., Ye, J.C.: Deep learning with domain adaptation for accelerated projection reconstruction MR. CoRR **abs/1703.01135** (2017), http://arxiv.org/abs/1703.01135
7. Huang, W., Li, H., Pan, J., Cruz, G., Rueckert, D., Hammernik, K.: Neural implicit k-space for binning-free non-cartesian cardiac mr imaging (2023)
8. Jung, H., Sung, K., Nayak, K.S., Kim, E.Y., Ye, J.C.: k-t focuss: A general compressed sensing framework for high resolution dynamic mri. Magnetic Resonance in Medicine **61**(1), 103–116. https://doi.org/https://doi.org/10.1002/mrm.21757
9. Kunz, J.F., Ruschke, S., Heckel, R.: Implicit neural networks with fourier-feature inputs for free-breathing cardiac mri reconstruction (2024)
10. Lee, D., Yoo, J., Ye, J.C.: Compressed sensing and parallel mri using deep residual learning. In: The International Society for Magnetic Resonance in Medicine. ISMRM (2017)
11. Lee, D., Yoo, J., Ye, J.C.: Deep artifact learning for compressed sensing and parallel mri. arXiv preprint arXiv:1703.01120 (2017)
12. Li, G., Zhao, L., Sun, J., Lan, Z., Zhang, Z., Chen, J., Lin, Z., Lin, H., Xing, W.: Rethinking multi-contrast mri super-resolution: Rectangle-window cross-attention transformer and arbitrary-scale upsampling. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 21230–21240 (2023)
13. Lingala, S.G., Hu, Y., DiBella, E., Jacob, M.: Accelerated dynamic mri exploiting sparsity and low-rank structure: k-t slr. IEEE Transactions on Medical Imaging **30**(5), 1042–1054 (2011). https://doi.org/10.1109/TMI.2010.2100850
14. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: Nerf: Representing scenes as neural radiance fields for view synthesis. CoRR **abs/2003.08934** (2020), https://arxiv.org/abs/2003.08934
15. Müller, T., Evans, A., Schied, C., Keller, A.: Instant neural graphics primitives with a multiresolution hash encoding. ACM Trans. Graph. **41**(4), 102:1–102:15 (Jul 2022). https://doi.org/10.1145/3528223.3530127, https://doi.org/10.1145/3528223.3530127
16. Pumarola, A., Corona, E., Pons-Moll, G., Moreno-Noguer, F.: D-nerf: Neural radiance fields for dynamic scenes. CoRR **abs/2011.13961** (2020), https://arxiv.org/abs/2011.13961
17. Tancik, M., Srinivasan, P.P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J.T., Ng, R.: Fourier features let networks learn high frequency functions in low dimensional domains. CoRR **abs/2006.10739** (2020), https://arxiv.org/abs/2006.10739
18. Wang, S., Su, Z., Ying, L., Peng, X., Zhu, S., Liang, F., Feng, D., Liang, D.: Accelerating magnetic resonance imaging via deep learning. In: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). pp. 514–517 (2016). https://doi.org/10.1109/ISBI.2016.7493320
19. Yoo, J., Jin, K.H., Gupta, H., Yerly, J., Stuber, M., Unser, M.: Time-dependent deep image prior for dynamic mri (2021)