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) is a clinically important medical imaging modality that provides excellent soft tissue contrast without any ionizing radiation. However, MRI acquisitions suffer from an inherent trade-off between scan time, resolution, and signal-to-noise ratio (SNR) [1]. Shortening MRI scan time while maintaining image quality through computation reduces costs to patients and hospitals and reduces sensitivity to motion artifacts.

MRI scans can be accelerated by sampling the acquired Fourier (k-space) data below the Nyquist sampling rate and solving the ill-posed reconstruction problem to produce a high fidelity image from under-sampled measurements. To reduce scan time, clinics routinely use parallel imaging, [2], [3], [4], which interpolates missing data with the encoding capability of the multi-coil signal receive array, and compressed sensing [5], which exploits random sampling and sparsity. Recently, machine learning algorithms for accelerated MRI reconstruction employing training databases yield state of the art results [6], [7], [8], [9], [10], [11], [12]. End-to-end methods learn a point-wise mapping between under-sampled and fully-sampled data but are susceptible to test time shifts in the measurement operator. More recently, generative methods have been proposed that learn a prior over clean images and are robust to test time shifts in the forward operator.

All the MRI reconstruction algorithms previously discussed assume that a discretized grid of voxels represents the desired MRI image. However, alternative representations can also be employed, such as Implicit Neural Representations (INRs) [13], [14]. An INR trains a neural network to take as input a spatial coordinate and output the signal value of interest at that spatial coordinate, thus providing a continuous image representation, and in some sense, implicitly storing the image in the weights of the neural network. Recent work utilizes INRs for accelerating MRI, particularly in scan-specific settings where only data from the under-sampled scan itself are available. By combining traditional regularization with the inherent regularization of the dense neural network architecture, positional encoding, and periodic activations, previous work demonstrates that INRs help further accelerate MRI scans [15], [16], [17], [18], [19]. However, these methods apply unlearned spatial regularization priors and are mainly limited to 2D or low-resolution 3D acquisitions.

This work proposes further accelerating MRI with INR-based reconstructions by drawing inspiration from recent work in computer vision that uses diffusion priors to reconstruct INR-representations of 3D scenes from few camera views [20]. In particular, our method, termed INFusion, uses diffusion models pre-trained on MRI images to regularize the optimization of INRs from under-sampled measurements. First, we demonstrate that our proposed approach improves reconstruction performance in comparison to competing techniques in 2D, T_2-weighted brain MRI. Then, we propose a hybrid 3D approach with our diffusion regularization that enables INR application on large-scale 3D MR datasets with 2D diffusion models, and we show experiments on a realistic, 3D knee acquisition.

II. METHODS

A. The Accelerated MRI Inverse Problem

Given the linear, multi-channel MRI measurement model $A = PFS \in \mathbb{C}^{(MC)\times N}$, solving the following inverse problem yields an image from under-sampled measurements, $y \in \mathbb{C}^{MC}$, acquired in an accelerated scan,

$$\arg\min_x ||y - Ax||_2^2 + \lambda R(x), \qquad (1)$$

where $x \in \mathbb{C}^N$ is the discretized image to reconstruct, $S \in \mathbb{C}^{(NC)\times N}$ describes how the multi-coil receive array measures signal, $F \in \mathbb{C}^{(NC)\times(NC)}$ applies the Fourier Transform to each coil image, $P \in \mathbb{C}^{(MC)\times(NC)}$ indicates which k-space points are sampled (with $M < N$ for reduced scan time), and $\lambda R(\cdot) : \mathbb{C}^N \to \mathbb{R}$ imposes regularization (e.g., sparsity, low-rank).

B. Implicit Neural Representations

Let $r \in \mathbb{R}^3$ represent a physical coordinate corresponding to a spatial location of the MRI anatomy to image (without loss of generality assume that $r \in [0,1]^3$). Then, an Implicit Neural Representation (INR) is a fully connected neural network $I_\theta : \mathbb{R}^3 \to \mathbb{C}$ that takes a spatial coordinate, r, as input and outputs the signal value of interest of the MRI anatomy at that spatial location. Given a discretized image x, the INR can be trained by optimizing

$$\theta^* = \arg\min_\theta ||x - I_\theta(\boldsymbol{r})||_2^2, \qquad (2)$$

where $\boldsymbol{r} \in \mathbb{R}^{3\times N}$ is the spatial coordinates of the voxel locations in the discretized image. Then, I_{θ^*} can be evaluated at spatial coordinates not on the original grid of coordinates, $\boldsymbol{r}$, thus yielding a continuous representation.

C. Accelerated MRI with INRs

Given the discrete, under-sampled k-space measurements y, the following optimization problem estimates weights of an INR in the accelerated MRI setting:

$$\theta^* = \arg\min_\theta ||y - AI_\theta(\boldsymbol{r})||_2^2 + \lambda R(I_\theta(\boldsymbol{r})), \qquad (3)$$

recalling that $\boldsymbol{r}$ is the spatial coordinates of the discretized voxel locations. Then evaluating $I_{\theta^*}(\boldsymbol{r})$ yields the desired image. Previous work shows that the structure of the INR helps regularize the ill-posed reconstruction problem and yields higher fidelity images in comparison to directly solving for the discretized image [15], [16], [17], [18], [19]. However, previous work regularizes the optimization problem with a hand-crafted spatial prior R, leaving room for further improvements with learned regularizors.

D. 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 [9], [10], [11] enabling applications with INRs. As it is computationally expensive to evaluate the diffusion model at every 3D coordinate, we also propose a stochastic
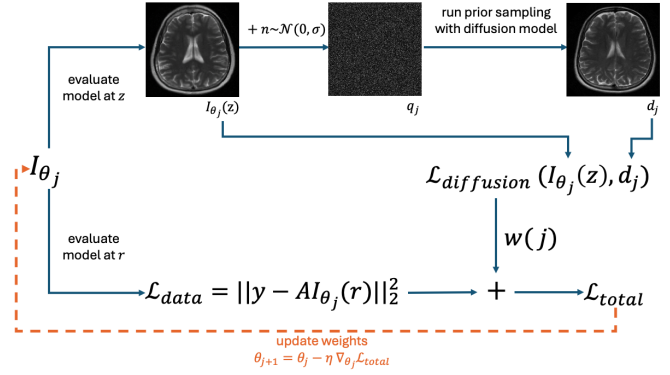

Procedure on iteration j of optimization

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.

regularizer 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 [21], [22] 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 [23].
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. In 3D settings, evaluating the INR at all spatial coordinates and backpropagating through the prior sampling procedure becomes computationally prohibitive. For 3D problems, we propose only encoding the $x - y$ dimension with spatial coordinates and having the model return discretized values corresponding to the third z dimension. In addition, we compute our proposed diffusion regularization with respect to two random slices at each iteration.

E. INR and INFusion Hyperparameters

For all experiments, our INRs consisted of fully-connected neural networks with relu activations and 4 hidden layers using spatial coordinates encoded with 128 gaussian fourier features [14]. INRs for 2D and 3D experiments had size 256 and size 3200 hidden layers respectively.

For the INFusion procedure, we selected $\sigma_{\min}$ to decrease from [1,0] linearly as function of the current optimization

iteration j. Similarly, $w(j)$ decreased exponentially from [1,0]. We ran 10 iterations of the diffusion sampling formulation in [22] to generate the prior sample d_j.

F. Generative Diffusion Model Training

We trained two diffusion models using a U-Net style architecture with 65 million parameters and used the "EDM" hyper-parameters, loss function, and diffusion noise schedule [22]. The first model was trained on 8500 T_2 axial brain slices from fastMRI [24] of dimension 192×192 and the second was trained on 15720 sagittal knee slices from SKM-TEA [25].

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 [24] dataset, cropped both to 192×192 resolution, and retrospectively under-sampled with a 2D-Poisson pattern generated with BART [26] 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 [27] 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. We also 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. Both 3D experiments compared discretized images produced by an INR trained with and without our INFusion method. The smaller 3D brain experiment encoded all three spatial dimensions with coordinates and computed diffusion regularization across all 16 slices each iteration. The larger knee experiment incorporated the previously discussed modifications for 3D computational feasibility. Note that our diffusion model was trained on $x - y$ images, but under-sampling was performed in the $y - z$ direction.

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
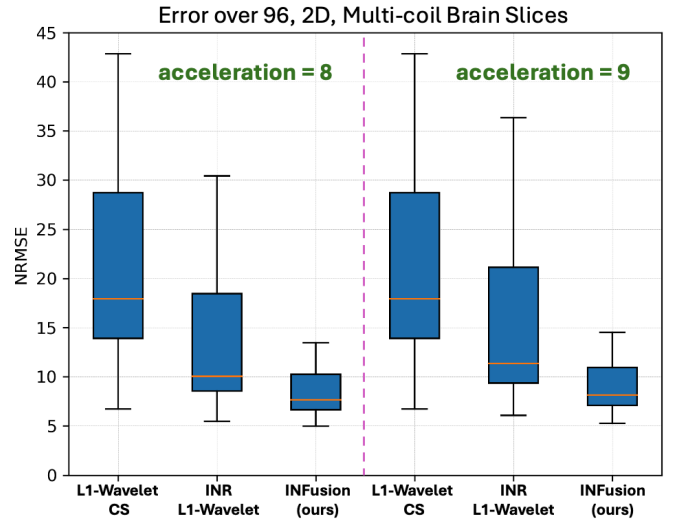

Fig. 2. Quantitative comparison of standard L1-Wavelet CS and INRs trained with Wavelet and our proposed regularization at $R = 8$ and $R = 9$ across 96 samples. Our proposed INFusion approach improves NRMSE across the test dataset.

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.

V. DISCUSSION

INFusion exploits the generative capabilities of diffusion models to regularize INR training from scan-specific, under-sampled MR measurements. INR-based accelerated MRI inverse problems solved with our proposed diffusion regularization yield higher fidelity images with reduced scan time. 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. Full 3D diffusion models are challenging to train, due to computational burden and limited data availability, but INFusion enables regularization of 3D reconstruction problems using 2D diffusion models. In addition, the 2D diffusion models can be trained on data from any 2D orientation that is available.

Our implementation takes approximately 15 minutes to reconstruct a single 2D slice on a Nvidia A100 GPU, so clinical adoption requires a more efficient implementation. Future research work will investigate the effectiveness of INRs as an image representation and explore the trade-offs between using diffusion models as regularization versus posterior sampling directly on the voxel grid.
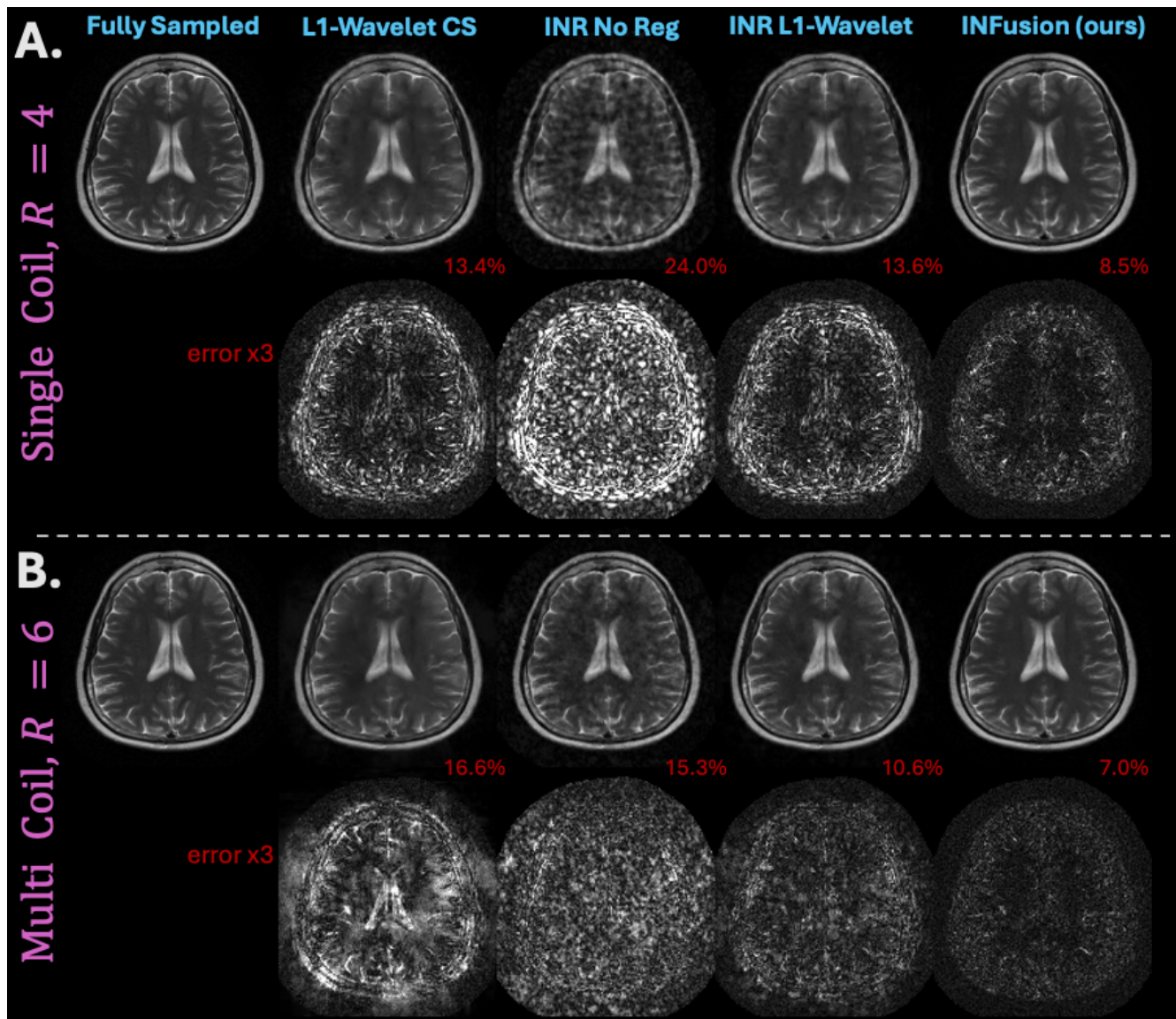
Fig. 3. Reconstructions of discrete images on the (A.) 2D single-coil brain data at $R = 4$ and (B.) 2D multi-coil brain data at $R = 6$ with standard L1-wavelet Compressed Sensing (CS) and INRs with none, L1-wavelet, and our proposed diffusion regularization. The proposed INFusion approach yields images with lowest NRMSE and best qualitative image quality.

REFERENCES

[1] Dwight Nishimura, *Principles of Magnetic Resonance Imaging*, Stanford University, 2010.

[2] Mark Griswold, Peter Jakob, Robin Heidemann, Mathias Nittka, Vladimir Jellus, Jianmin Wang, Berthold Kiefer, and Axel Haase, "Generalized autocalibrating partially parallel acquisitions (grappa)," *Magnetic Resonance in Medicine*, 2002.

[3] klaas pruessmann, markus weiger, markus scheidegger, and peter boesiger, "sense: sensitivity encoding for fast mri," *Magnetic Resonance in Medicine*, 1999.

[4] Martin Uecker, Peng Lai, Mark J. Murphy, Patrick Virtue, Michael Elad, John M. Pauly, Shreyas S. Vasanawala, and Michael Lustig, "Espirit—an eigenvalue approach to autocalibrating parallel mri: Where sense meets grappa," *Magnetic Resonance in Medicine*, 2013.

[5] Michael Lustig, David Donoho, and John M. Pauly, "Sparse mri: The application of compressed sensing for rapid mr imaging," 2007.

[6] Hemant K Aggarwal, Merry P Mani, and Mathews Jacob, "Modl: Model based deep learning architecture for inverse problems," *IEEE Transactions on Medical Imaging*, 2018.

[7] Kerstin Hammernik, Teresa Klatzer, Erich Kobler, Michael P. Recht, Daniel K. Sodickson, Thomas Pock, and Florian Knoll, "Learning a variational network for reconstruction of accelerated mri data," *Magnetic Resonance in Medicine*, 2017.

[8] Burhaneddin Yaman, Seyed Amir Hossein Hosseini, Steen Moeller, Jutta Ellermann, Kâmil Uğurbil, and Mehmet Akçakaya, "Self-supervised learning of physics-guided reconstruction neural networks without fully sampled reference data," *Magnetic Resonance in Medicine*, 2020.

[9] Ajil Jalal, Marius Arvinte, Giannis Daras, Eric Price, Alexandros G Dimakis, and Jon Tamir, "Robust compressed sensing mri with deep generative priors," 2021.

[10] Jyothi Rikhab Chand and Mathews Jacob, "Multi-scale energy (muse) plug and play framework for inverse problems," *arXiv*, 2023.

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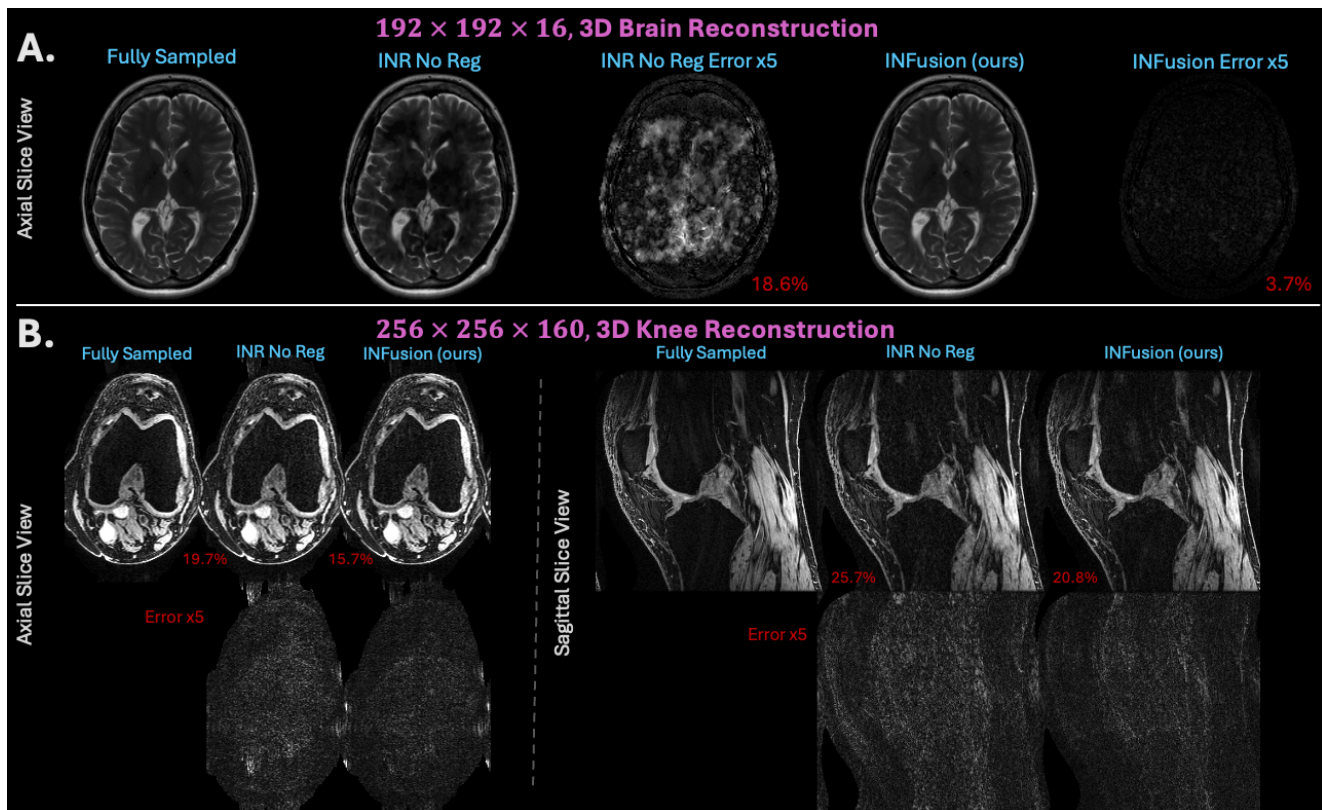
Fig. 4. (A.) INFusion outperforms no regularization on the modest, single-coil 3D dataset, but encoding all three spatial dimensions in the INR required prohibitively large GPU memory usage and compute. (B.) Applying diffusion regularization to random slices at each iteration and encoding only the x-y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

"Bayesian mri reconstruction with joint uncertainty estimation using diffusion models," *Magnetic Resonance in Medicine*, 2023.

[12] Reinhard Heckel, Mathews Jacob, Akshay Chaudhari, Or Perlman, and Efrat Shimron, "Deep learning for accelerated and robust mri reconstruction," *Magnetic Resonance Materials in Physics, Biology and Medicine*, 2024.

[13] Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein, "Implicit neural representations with periodic activation functions," *Advances in Neural Information Processing Systems*, 2020.

[14] 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*, 2020.

[15] Ruimin Feng, Qing Wu, Jie Feng, Huajun She, Chunlei Liu, Yuyao Zhang, and Hongjiang Wei, "Imjense: Scan-specific implicit representation for joint coil sensitivity and image estimation in parallel mri," *IEEE Transactions on Medical Imaging*, 2024.

[16] Liyue Shen, John Pauly, and Lei Xing, "Nerp: Implicit neural representation learning with prior embedding for sparsely sampled image reconstruction," *IEEE Transactions on Neural Networks and Learning systems*, 2022.

[17] Wenqi Huang, Hongwei Bran Li, Jiazhen Pan, Gastao Cruz, Daniel Rueckert, and Kerstin Hammernik, "Neural implicit k-space for binning-free non-cartesian cardiac mr imaging," *Information Processing in Medical Imaging*, 2023.

[18] Johannes F. Kunz, Stefan Ruschke, and Reinhard Heckel, "Fourier-feature inputs for free-breathing cardiac mri reconstruction," *IEEE Transactions on Computational Imaging*, 2024.

[19] Tae Jun Jang and Chang Min Hyun, "Nerf solves undersampled mri reconstruction," *arXiv*, 2024.

[20] Rundi Wu, Ben Mildenhall, Philipp Henzler, Keunhong Park, Ruiqi Gao, Daniel Watson, Pratul P. Srinivasan, Dor Verbin, Jonathan T. Barron, Ben Poole, and Aleksander Holynski, "Reconfusion: 3d reconstruction with diffusion priors," *Conference on Computer Vision and Pattern Recognition*, 2024.

[21] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole, "Score-based generative modeling through stochastic differential equations," *International Conference on Learning Representations*, 2020.

[22] Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine, "Elucidating the design space of diffusion-based generative models," *Advances in Neural Information Processing Systems*, 2022.

[23] Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang, "The unreasonable effectiveness of deep features as a perceptual metric," *Conference on Computer Vision and Pattern Recognition*, 2018.

[24] Florian Knoll, Jure Zbontar, and et. al., "fastmri: A publicly available raw k-space and dicom dataset of knee images for accelerated mr image reconstruction using machine learning," *Radiology: Artificial Intelligence*, 2020.

[25] Arjun D Desai, Andrew M Schmidt, Elka B Rubin, Christopher M Sandino, Marianne S Black, Valentina Mazzoli, Kathryn J Stevens, Robert Boutin, Christopher Ré, Garry E Gold, Brian A Hargreaves, and Akshay S Chaudhari, "Skm-tea: A dataset for accelerated mri reconstruction with dense image labels for quantitative clinical evaluation," *Advances in Neural Information Processing Systems*, 2021.

[26] Jonathan Tamir, Joseph Cheng, Martin Uecker, and Michael Lustig, "Generalized magnetic resonance image reconstruction using the berkeley advanced reconstruction toolbox," *ISMRM Workshop on Data Sampling and Image Reconstruction*, 2016.

[27] Jong Chul Ye, "Compressed sensing mri: a review from signal processing perspective," *BMC Biomedical Engineering*, 2019.