INFusion: Diffusion Regularized Implicit Neural Representations for 2D and 3D accelerated MRI reconstruction

Yamin Arefeen*†, Brett Levac*, Zach Stoebner*, Jonathan I Tamir*
*The University of Texas at Austin
Chandra Family Department of Electrical and Computer Engineering
†MD Anderson Cancer Center
Department of Imaging Physics

***Abstract*—Implicit Neural Representations (INRs) are a learning-based approach to accelerate Magnetic Resonance Imaging (MRI) acquisitions, particularly in scan-specific settings when only data from the under-sampled scan itself are available. Previous work demonstrates that INRs improve rapid MRI through inherent regularization imposed by neural network architectures. Typically parameterized by fully-connected neural networks, INRs support continuous image representations by taking a physical coordinate location as input and outputting the intensity at that coordinate. Previous work has applied unlearned regularization priors during INR training and have been limited to 2D or low-resolution 3D acquisitions. Meanwhile, diffusion-based generative models have received recent attention as they learn powerful image priors decoupled from the measurement model. This work proposes INFusion, a technique that regularizes the optimization of INRs from under-sampled MR measurements with pre-trained diffusion models for improved image reconstruction. In addition, we propose a hybrid 3D approach with our diffusion regularization that enables INR application on large-scale 3D MR datasets. 2D experiments demonstrate improved INR training with our proposed diffusion regularization, and 3D experiments demonstrate feasibility of INR training with diffusion regularization on 3D matrix sizes of $256 \times 256 \times 80$.**

I. Introduction

Magnetic Resonance Imaging (MRI) suffers from lengthy acquisition times, but sampling the acquired Fourier (k-space) data below the Nyquist rate and solving the ill-posed reconstruction problem accelerates acquisitions, reducing costs to patients. Widespread reconstruction algorithms include parallel imaging [1–3] and compressed sensing [4], but deep learning based techniques exploiting training databases yield state of the art results [5–10], with diffusion models receiving recent attention as they decouple the measurement model and learned prior.

The machinery of Implicit Neural Representations (INRs) [11,12] serves as an alternative learning-based approach for accelerating MRI, particularly in scan-specific settings where only data from the under-sampled scan itself are available. Previous INR work leverages traditional regularization [13], the inherent regularization of the dense neural network architecture and positional encoding [12,14–17], and periodic activations [11]. However, previous work applies unlearned regularization priors when applying INRs for MRI reconstruction and are mainly limited to 2D or low-resolution 3D acquisitions [13–17].

Drawing inspiration from recent work in computer vision that uses diffusion priors to reconstruct INR-representations of 3D scenes from few camera views [18], we present INFusion, a technique to regularize the optimization of INRs from under-sampled MR measurements with pre-trained diffusion models for improved image reconstruction. In addition, we propose a hybrid 3D approach with our diffusion regularization that enables INR application on large-scale 3D MR datasets.

II. Methods

A. MRI Inverse Problems with INRs

The model-based accelerated MRI inverse problem reads:

$$\arg\min_{x} ||y - Ax||_2^2 + \lambda R(x), \tag{1}$$

where $y \in \mathbb{C}^M$ are the discrete k-space measurements, $A \in \mathbb{C}^{M \times N}$ is the linear measurement model, $x \in \mathbb{C}^N$ is the discretized image to reconstruct, and $\lambda R(\cdot) : \mathbb{C}^N \rightarrow \mathbb{R}$ imposes regularization (e.g., sparsity, low-rank).

Let $I_\theta : \mathbb{R}^C \rightarrow \mathbb{C}$ be an INR parameterized by weights θ, which represents the image as a function of the continuous coordinate domain. Then, the following optimization problem estimates weights of the INR through comparison to the acquired discrete k-space measurements [13–17]:

$$\theta^* = \arg\min_{\theta} ||y - AI_\theta(r)||_2^2 + \lambda R(I_\theta(r)), \tag{2}$$

where $r \in \mathbb{R}^{N \times C}$ indexes the spatial coordinates of the discretized voxel locations of x. Evaluating $I_{\theta^*}(r)$ then yields the desired image.

B. INFusion: Regularizing INRs with Diffusion Models

Our proposed INFusion method regularizes the INR MRI inverse problem with pre-trained generative diffusion models as they are powerful priors that decouple from the measurement model [8-10] enabling applications with INRs. As it is computationally expensive to evaluate the diffusion model at every 3D coordinate, we also propose a stochastic regularizer

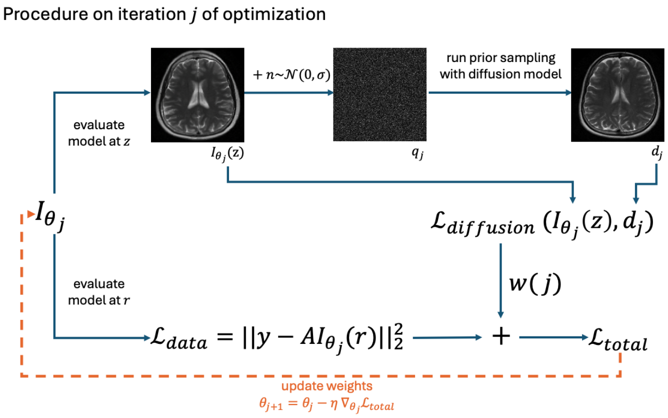

Fig. 1. Our proposed approach, INFusion, first computes standard data consistency loss L_{data} by evaluating the INR at discrete voxel locations, applying the MRI forward model to the resultant image, and then comparing to the acquired k-space data. A second loss $L_{diffusion}$ is computed by adding gaussian noise to random slices produced by the INR, running prior sampling with the diffusion model, and comparing the resultant prior sample to the INR produced random slices with perceptual loss. The diffusion loss guides the INR to produce images that better match the learned prior.

in which the diffusion model is queried at random spatial coordinates $z \subset r$ in each iteration.

During each iteration, j, of optimization, we propose the following procedure:

1: Compute data consistency loss, $L_{data} = ||y - AI_{\theta_j}(r)||_2^2$.
2: Select spatial coordinates z so that evaluating $I_{\theta_j}(z)$ yields random discretized 2D slices from the current estimate of image volume x.
3: Compute $q_j = I_{\theta_j}(z) + n$ where $n \sim \mathcal{N}(0, \sigma^2)$ and $\sigma \sim \mathcal{U}(\sigma_{min}, 1)$.
4: Run diffusion [19,20] with initialization q_j to give a prior sample d_j.
5: Compute $L_{diffusion} = LPIPS(I_{\theta_j}(z), d_j)$ with LPIPS perceptual loss [21].
6: Compute $L_{total} = L_{data} + w(j)L_{diffusion}$ with adaptive scalar weighting $w(j)$, and backpropagate to update θ_j.

Fig. 1 presents a schematic summary of the proposed procedure.

III. Experiments

A. In-vivo 2D Single- and Multi- Coil Brain

We used axial, T2-weighted, single-coil and 4-coil 2D k-space from the fastMRI [22] dataset, cropped both to 192×192 resolution, and retrospectively under-sampled with a 2D-Poisson pattern generated with BART [23] corresponding to an acceleration rate of $R = 4$ for single-coil (R times less than Nyquist) and $R = 6$ for multi-coil. Both experiments compared the discretized image produced by solving the standard MRI inverse problem with L1-Wavelet regularization [24] to an INR trained with (i) none, (ii) L1-Wavelet, or (iii) our proposed diffusion regularization.

Second, we took 96 additional, multi-coil 2D samples from the fastMRI dataset, and retrospectively under-sampled them with 2D-Poisson under-sampling masks with an acceleration rate of $R = 8$ and $R = 9$. These samples had coil counts ranging from $4 - 20$. We compared error of the discretized images produced by solving the standard MRI inverse problem with L1-Wavelet regularization and INRs with L1-Wavelet and our proposed diffusion regularization.

B. In-vivo small 3D Single Coil Brain and large 3D Multi-coil Knee

We used a multi-slice single-coil $192 \times 192 \times 16$ dataset from fastMRI with axial brain slices, and treated it as 3D by retrospectively under-sampling it in the $y - z$ direction with a $R = 2$ Poisson under-sampling mask.

Finally, we used k-space of an 8-coil 3D knee volume from the SKM-TEA [25] dataset, resized it to $256 \times 256 \times 80$, and uniformly under-sampled by $R = 4 \times 1$ in the $y - z$ directions. Fitting the problem in GPU memory and solving it in feasible time required two changes to our INR procedure:

- Only the $x - y$ dimensions were encoded with coordinates, and the model instead returned 80 discretized slice values at each coordinate.
- Diffusion regularization was computed with respect to two random slices each iteration.

Both 3D experiments compared discretized images produced by an INR trained with and without our INFusion method. Note that our diffusion model was trained on $x - y$ images, but under-sampling was performed in the $y - z$ direction.

C. INR Details

All INRs consisted of fully-connected neural-networks with ReLU activations and 128 gaussian Fourier-features [12]. 2D, 3D brain, and 3D knee networks had 165K, 1643K, and 21411K parameters, respectively.

D. Diffusion Model Details

We trained separate diffusion models, using an architecture and training procedure provided in [20], for the brain model with 8500 axial fastMRI slices and the knee with 15720 sagittal SKM-TEA slices.

IV. Results

INFusion achieved lower normalized-root-mean-squared-error (NRMSE) on the single (Fig. 3, A) and multi (Fig. 3, B) coil brain slice in comparison to INRs trained with Wavelet or without regularization and the standard Compressed Sensing (CS) L1-Wavelet MRI inverse problem. The box plot in Fig. 2 shows that INFusion improves reconstruction performance with respect NRMSE over 96 slices in comparison to standard CS L1-Wavelet and INRs with Wavelet regularization.

In Fig. 4, A, INFusion outperforms no regularization on the modest, single-coil 3D dataset, but encoding all three spatial dimensions in the INR required prohibitively large GPU memory usage and compute. Fig. 4, B shows that our proposed approach of diffusion regularization with random slices and just encoding the $x - y$ dimensions in coordinates enables training of INRs on $256 \times 256 \times 80$ matrix size k-space.

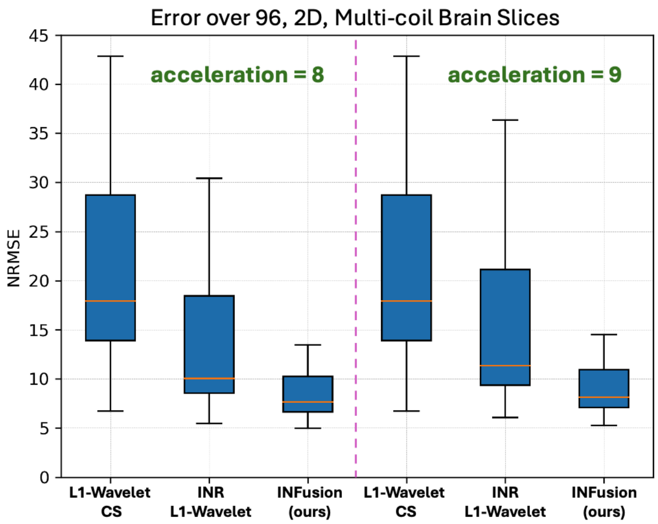

Fig. 2. Quantitative comparison of standard L1-Wavelet CS and INRs trained with Wavelet and our proposed regularization at $R = 8$ and $R = 9$ across 96 samples. Our proposed INFusion approach improves NRMSE across the dataset.

V. Discussion

INFusion exploits the generative capabilities of diffusion models to regularize INR training from scan-specific MR measurements. We show improved performance in 2D and enable application in realistic 3D settings where the diffusion models were also trained on a different orientation than the under-sampling. In this way, we regularize 3D reconstructions without an explicit 3D prior.

Future work will investigate the effectiveness of INRs as an image representation and explore the trade-offs between using diffusion models as regularization for INRs versus using them for posterior sampling directly on the discretized voxel grid.

Acknowledgment

This work was supported in part by NSF CCF-2239687 (CAREER), NSF IFML 2019844, and JCCO postdoctoral fellowship.

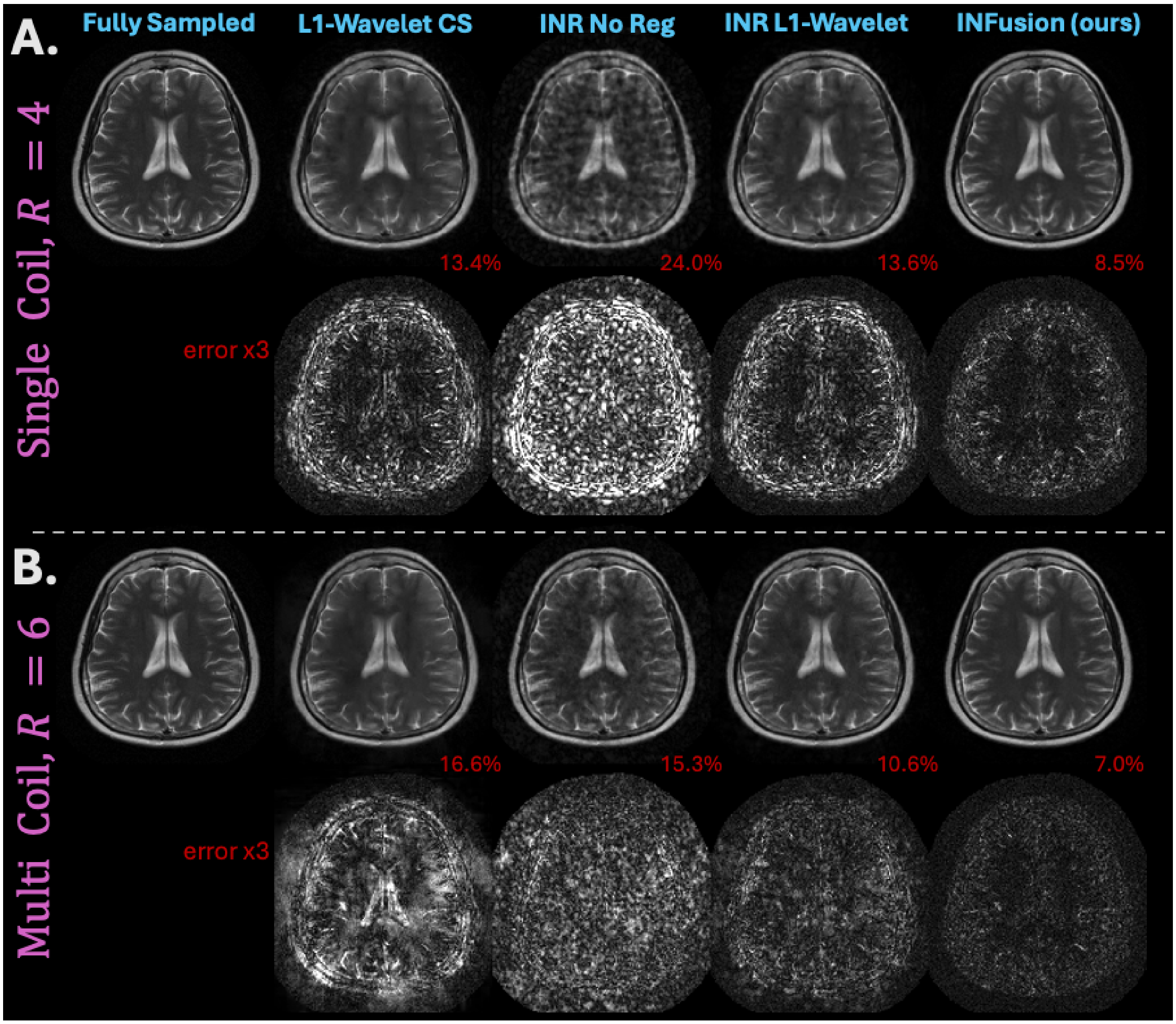

Fig. 3. Reconstructions of discrete images on the (A.) 2D single-coil brain data at $R = 4$ and (B.) 2D multi-coil brain data at $R = 6$ with standard L1-wavelet Compressed Sensing (CS) and INRs with none, L1-wavelet, and our proposed diffusion regularization. The proposed INFusion approach yields images with lowest NRMSE.

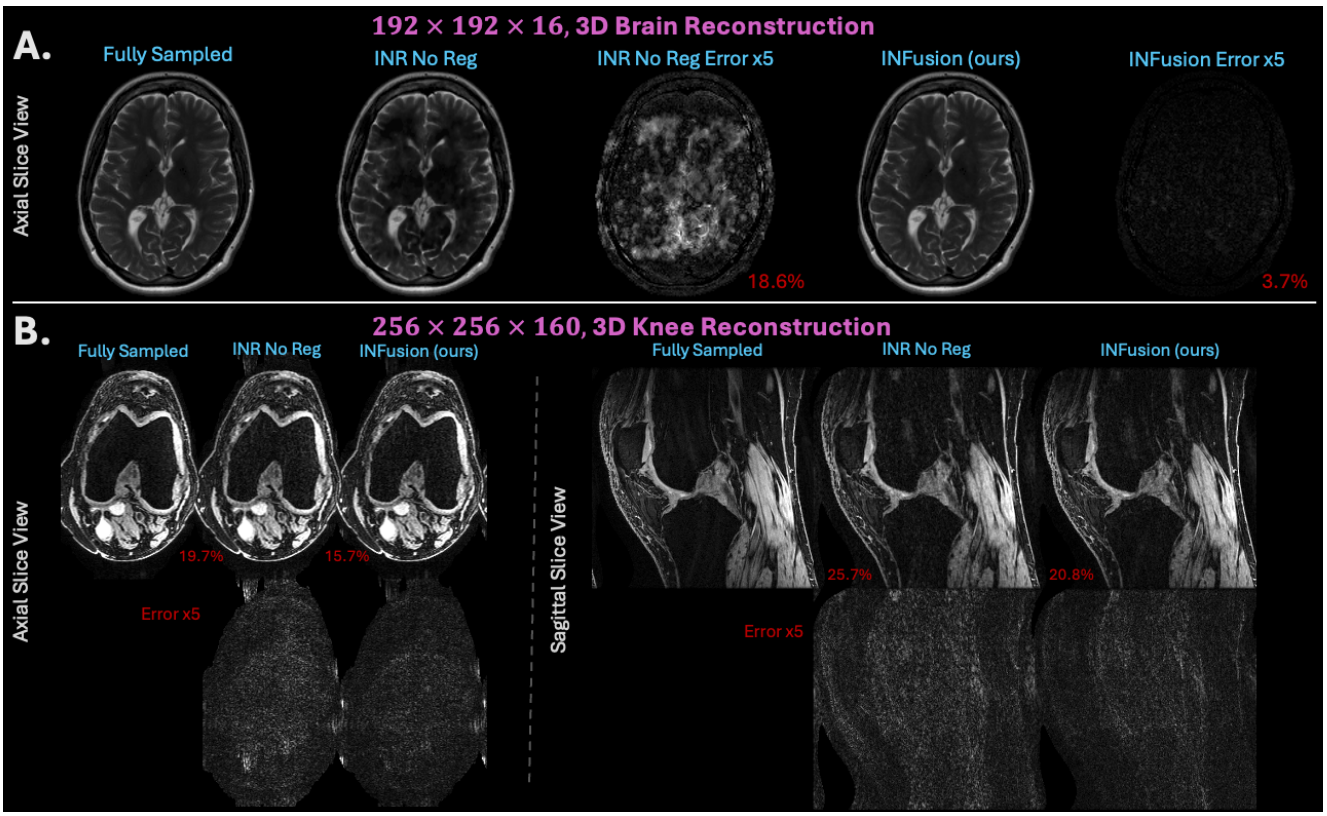

Fig. 4. (A.) INFusion outperforms no regularization on the modest, single-coil 3D dataset, but encoding all three spatial dimensions in the INR required prohibitively large GPU memory usage and compute. (B.) Applying diffusion regularization to random slices at each iteration and encoding only the x-y coordinates enables training of INRs with diffusion regularization on a realistically sized $256 \times 256 \times 80$ 3D k-space volume.

REFERENCES

[1] Griswold MA, Jakob PM, Heidemann RM, Nittka M, Jellus V, Wang J, et al. Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magn Reson Med. 2002;47: 1202–1210.

[2] Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: sensitivity encoding for fast MRI. Magn Reson Med. 1999;42: 952–962.

[3] Uecker M, Lai P, Murphy MJ, Virtue P, Elad M, Pauly JM, et al. ESPIRiT—an eigenvalue approach to autocalibrating parallel MRI: where SENSE meets GRAPPA. Magn Reson Med. 2014;71: 990–1001.

[4] Lustig M, Donoho D, Pauly JM. Sparse MRI: The application of compressed sensing for rapid MR imaging. Magn Reson Med. 2007;58: 1182–1195.

[5] Aggarwal HK, Mani MP, Jacob M. MoDL: Model-Based Deep Learning Architecture for Inverse Problems. IEEE Trans Med Imaging. 2019;38: 394–405.

[6] Hammernik K, Klatzer T, Kobler E, Recht MP, Sodickson DK, Pock T, et al. Learning a variational network for reconstruction of accelerated MRI data. Magn Reson Med. 2018;79: 3055–3071.

[7] Yaman B, Hosseini SAH, Moeller S, Ellermann J, Uğurbil K, Akçakaya M. Self-supervised learning of physics-guided reconstruction neural networks without fully sampled reference data. Magn Reson Med. 2020;84: 3172–3191.

[8] A. Jalal, M. Arvinte, G. Daras, E. Price, A. G. Dimakis, and J. Tamir, "Robust compressed sensing mri with deep generative priors," Advances in Neural Information Processing Systems, vol. 34, pp. 14 938–14 954, 2021.

[9] Jyothi Chand MJ. Multi-Scale Energy (MuSE) plug and play framework for inverse problems. arXiv. 2023.

[10] Luo G, Blumenthal M, Heide M, Uecker M. Bayesian MRI reconstruction with joint uncertainty estimation using diffusion models. Magn Reson Med. 2023;90: 295–311.

[11] Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. Advances in Neural Information Processing Systems, 33:7462–7473, 2020.

[12] Matthew Tancik, Pratul P. Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan T. Barron, and Ren Ng. Fourier features let networks learn high frequency functions in low dimensional domains. Advances in Neural Information Processing Systems, 33:7537–7547, 2020

[13] Feng R, Wu Q, Feng J, She H, Liu C, Zhang Y, et al. IMJENSE: Scan-specific Implicit Representation for Joint Coil Sensitivity and Image Estimation in Parallel MRI. IEEE Trans Med Imaging. 2023

[14] Shen L, Pauly J, Xing L. NeRP: Implicit Neural Representation Learning With Prior Embedding for Sparsely Sampled Image Reconstruction. IEEE Trans Neural Netw Learn Syst. 2022

[15] Huang W, Li HB, Pan J, Cruz G, Rueckert D, Hammernik K. Neural Implicit k-Space for Binning-Free Non-Cartesian Cardiac MR Imaging. Information Processing in Medical Imaging. Springer Nature Switzerland; 2023. pp. 548–560.

[16] J. F. Kunz, S. Ruschke, and R. Heckel, Implicit neural networks with Fourier-feature inputs for free-breathing cardiac MRI reconstruction, 2023, arXiv:2305.06822.

[17] Jang TJ, Hyun CM., NeRF Solves Undersampled MRI Reconstruction, 2024, arXiv:2402.13226

[18] Wu R, Mildenhall B, Henzler P, Park K, Gao R, Watson D, et al. ReconFusion: 3D Reconstruction with Diffusion Priors, 2023, arXiv:2312.02981

[19] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. 2020. Score-Based Generative Modeling through Stochastic Differential Equations. In International Conference on Learning Representations.

[20] Karras T, Aittala M, Aila T, Laine S. Elucidating the design space of diffusion-based generative models. Adv Neural Inf Process Syst. 2022;35: 26565–26577.

[21] Zhang R, Isola P, Efros AA, Shechtman E, Wang O. The unreasonable effectiveness of deep features as a perceptual metric. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018; 586–595.

[22] Knoll F, Zbontar J, Sriram A, Muckley MJ, Bruno M, Defazio A, 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. 2020;2: e190007.

[23] Tamir JI, Ong F, Cheng JY, Uecker M, Lustig M. Generalized magnetic resonance image reconstruction using the Berkeley Advanced Reconstruction Toolbox. ISMRM Workshop on Data Sampling and Image Reconstruction, Sedona, AZ. 2016.

[24] Ye JC. Compressed sensing MRI: a review from signal processing perspective. BMC Biomed Eng. 2019;1: 8.

[25] A.D. Desai, A.M. Schmidt, E.B. Rubin, C.M. Sandino, M.S. Black, V. Mazzoli, K.J. Stevens, R. Boutin, C. Re, G.E. Gold, B. Hargreaves, A. Chaudhari, SKM-TEA: a dataset for accelerated MRI reconstruction with dense image labels for quantitative clinical evaluation, in: Thirty-Fifth Conference on Neural Information Processing Systems Datasets and Benchmarks, 2021, Track (Round 2)