Coarse-to-Fine 3D MRI Reconstruction via Resolution-Agnostic Neural Operators

Jiayun Wang[1*] **Ram Daftari**[2,3*] **Weijie Gan**[4] **Valentin Duruisseaux**[2]
Ruibo Wang[5] **Armeet Singh Jatyani**[2] **Boris Bonev**[6] **Anima Anandkumar**[2]

[1]**Georgia Tech** [2]**Caltech** [3]**IIT Kanpur** [4]**Washington University in St. Louis** [5]**TU Delft** [6]**NVIDIA**

pjwang@gatech.edu, ramdd24@iitk.ac.in, anima@caltech.edu

Abstract

Deep-learning-based 3D MRI reconstruction is limited in practice by the GPU memory and compute cost of full-volume training; 3D acquisitions are thus most often reconstructed by 2D models applied slice-by-slice and then stacked, which ignores through-plane (stacking-axis) structure. We reconstruct in full 3D and ask whether a model can be trained at low resolution (fast and memory-feasible) yet deployed at higher resolution. This coarse-to-fine strategy needs reliable cross-resolution performance, but standard networks are resolution-specific and degrade substantially when evaluated outside the training setting. We therefore use resolution-agnostic neural operators (NOs) and propose *KIR-NO* (k-space-to-Image Reconstruction Neural Operator), a discretization-agnostic framework for accelerated 3D MRI reconstruction. KIR-NO is trained with full backpropagation on low-resolution volumes and applied zero-shot to higher-resolution volumes at inference. Its core building block is a 3D neural operator based on discrete-continuous convolutions: convolutional filters are parameterized in a continuous basis and sampled at arbitrary voxel resolutions while preserving the local inductive bias of convolution. KIR-NO integrates these learned operators with physics-based consistency: a k-space neural-operator head first refines the undersampled Fourier measurements, and image-space neural-operator cascades then refine the reconstruction through interleaved data-consistency updates. Across 3D knee and brain MRI benchmarks, KIR-NO improves over classical and CNN baselines: it improves over a state-of-the-art 3D CNN by 2.2 dB PSNR at half resolution under $4\times$ acceleration on SKM-TEA, with gains persisting at $8\times$ and $16\times$. Together, these results show that resolution-agnostic operators make coarse-to-fine 3D MRI reconstruction – cheap low-resolution training with high-resolution deployment – practical across anatomies. Our code is available at `https://github.com/neuraloperator/KIR-NO`.

1 Introduction

Deep learning has emerged as a powerful tool for medical imaging, e.g. MRI reconstruction from highly undersampled measurements [1–6], where it can shorten scans and improve patient comfort. Although most existing work focuses on 2D reconstruction due to GPU memory and compute constraints [7, 8], there is growing interest in 3D MRI reconstruction. Compared with slice-wise 2D approaches, 3D models leverage volumetric context to improve through-plane consistency and reconstruction fidelity [7], which is especially important when lesions span multiple slices and 3D morphology aids diagnosis [9]. It is enabled by public volumetric datasets (e.g., SKM-TEA [10]) and by hardware that increasingly makes whole-volume processing feasible.

*Equal contribution.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

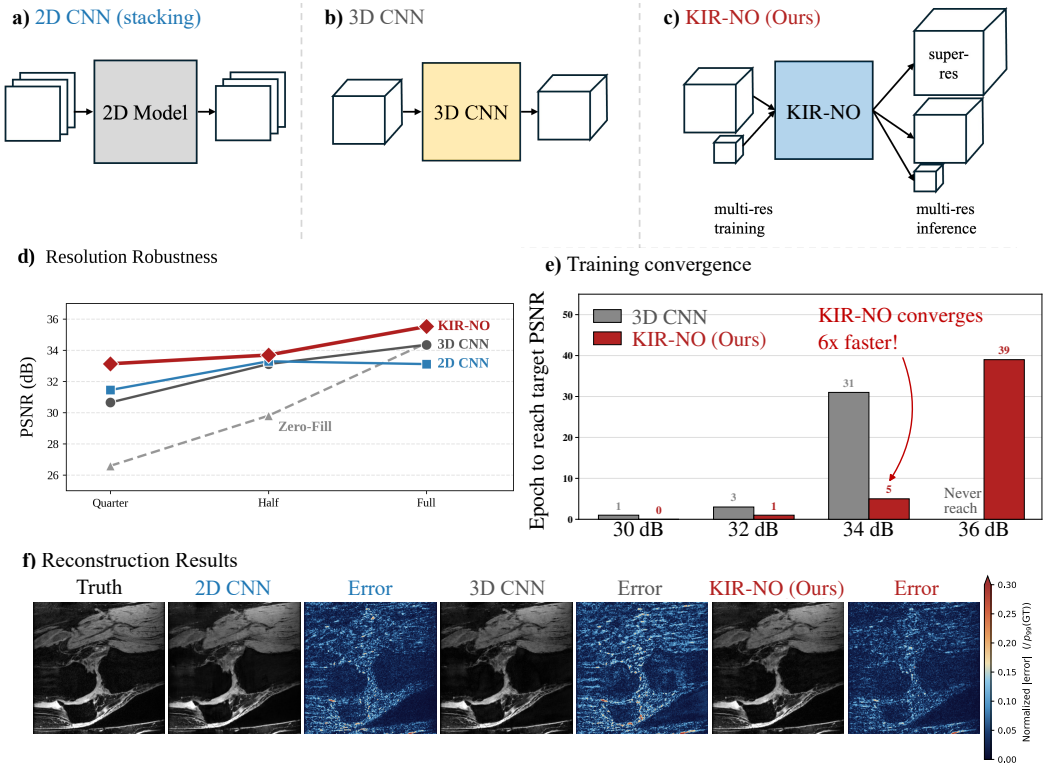

Figure 1: (a–c) Three deep learning architectures for 3D MRI reconstruction. **(a) 2D CNN:** GPU-memory limits force slice-wise 2D models that stack into a volume, losing inter-slice consistency. **(b) 3D CNN:** fixed discrete kernels are tied to the training grid, so cross-resolution transfer is unreliable. **(c) 3D Neural Operator (KIR-NO, ours):** discretization-agnostic — trained on memory-friendly coarse volumes and deployed at higher resolutions without retraining. **(d) Resolution robustness:** PSNR at quarter/half/full resolution on SKM-TEA (quarter-trained); KIR-NO leads throughout. **(e) Convergence** (SKM-TEA, half-res, $4\times$): first epoch to reach each PSNR threshold; KIR-NO reaches 34 dB $\sim 6\times$ faster than the 3D CNN. **(f) Qualitative:** a knee slice with error maps for 2D CNN, 3D CNN, and KIR-NO; KIR-NO has the lowest residual.

The computational gap between 2D and 3D is significant at training time. While high-resolution 3D inference only holds forward activations, training additionally requires storing activations across all layers for backpropagation and computing gradients across an additional spatial dimension, making naive full-resolution training prohibitively expensive. This motivates a *coarse-to-fine* paradigm: *train at low resolution, where backpropagation remains feasible, and deploy zero-shot at full resolution.* The memory saving here is a property of the training *recipe*, thus the decisive question is not model size but whether the learned model transfers reliably across resolutions. Standard CNNs, however, overfit to the discretizations seen during training [8]: resizing or interpolating discrete kernels alters the effective receptive field and provides no principled mechanism for resolution transfer. In MRI, such unprincipled transfer can introduce artifacts or hallucinated structures in critical anatomy and may mislead clinical interpretation.

Neural operators provide a principled alternative by learning function-to-function maps rather than mappings in fixed dimensions [11, 12], naturally supporting resolution-agnostic reconstruction. In particular, local neural operators built on discrete-continuous convolutions (DISCO) [13, 14] serve as a resolution-agnostic convolution layer, with strong empirical results on 2D MRI [5], CT [15] and 3D photoacoustic tomography [16], suggesting that they adapt more naturally than standard convolutions for varying imaging resolution and undersampling.

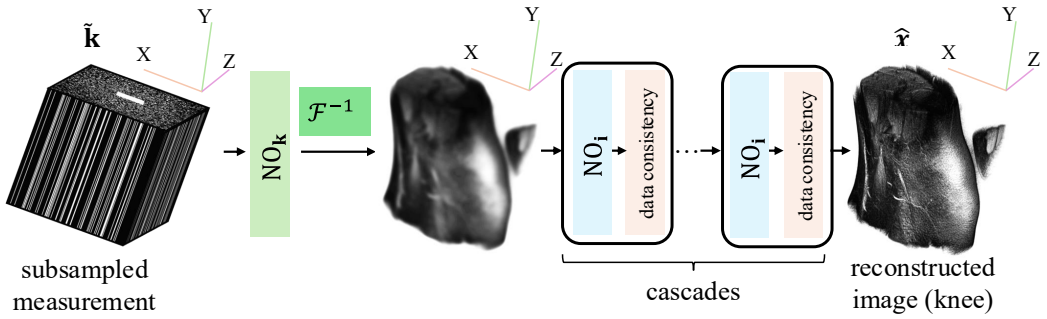

Figure 2: Our **KIR-NO** architecture. Undersampled multi-coil k-space measurements $\tilde{\mathbf{k}}$ are first refined by a k-space neural-operator head NO_k (Eq. 1), then mapped to image space (inverse Fourier transform $\mathcal{F}^{-1}$ and coil combination) and refined by an unrolled cascade of image-space neural operators NO_i (Eq. 2, where ℓ indexes the cascade) interleaved with data-consistency updates, producing the final 3D reconstruction $\hat{\mathbf{x}}$. Operator layers learn continuous kernels built on discrete-continuous convolutions, in contrast to fixed-resolution conventional CNN kernels.

Our approach. Our central novelty is making resolution-agnostic reconstruction work *natively in 3D*: while the prior 2D neural-operator/DISCO line targets robustness to undersampling patterns and field-of-view [5], we target robustness to *resolution*, a coarse-to-fine capability specific to 3D. Extending 2D DISCO to 3D is non-trivial: the 2D kernel bases [13, 14] are defined on planar supports and must be redesigned for volumetric kernels that stay compactly supported, so that each layer remains local and memory-feasible on dense 3D grids. We therefore design two new 3D continuous kernel bases — piecewise-linear hat functions in spherical coordinates and 3D Morlet wavelets — that are strictly local, orientation-selective, and GPU-friendly. Using these bases, we build *KIR-NO*, a k-space 3D DISCO head refines the undersampled measurements, then (after an inverse FFT) an unrolled cascade of image-space 3D DISCO operators alternates with k-space data-consistency updates. Sampling each kernel at the target voxel resolution makes KIR-NO a resolution-agnostic alternative to `Conv3D` that preserves convolutional locality and weight sharing, leading to more reliable high-resolution transfer than 3D CNNs.

Empirically across knee and brain MRI, KIR-NO outperforms classical and CNN baselines, and diffusion baselines given a single sampling step. On SKM-TEA it improves over a 3D E2E-VarNet by $+2.2\,\mathrm{dB}$ at the in-domain half resolution ($4\times$) and, most notably, by $+3\,\mathrm{dB}$ (0.940 vs 0.870 SSIM) under zero-shot super-resolution to full resolution, with gains persisting at $8\times$ and $16\times$.

Contributions. We (i) introduce a coarse-to-fine paradigm for accelerated 3D MRI enabled by resolution-agnostic operators: train at low resolution, deploy zero-shot at high resolution; (ii) design two new 3D DISCO kernel bases (piecewise-linear-spherical and 3D Morlet); and (iii) build *KIR-NO*, a dual-domain k-space/image neural-operator reconstructor. The approach could extend to other volumetric inverse problems (e.g., seismic imaging, 3D microscopy).

2 Related Work

Accelerated MRI. Medical imaging tasks are often inverse problems: recovering images from indirect, noisy measurements under a known forward model [17]. In MRI, this model samples the image Fourier transform at discrete k-space frequencies; accelerated MRI undersamples these frequencies [2], making reconstruction ill-posed. This is especially challenging in 3D, where $O(n^3)$ voxels sharply increase memory and compute requirements [7, 10]. MRI reconstruction is often cast as regularized optimization alternating between *data consistency* and prior-driven regularization, using classical sparsity priors [18] or learned plug-and-play denoisers [19, 20].

Deep Learning for MRI Reconstruction. Early deep methods learn a feed-forward CNN that maps the zero-filled reconstruction to a clean image [21, 22]. *Unrolled* networks instead preserve the optimization structure by replacing the hand-crafted prior with a trainable module and unrolling iterations into *cascades*, trained end-to-end [3, 4]. Our main CNN baseline, E2E-VarNet [4], alternates multi-coil k-space data fidelity with U-Net image refinement and provided sensitivity maps. In 3D,

storing per-cascade activations scales as $O(n^3)$, making full-volume training memory-prohibitive at clinical resolutions; many models therefore train on 2D slices and apply them slice-by-slice, sacrificing through-plane consistency. Wang et al. [7] address this with memory-efficient unrolled 3D training, but remain tied to a fixed training discretization, motivating our resolution-agnostic 3D framework. Score-based diffusion priors [6] decouple the prior from the forward operator but require many sampling steps and are substantially slower at inference.

Neural Operators. Unlike standard neural networks, neural operators learn function-to-function maps [11, 12], are universal approximators for nonlinear operators [23], accept inputs and can be queried at arbitrary discretizations, enabling training across mixed resolutions. Physics-informed variants [24–26] add residual losses from governing equations; in contrast, KIR-NO incorporates MRI acquisition physics through data-consistency and k-space processing rather than an additional loss. The Fourier neural operator (FNO) [27, 28] truncates high-frequency modes and is equivalent to a global convolution; FNOSeg3D [8] adapts it for 3D MRI segmentation. Reconstruction, however, requires capturing intricate anatomical structure with local inductive biases, where FNO's high-frequency truncation is detrimental. DISCO [13, 14] retains all frequencies and learns local biases, and Jatyani et al. [5] show DISCO-based operators are robust for 2D MRI across undersampling patterns and extended field-of-view. That line, however, is confined to 2D and targets robustness to the *sampling* operator; we instead target cross-*resolution* transfer in *3D* (coarse-to-fine, train-low/deploy-high), which raises new continuous-basis-design and volumetric-memory challenges absent in 2D.

3 Methods

KIR-NO reconstructs 3D MRI volumes from undersampled multi-coil k-space by combining physics-based data consistency with neural-operator priors. It comprises a forward operator $\mathcal{A}$, an unrolled architecture of data-consistency and image-space neural-operator updates, and an optional k-space neural-operator head. To transfer across discretizations while preserving local structure, the neural operators are parameterized as local integration operators via 3D DISCO convolutions.

We denote the MRI forward imaging operator by $\mathcal{A}$ and model the acquired subsampled k-space measurements as $\tilde{\mathbf{k}} = \mathcal{A}(\mathbf{x}_\star)+\epsilon$, where $\mathbf{x}_\star$ is the unknown ground-truth image and ϵ represents measurement noise and modeling error. Here, $\mathcal{A}$ consists of the multi-coil sensitivity encoding, followed by a Fourier transform and the application of an undersampling mask. Classical compressed sensing MRI reconstructs an image by solving a regularized inverse problem $\hat{\mathbf{x}} = \operatorname{argmin}_{\mathbf{x}} f(\mathbf{x}) + g(\mathbf{x})$, where the data-fidelity term $f(\mathbf{x}) = \frac{1}{2}\left\|\mathcal{A}(\mathbf{x}) - \tilde{\mathbf{k}}\right\|_2^2$ enforces agreement with the acquired measurements, while $g(\mathbf{x}) = \lambda\Psi(\mathbf{x})$ encodes prior knowledge about plausible images, such as sparsity [18].

3.1 Unrolled Neural-Operator Reconstruction

Motivated by this optimization perspective, KIR-NO (Fig. 2) uses an unrolled neural operator: instead of a hand-crafted regularizer Ψ, it learns the prior from data while keeping explicit dependence on the acquisition model through data-consistency updates.

k-space Neural Operator. We first augment KIR-NO with a k-space neural-operator (kNO) head that operates directly on the coil-combined undersampled measurements before any inverse Fourier transform. Given the acquired subsampled k-space measurements $\tilde{\mathbf{k}}$, we compute

$$\tilde{\mathbf{k}}' = \tilde{\mathbf{k}} + \mathrm{NO}_k(\tilde{\mathbf{k}}), \tag{1}$$

where NO_k denotes a neural operator acting on k-space samples. The refined multi-coil measurements $\tilde{\mathbf{k}}'$ are then mapped to image space by the adjoint operator, $\mathbf{x}^0 = \mathcal{A}^*(\tilde{\mathbf{k}}')$ (an inverse Fourier transform followed by adjoint coil combination), providing the initial image estimate that is passed to the subsequent image-space cascades.

Operating directly in k-space lets the head exploit acquisition-specific structure (e.g., omitted phase-encoding lines or sampling gaps) and supply a learned Fourier-domain initialization of missing spectral content before image-space refinement. Empirically this helps: KIR-NO outperforms the image-space-only variant (*KIR-NO w/o kNO*) by up to $+6.6\,\mathrm{dB}$ at full resolution (Table 10).

Image-Space Cascades. After the kNO head, KIR-NO unrolls a sequence of image-space cascades, each representing one learned reconstruction update.

Starting from the initial image estimate $\mathbf{x}^0 = \mathcal{A}^*(\tilde{\mathbf{k}}')$ produced by the kNO head (or the adjoint $\mathcal{A}^*(\tilde{\mathbf{k}})$ when the kNO head is omitted), each cascade updates the current reconstruction according to

$$\mathbf{x}^{\ell+1} = \mathbf{x}^\ell - \eta^\ell \mathcal{A}^*\big(\mathcal{A}(\mathbf{x}^\ell) - \tilde{\mathbf{k}}\big) + \lambda^\ell \mathrm{NO}_i^\ell(\mathbf{x}^\ell), \tag{2}$$

where

1. $\mathcal{A}^*\big(\mathcal{A}(\mathbf{x}^\ell) - \tilde{\mathbf{k}}\big)$ is a data-fidelity term that backprojects the measurement residual into image space via the adjoint of $\mathcal{A}$;
2. $\mathrm{NO}_i^\ell(\mathbf{x}^\ell)$ is a learned neural-operator prior replacing the hand-crafted regularizer Ψ (removing aliasing, suppressing noise, recovering missing content);
3. η^ℓ, λ^ℓ weight the fidelity and prior terms, respectively.

The kNO head and image-space cascades are trained jointly end-to-end, forming a single reconstruction model rather than independent optimization iterations with fixed update rules. At inference, the trained KIR-NO maps k-space measurements $\tilde{\mathbf{k}}$ to a reconstructed 3D image $\hat{\mathbf{x}}$.

3.2 Neural Operator Parameterization

Both the image-space cascades and the kNO head use neural operators. Unlike standard networks, which learn maps between fixed-dimensional arrays tied to the training grid [5], neural operators learn maps between functions [11, 12] and can be queried at any discretization — exactly what our coarse-to-fine (train-low, deploy-high) setting requires.

Because MRI reconstruction also requires preserving local anatomical structure, we use local neural operators based on discrete-continuous convolutions (DISCO). This combines resolution-agnostic operator learning with convolutional locality, and serves as the building block for both the image-space cascades and the k-space neural-operator head.

3.2.1 Local Integration Operators

Local inductive bias is usually imposed with compactly supported convolutional kernels (as in CNNs), but standard discrete kernels are not resolution-agnostic: Liu-Schiaffini et al. [14] show they converge to pointwise linear operators as resolution increases, rather than to a local integration operator.

For a kernel κ and input function g defined over a compact domain $D \subset \mathbb{R}^d$, the *local convolution operator* maps g to an output at location v. Given input points $(u_j)_{j=1}^m \subset D$ with quadrature weights q_j and output points $v_i \in D$, DISCO [13, 14] approximates this operator by

$$(\kappa \star g)(v_i) = \int_D \kappa(u - v_i) \cdot g(u)\,\mathrm{d}u \approx \sum_{j=1}^m \kappa(u_j - v_i) \cdot g(u_j) q_j. \tag{3}$$

We parameterize the kernel as a linear combination of basis functions $\kappa(\mathbf{x}) = \sum_{\ell=1}^L \theta^\ell \kappa^\ell(\mathbf{x})$ where the coefficients θ^ℓ are learnable parameters and the basis functions κ^ℓ are defined in continuous space. The kernel is therefore parameterized independently of the grid on which it is evaluated. For a given resolution, we sample the continuous kernel on the corresponding 3D grid to obtain a discrete convolution stencil that can be applied with standard GPU-accelerated convolution routines. Across resolutions, only this sampled stencil changes; the learned coefficients θ^ℓ remain the same. This provides a resolution-agnostic local integration operator, which we use as the building block for both the measurement-space and image-space components of KIR-NO.

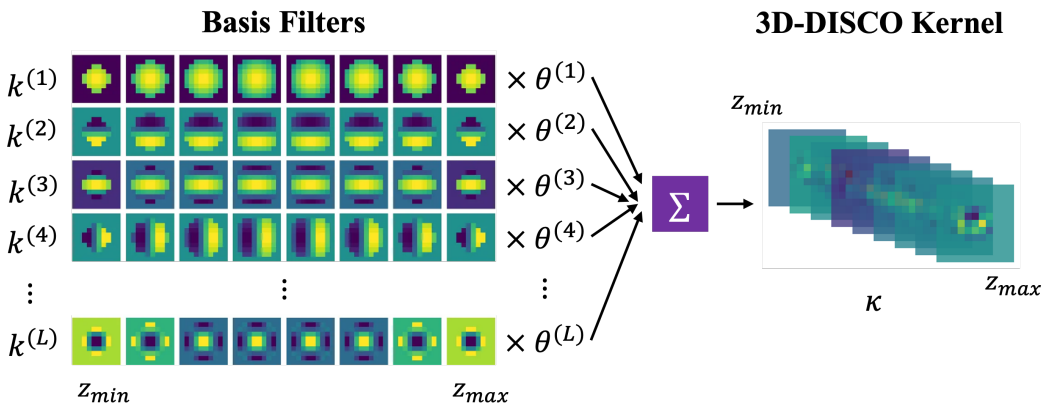

Figure 3: **3D Morlet-wavelet DISCO basis.** Left: cross-section slices through individual basis functions. Right: a 3D DISCO kernel formed as a learned linear combination of these functions. The discretization shown is for illustration only; the basis is continuous, so the same learned coefficients can be sampled at any target resolution.

3.2.2 DISCO Kernel Bases

We introduce two continuous bases of functions used to parameterize the 3D kernels in KIR-NO: 1) a Morlet-wavelet basis, which provides smooth, oriented, and frequency-selective filters. 2) a piecewise linear basis, which provides strictly local interpolation over learned control nodes.

3D Morlet-Wavelet DISCO Kernels. This extends the 2D basis of Liu-Schiaffini et al. [14] to 3D. Although referred to as a "Morlet" basis for consistency with Liu-Schiaffini et al. [14], it is more precisely a Hann-windowed real Fourier basis rather than a Gaussian-modulated complex Morlet wavelet. Each basis function is constructed as a localized oscillatory pattern: a separable product of one-dimensional Fourier modes along the x_1, x_2, and x_3 coordinate directions defines the oscillatory component, while a compactly supported Hann-type window localizes the filter in space. Specifically, for local offset $\mathbf{x} = (x_1, x_2, x_3)$ and radial distance $r = \|\mathbf{x}\|_2$, we define

$$\kappa_{\mathrm{MW}}^{p,m,n}(\mathbf{x}) = W(r)\, H_{p,m,n}(\mathbf{x}), \qquad W(r) = \mathbf{1}_{[r \le r_c]} \cos^2(\pi r/(2r_c)) \tag{4}$$

with separable Fourier component $H_{p,m,n}(\mathbf{x}) = h_n(x_1)\, h_m(x_2)\, h_p(x_3)$ and 1D factors

$$h_j(s) = \cos(k_j \pi s/w) \text{ for even } j, \qquad h_j(s) = \sin(k_j \pi s/w) \text{ for odd } j, \tag{5}$$

where w denotes the half-width of the kernel support along each axis, with $w = r_c$ in our implementation, and k_j is the harmonic order for index j. Even indices $j \in \{p, m, n\}$ select cosine modes and odd indices select sine modes, so the multi-index (p, m, n) enumerates combinations of real Fourier modes across the three spatial directions. Larger k_j yield higher-frequency basis functions, while the window W localizes them within the support radius r_c.

The final kernel is learned as a linear combination of these filters and can be discretized at arbitrary resolutions. Figure 3 shows representative basis filters and the resulting DISCO kernel construction.

3D Piecewise Linear DISCO Kernels. This basis represents each continuous kernel with local piecewise-linear "hat" functions $\Lambda(t) = \max\{1 - |t|, 0\}$ in spherical coordinates centered at the convolution point. For a local offset $\mathbf{x}$, let (r, θ, φ) denote its spherical coordinates, with $r \in [0, r_c]$, $\theta \in [0, \pi]$, and $\varphi \in [-\pi, \pi)$. For kernel shape $(n_r, n_\theta, n_\varphi)$, control nodes lie on n_r evenly spaced radii $r_a = a\,\Delta r$, $a = 0, \ldots, n_r{-}1$, with $\Delta r = r_c/(n_r{-}1)$. Since the angles are undefined at $r = 0$, the origin carries a single isotropic node, while each of the $n_r{-}1$ outer shells carries an $n_\theta \times n_\varphi$ angular grid

$$\theta_b = \left(b - \tfrac{1}{2}\right) \Delta\theta, \qquad \varphi_c = -\pi + \left(c - \tfrac{1}{2}\right) \Delta\varphi, \tag{6}$$

with $\Delta\theta = \pi/n_\theta$ and $\Delta\varphi = 2\pi/n_\varphi$, giving $L = (n_r{-}1)\, n_\theta n_\varphi + 1$ basis functions (101 for $[5, 5, 5]$).

The origin node defines the isotropic function $\kappa^0_{\mathrm{PL}}(\mathbf{x}) = \Lambda(r/\Delta r)$, and each shell node $\ell = (a, b, c)$, $a \geq 1$, defines a product of one-dimensional hat functions with a hard radial cutoff,

$$\kappa^{\ell}_{\mathrm{PL}}(\mathbf{x}) = \mathbf{1}_{[r \leq r_c]}\, \Lambda\big(\tfrac{r-r_a}{\Delta r}\big)\, \Lambda\big(\tfrac{\theta-\theta_b}{\Delta\theta}\big)\, \Lambda\Big(\tfrac{d_{\mathrm{circ}}(\varphi,\varphi_c)}{\Delta\varphi}\Big), \tag{7}$$

where $d_{\mathrm{circ}}(\varphi, \varphi_c) = \min(|\varphi - \varphi_c|,\, 2\pi - |\varphi - \varphi_c|)$ enforces 2π-periodicity in azimuth.

Rather than mixing fixed oscillatory filters, the model learns nodal kernel values θ^{ℓ} and evaluates the kernel by piecewise multilinear interpolation between nearby control nodes, enabling independent radial and angular resolution. Although less frequency-selective than Morlet, this strictly local basis better matches locally piecewise-smooth MRI images and sparse k-space residuals. Used in both the image-space cascades and the k-space neural-operator head, it improves over Morlet by $+2.11\,\mathrm{dB}$ PSNR at quarter resolution under an identical kernel configuration and training budget — despite using fewer coefficients — so KIR-NO adopts it as the default DISCO parameterization (Table 11).

4 Experiments

We evaluate on SKM-TEA [10], a large-scale multi-coil qDESS knee MRI dataset, and BraTS 2021 [29]. Models are trained on quarter- or half-resolution volumes and evaluated zero-shot at the other resolutions, without fine-tuning unless noted. Dataset and preprocessing details appear in Appendix A; implementation details in Appendix B. We compare ***KIR-NO*** against *(a)* classical reconstruction methods, including Zero-Filled (adjoint SENSE [30]) and ℓ_1-Wavelet compressed sensing [17]; *(b)* conventional CNN baselines, including 2D E2E-VarNet (per-slice), E2E-VarNet3D [4], and a 3D U-Net [31]; and *(c)* diffusion baselines, Diffusion-3D Physics [32] and Diffusion-CSGM [33]. For all methods, the coil sensitivity maps are supplied as a known forward operator rather than estimated by the network. They are estimated from scanner calibration for SKM-TEA and simulated for BraTS (Appendix A). In particular, our E2E-VarNet baselines use these provided maps in place of the original method's learned sensitivity-map module [4], ensuring that every method shares an identical forward model.

Two axes of discretization-agnosticism. Our experiments isolate two independent axes. *(i) Image-space resolution* (Tables 1, 2): the image-space operators (NO_i) are re-sampled on a finer voxel grid at inference, so a model trained at 256 reconstructs at 512 with no retraining (the zero-shot super-resolution setting). Crucially, at every evaluation resolution, the k-space head its not re-discretized: it keeps its trained kernel at ever evaluation resolution and only the image head after the inverse transform is resampled onto the finer grid. The model thus receives no additional measurements at higher resolution, so any gain at full resolution comes from the learned model rather than from additional acquired frequencies. Because the image-space kernels are continuous, they are sampled at the new voxel spacing and keep the same physical support, whereas a CNN applies its fixed discrete kernels on the finer grid, which shrinks their receptive field in physical units. *(ii) k-space sampling rate* (Table 3): at fixed image resolution (256), one trained model handles all undersampling factors ($4\times$ /$8\times$ /$16\times$) zero-shot, only the mask changing, exercising the k-space head (NO_k) and data-consistency at corresponding sampling rate. In neither case is k-space interpolated to change resolution; sub-native resolutions are formed by image-space resizing followed by an FFT (Appendix A).

Zero-shot super-resolution on SKM-TEA. Table 1 reports both training regimes. Trained at half resolution, KIR-NO leads decisively at the in-domain half setting ($36.54\,\mathrm{dB}$ / 0.924 SSIM; $+2.23\,\mathrm{dB}$ over E2E-VarNet3D and $+1.50\,\mathrm{dB}$ over the strongest CNN) and, most notably, under zero-shot *upward* super-resolution to full resolution ($38.36\,\mathrm{dB}$ / 0.940 SSIM, best on all three metrics, $+3.03\,\mathrm{dB}$ over the best CNN). The CNN baselines transfer stably in PSNR at full resolution (all $\sim$35.2–35.3 dB) but lose structural fidelity (3D U-Net SSIM drops to 0.827), whereas KIR-NO's SSIM rises to 0.940. At the coarsest zero-shot-*downward* setting (quarter), KIR-NO is competitive ($29.64\,\mathrm{dB}$) though the CNNs are marginally higher on this grid. Trained at quarter resolution, KIR-NO again leads at in-domain quarter ($33.13\,\mathrm{dB}$ / 0.894) and at full-resolution SR ($35.52\,\mathrm{dB}$ / 0.908).

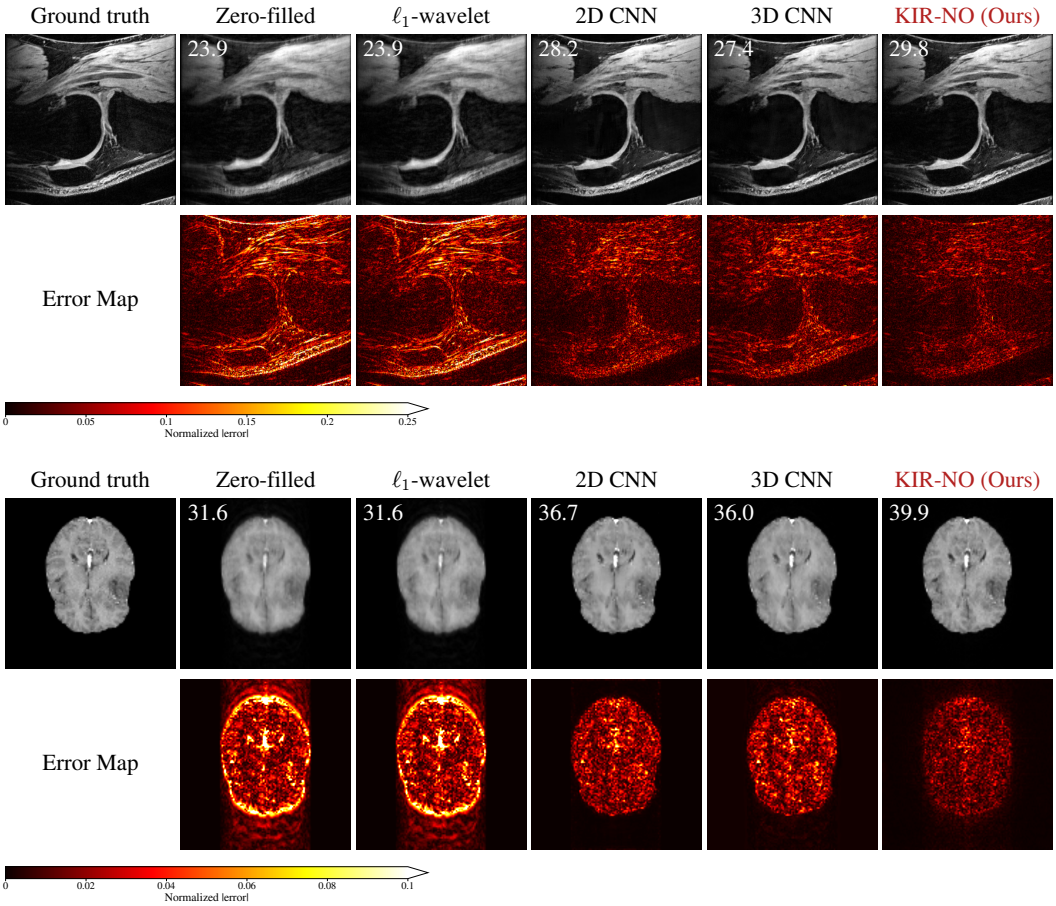

Figure 4: **Qualitative comparison on knee (top) and brain (bottom) MRI reconstruction.** Reconstructions of a representative slice with corresponding error maps (warmer colors indicate larger residuals). All models are trained and evaluated on half resolution data. The top left number is the PSNR (dB) on the shown slice. The ground truth (GT) is compared to classical baselines (Zero-Filled, ℓ_1-Wavelet) and learned baselines — 2D E2E-VarNet, 3D E2E-VarNet, and our KIR-NO. Classical methods show strong aliasing; KIR-NO has the smallest residuals on both samples. Dataset-level results are in Tables 1 and 2.

Zero-shot super-resolution on BraTS. Table 2 applies the same two-regime protocol to BraTS brain MRI (8-coil simulated k-space); every row is evaluated on the same 20 test volumes at all three resolutions, and the E2E-VarNets are trained with the same budget as KIR-NO (100 volumes per epoch for 120 epochs). Trained at quarter resolution, KIR-NO leads in-domain by $+1.67$ dB over the 2D E2E-VarNet (40.14 vs 38.47 dB) with the best NMSE (0.0075), while the CNN baselines transfer upward more strongly on this simulated 8-coil data (2D E2E-VarNet 39.84 / 41.42 dB vs KIR-NO 39.07 / 40.34 dB at half / full resolution). Trained at half resolution, KIR-NO leads in-domain (43.37 dB; $+1.24$ dB over the 2D and $+2.41$ dB over the 3D E2E-VarNet) and at full resolution (43.72 dB; $+0.05$ / $+0.64$ dB), with the best NMSE at both. Figure 4 shows representative reconstructions and error maps for both anatomies.

Robustness to acceleration. Table 3 reports half-resolution reconstruction at $4\times$, $8\times$, and $16\times$ acceleration. KIR-NO leads at every rate, improving over the 3D CNN baseline by $+2.23$ dB, $+2.40$ dB, and $+1.97$ dB, respectively. KIR-NO retains the highest SSIM at every rate (0.924, 0.895, 0.776 at $4\times$/$8\times$/$16\times$, versus E2E-VarNet3D's 0.883, 0.847, 0.716), remaining roughly $+0.04$–0.06 SSIM ahead across rates.

Table 1: Zero-shot super-resolution on SKM-TEA (3D Knee MRI), $4\times$ acceleration. KIR-NO is trained *once per regime* and evaluated at all three resolutions *without retraining or fine-tuning*; the in-domain (training) resolution is indicated in each panel. Classical methods are training-free. Table 6 extends the comparison. Single-cascade, single-step models are reported here; deeper-cascade and multi-step results appear in Appendices E.8 and D.

	Quarter Res (128)			**Half Res (256)**			**Full Res SR (512)**		
Method	PSNR↑	SSIM↑	NMSE↓	PSNR↑	SSIM↑	NMSE↓	PSNR↑	SSIM↑	NMSE↓
Learning-Free Methods (training-free)									
Zero-Filled	$26.59_{\pm 1.31}$	$0.700_{\pm 0.036}$	$0.1074_{\pm 0.0222}$	$29.80_{\pm 1.49}$	$0.797_{\pm 0.031}$	$0.0616_{\pm 0.0148}$	$34.46_{\pm 1.30}$	$0.882_{\pm 0.016}$	$0.0241_{\pm 0.0061}$
ℓ_1-Wavelet	$26.62_{\pm 1.31}$	$0.701_{\pm 0.036}$	$0.1067_{\pm 0.0221}$	$29.82_{\pm 1.49}$	$0.797_{\pm 0.031}$	$0.0612_{\pm 0.0147}$	$34.47_{\pm 1.30}$	$0.882_{\pm 0.016}$	$0.0240_{\pm 0.0061}$
Trained @ half resolution (in-domain: Half; zero-shot: Quarter, Full)									
2D E2E-VarNet (CNN)	$\mathbf{30.41}_{\pm 1.47}$	$\mathbf{0.849}_{\pm 0.029}$	$\mathbf{0.0452}_{\pm 0.0122}$	$35.04_{\pm 1.58}$	$0.906_{\pm 0.020}$	$0.0186_{\pm 0.0054}$	$35.33_{\pm 1.24}$	$0.896_{\pm 0.016}$	$0.0195_{\pm 0.0041}$
3D E2E-VarNet (CNN)	$30.08_{\pm 1.55}$	$0.828_{\pm 0.040}$	$0.0490_{\pm 0.0140}$	$34.31_{\pm 1.73}$	$0.883_{\pm 0.027}$	$0.0223_{\pm 0.0073}$	$35.32_{\pm 1.38}$	$0.870_{\pm 0.025}$	$0.0199_{\pm 0.0060}$
3D U-Net	$30.33_{\pm 1.39}$	$0.847_{\pm 0.026}$	$0.0458_{\pm 0.0114}$	$34.64_{\pm 1.51}$	$0.879_{\pm 0.017}$	$0.0204_{\pm 0.0056}$	$35.24_{\pm 1.23}$	$0.827_{\pm 0.015}$	$0.0200_{\pm 0.0048}$
Diffusion-3D Physics	$26.99_{\pm 1.33}$	$0.716_{\pm 0.035}$	$0.0982_{\pm 0.0214}$	$32.42_{\pm 1.57}$	$0.827_{\pm 0.027}$	$0.0341_{\pm 0.0104}$	$35.02_{\pm 1.30}$	$0.892_{\pm 0.015}$	$0.0212_{\pm 0.0056}$
Diffusion-CSGM	$26.61_{\pm 1.31}$	$0.701_{\pm 0.036}$	$0.1069_{\pm 0.0222}$	$29.89_{\pm 1.50}$	$0.799_{\pm 0.031}$	$0.0603_{\pm 0.0146}$	$34.61_{\pm 1.31}$	$0.885_{\pm 0.016}$	$0.0233_{\pm 0.0060}$
KIR-NO (Ours)	$29.64_{\pm 1.27}$	$0.770_{\pm 0.027}$	$0.0534_{\pm 0.0115}$	$\mathbf{36.54}_{\pm 1.60}$	$\mathbf{0.924}_{\pm 0.016}$	$\mathbf{0.0132}_{\pm 0.0040}$	$\mathbf{38.36}_{\pm 1.25}$	$\mathbf{0.940}_{\pm 0.010}$	$\mathbf{0.0097}_{\pm 0.0021}$
Trained @ quarter resolution (in-domain: Quarter; zero-shot: Half, Full)									
2D E2E-VarNet (CNN)	$31.45_{\pm 1.46}$	$0.874_{\pm 0.027}$	$0.0357_{\pm 0.0099}$	$33.29_{\pm 1.43}$	$\mathbf{0.878}_{\pm 0.023}$	$0.0275_{\pm 0.0068}$	$33.11_{\pm 1.19}$	$0.852_{\pm 0.020}$	$0.0325_{\pm 0.0063}$
3D E2E-VarNet (CNN)	$30.65_{\pm 1.50}$	$0.852_{\pm 0.030}$	$0.0430_{\pm 0.0124}$	$33.11_{\pm 1.52}$	$0.850_{\pm 0.040}$	$0.0291_{\pm 0.0088}$	$34.35_{\pm 1.40}$	$0.845_{\pm 0.040}$	$0.0252_{\pm 0.0088}$
3D U-Net	$31.43_{\pm 1.43}$	$0.860_{\pm 0.022}$	$0.0358_{\pm 0.0100}$	$33.63_{\pm 1.46}$	$0.859_{\pm 0.023}$	$0.0256_{\pm 0.0069}$	$34.88_{\pm 1.28}$	$0.855_{\pm 0.027}$	$0.0218_{\pm 0.0052}$
KIR-NO (Ours)	$\mathbf{33.13}_{\pm 1.44}$	$\mathbf{0.894}_{\pm 0.020}$	$\mathbf{0.0242}_{\pm 0.0069}$	$\mathbf{33.69}_{\pm 1.45}$	$0.872_{\pm 0.021}$	$\mathbf{0.0251}_{\pm 0.0060}$	$\mathbf{35.52}_{\pm 1.22}$	$\mathbf{0.908}_{\pm 0.014}$	$\mathbf{0.0187}_{\pm 0.0038}$

Table 2: Zero-shot super-resolution on BraTS 2021 (3D Brain MRI) reconstruction at $4\times$ acceleration, 8-coil. KIR-NO is trained *once per regime* and evaluated at all three resolutions; the in-domain (training) resolution is indicated in each panel.

	Quarter (60)			**Half (120)**			**Full (240)**		
Method	PSNR↑	SSIM↑	NMSE↓	PSNR↑	SSIM↑	NMSE↓	PSNR↑	SSIM↑	NMSE↓
Learning-Free Methods (training-free)									
Zero-Filled	$30.59_{\pm 1.45}$	$0.756_{\pm 0.028}$	$0.0663_{\pm 0.0135}$	$34.14_{\pm 1.08}$	$0.801_{\pm 0.024}$	$0.0355_{\pm 0.0109}$	$37.04_{\pm 1.15}$	$0.853_{\pm 0.022}$	$0.0243_{\pm 0.0078}$
ℓ_1-Wavelet	$30.60_{\pm 1.45}$	$0.756_{\pm 0.028}$	$0.0661_{\pm 0.0134}$	$34.16_{\pm 1.08}$	$0.800_{\pm 0.024}$	$0.0354_{\pm 0.0108}$	$37.05_{\pm 1.15}$	$0.852_{\pm 0.022}$	$0.0242_{\pm 0.0078}$
Trained @ half resolution (in-domain: Half; zero-shot: Quarter, Full)									
2D E2E-VarNet (CNN)	$36.51_{\pm 1.65}$	$\mathbf{0.939}_{\pm 0.014}$	$0.0170_{\pm 0.0038}$	$42.13_{\pm 1.28}$	$\mathbf{0.981}_{\pm 0.004}$	$0.0057_{\pm 0.0017}$	$43.67_{\pm 1.43}$	$\mathbf{0.974}_{\pm 0.007}$	$0.0053_{\pm 0.0017}$
3D E2E-VarNet (CNN)	$\mathbf{36.70}_{\pm 1.51}$	$0.936_{\pm 0.012}$	$\mathbf{0.0163}_{\pm 0.0039}$	$40.96_{\pm 1.25}$	$0.960_{\pm 0.007}$	$0.0074_{\pm 0.0022}$	$43.08_{\pm 1.20}$	$0.973_{\pm 0.006}$	$0.0059_{\pm 0.0014}$
Diffusion-3D Physics	$30.80_{\pm 1.43}$	$0.674_{\pm 0.055}$	$0.0631_{\pm 0.0130}$	$34.50_{\pm 1.08}$	$0.823_{\pm 0.026}$	$0.0328_{\pm 0.0102}$	$37.28_{\pm 1.16}$	$0.881_{\pm 0.020}$	$0.0230_{\pm 0.0074}$
Diffusion-CSGM	$30.68_{\pm 1.45}$	$0.759_{\pm 0.028}$	$0.0649_{\pm 0.0133}$	$34.28_{\pm 1.08}$	$0.803_{\pm 0.024}$	$0.0345_{\pm 0.0107}$	$37.21_{\pm 1.16}$	$0.855_{\pm 0.022}$	$0.0234_{\pm 0.0077}$
KIR-NO (Ours)	$34.50_{\pm 1.76}$	$0.828_{\pm 0.040}$	$0.0265_{\pm 0.0028}$	$\mathbf{43.37}_{\pm 1.34}$	$0.960_{\pm 0.007}$	$\mathbf{0.0043}_{\pm 0.0016}$	$\mathbf{43.72}_{\pm 1.12}$	$0.933_{\pm 0.011}$	$\mathbf{0.0052}_{\pm 0.0015}$
Trained @ quarter resolution (in-domain: Quarter; zero-shot: Half, Full)									
2D E2E-VarNet (CNN)	$38.47_{\pm 1.55}$	$\mathbf{0.969}_{\pm 0.004}$	$0.0109_{\pm 0.0026}$	$\mathbf{39.84}_{\pm 1.26}$	$0.953_{\pm 0.014}$	$\mathbf{0.0095}_{\pm 0.0027}$	$\mathbf{41.42}_{\pm 1.19}$	$0.958_{\pm 0.013}$	$\mathbf{0.0086}_{\pm 0.0022}$
3D E2E-VarNet (CNN)	$37.73_{\pm 1.47}$	$0.964_{\pm 0.004}$	$0.0129_{\pm 0.0031}$	$37.84_{\pm 1.46}$	$\mathbf{0.972}_{\pm 0.004}$	$0.0146_{\pm 0.0022}$	$40.55_{\pm 1.15}$	$\mathbf{0.979}_{\pm 0.004}$	$0.0104_{\pm 0.0017}$
KIR-NO (Ours)	$\mathbf{40.14}_{\pm 1.48}$	$0.958_{\pm 0.008}$	$\mathbf{0.0075}_{\pm 0.0022}$	$39.07_{\pm 1.18}$	$0.870_{\pm 0.021}$	$0.0114_{\pm 0.0034}$	$40.34_{\pm 1.00}$	$0.864_{\pm 0.019}$	$0.0111_{\pm 0.0027}$

Table 3: Half-resolution reconstruction at increasing acceleration rates on SKM-TEA. Each learned method is represented by a single trained model, evaluated at all three rates without retraining; only the k-space sampling mask changes between columns.

	$4\times$			$8\times$			$16\times$		
Method	PSNR	SSIM	NMSE	PSNR	SSIM	NMSE	PSNR	SSIM	NMSE
Zero-Filled	$29.80_{\pm 1.49}$	$0.797_{\pm 0.031}$	$0.0616_{\pm 0.0148}$	$28.74_{\pm 1.47}$	$0.758_{\pm 0.034}$	$0.0784_{\pm 0.0178}$	$26.31_{\pm 1.45}$	$0.668_{\pm 0.043}$	$0.1366_{\pm 0.0278}$
ℓ_1-Wavelet	$29.82_{\pm 1.49}$	$0.797_{\pm 0.031}$	$0.0612_{\pm 0.0147}$	$28.76_{\pm 1.47}$	$0.758_{\pm 0.034}$	$0.0780_{\pm 0.0177}$	$26.33_{\pm 1.45}$	$0.668_{\pm 0.043}$	$0.1360_{\pm 0.0277}$
2D E2E-VarNet (CNN)	$35.04_{\pm 1.58}$	$0.906_{\pm 0.020}$	$0.0186_{\pm 0.0054}$	$32.82_{\pm 1.49}$	$0.873_{\pm 0.024}$	$0.0307_{\pm 0.0076}$	$28.16_{\pm 1.45}$	$0.757_{\pm 0.035}$	$0.0893_{\pm 0.0188}$
3D E2E-VarNet (CNN)	$34.31_{\pm 1.73}$	$0.883_{\pm 0.027}$	$0.0223_{\pm 0.0073}$	$32.14_{\pm 1.66}$	$0.847_{\pm 0.032}$	$0.0364_{\pm 0.0105}$	$27.25_{\pm 1.53}$	$0.716_{\pm 0.047}$	$0.1102_{\pm 0.0241}$
KIR-NO (Ours)	$\mathbf{36.54}_{\pm 1.60}$	$\mathbf{0.924}_{\pm 0.016}$	$\mathbf{0.0132}_{\pm 0.0040}$	$\mathbf{34.54}_{\pm 1.54}$	$\mathbf{0.895}_{\pm 0.020}$	$\mathbf{0.0209}_{\pm 0.0058}$	$\mathbf{29.22}_{\pm 1.47}$	$\mathbf{0.776}_{\pm 0.035}$	$\mathbf{0.0703}_{\pm 0.0165}$

Robustness to sampling pattern. Table 14 examines sensitivity to the k-space sampling pattern. Beyond the Gaussian-density Cartesian mask used elsewhere, we retrain both E2E-VarNets and KIR-NO for each of equispaced, radial, and 2D Poisson-disc patterns at quarter resolution and $4\times$ acceleration (Appendix E.9). KIR-NO attains the highest PSNR and SSIM under every pattern (Table 14); the Poisson-disc pattern is the most favorable, reaching $35.30\,\mathrm{dB}$ ($+2.17\,\mathrm{dB}$ over the Gaussian default). This shows that KIR-NO's gains are not specific to the Gaussian Cartesian mask.

Convergence efficiency. Figure 1(e) reports the first epoch at which each model reaches a target PSNR on half-resolution validation. KIR-NO reaches $34\,\mathrm{dB}$ at epoch 5 versus 31 for the E2E-VarNet3D CNN baseline, about $6\times$ faster, and reaches $36\,\mathrm{dB}$ at epoch 39, a threshold the CNN

baseline never reaches within 120 epochs. Combined with the $+2.23\,\mathrm{dB}$ final-PSNR gain at half resolution (Table 3), these results show improved reconstruction quality and convergence efficiency.

Resolution-wise PSNR. Figure 1(d) plots PSNR at quarter, half, and full resolution for the zero-filled baseline and the quarter-trained 2D CNN, 3D CNN and KIR-NO of Table 1. KIR-NO leads at every resolution (33.13, 33.69 and $35.52\,\mathrm{dB}$). At full resolution the 2D and 3D CNN baselines (E2E-VarNet) fall below the training-free zero-filled reconstruction (33.11 and 34.35 vs $34.46\,\mathrm{dB}$), while KIR-NO stays $1.06\,\mathrm{dB}$ above it. The k-space head ablation is in Appendix E.5 (Table 10).

5 Conclusion

We introduced *KIR-NO*, a resolution-agnostic neural-operator framework for accelerated 3D MRI reconstruction built on discrete-continuous convolutions (DISCO). By parameterizing kernels in continuous space while preserving locality, KIR-NO enables a coarse-to-fine paradigm: train with full backpropagation on memory-feasible low-resolution volumes, and deploy zero-shot at high resolution. On SKM-TEA, KIR-NO improves reconstruction quality across acceleration rates and under zero-shot super-resolution, and converges faster than the CNN baseline; on BraTS it leads in-domain in both training regimes, and transfers upward more strongly from half resolution training. Ablations show that the k-space head drives the high-resolution gains, that additional image-space cascades further improve accuracy, and that the piecewise-linear DISCO basis outperforms the Morlet-wavelet basis.

Limitations. The full 3D model needs large GPU memory, and we use a single image-space cascade by default; gradient checkpointing enables deeper unrolling, which yields further gains for both KIR-NO and the CNN baseline (Table 13) at additional compute cost. Zero-shot super-resolution operates entirely in image space: the k-space head and data-consistency use the fixed training-resolution measurements, and only the post-transform image space operators are resampled onto the finer grid. On SKM-TEA the full-resolution in-plane ground truth contains acquired frequencies well beyond the training extent (416×512 acquisition matrix, zero-padded to 512×512 by the dataset), so it is a non-circular target for this prediction; the through-plane axis is interpolated and may inflate metrics there. The evaluation does not capture real-world noise, motion, or coil-geometry shifts of genuinely distinct high-resolution acquisitions. Experiments are restricted to two anatomies (SKM-TEA, BraTS) with single-seed runs; while we ablate different sampling patterns (Table 14), multi-contrast acquisitions, more aggressive accelerations, and matched comparisons against the broader score-based and unrolled-network landscape would further strengthen the empirical positioning.

Acknowledgment

This work is supported in part by ONR (MURI grant N000142312654 and N000142012786). R.D. is supported in part by the Undergraduate Research Fellowships (SURF) at Caltech. A.A. is supported in part by Bren endowed chair and the AI2050 senior fellow program at Schmidt Sciences.

References

[1] Mark A Griswold, Peter M Jakob, Robin M Heidemann, Mathias Nittka, Vladimir Jellus, Jianmin Wang, Berthold Kiefer, and Axel Haase. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 47(6):1202–1210, 2002.

[2] Michael Lustig, David L Donoho, Juan M Santos, and John M Pauly. Compressed sensing MRI. *IEEE signal processing magazine*, 25(2):72–82, 2008.

[3] 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*, 79(6):3055–3071, 2018.

[4] Anuroop Sriram, Jure Zbontar, Tullie Murrell, Aaron Defazio, C Lawrence Zitnick, Nafissa Yakubova, Florian Knoll, and Patricia Johnson. End-to-end variational networks for accelerated MRI reconstruction. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part II 23*, pages 64–73. Springer, 2020.

[5] Armeet Singh Jatyani, Jiayun Wang, Aditi Chandrashekar, Zihui Wu, Miguel Liu-Schiaffini, Bahareh Tolooshams, and Anima Anandkumar. A unified model for compressed sensing MRI across undersampling patterns. In *Conference on Computer Vision and Pattern Recognition (CVPR) Proceedings*, 2025.

[6] Hyungjin Chung and Jong Chul Ye. Score-based diffusion models for accelerated MRI. *Medical Image Analysis*, page 102479, 2022.

[7] Ke Wang, Michael Kellman, Christopher M Sandino, Kevin Zhang, Shreyas S Vasanawala, Jonathan I Tamir, Stella X Yu, and Michael Lustig. Memory-efficient learning for high-dimensional MRI reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 461–470. Springer, 2021.

[8] Ken CL Wong, Hongzhi Wang, and Tanveer Syeda-Mahmood. FNOSeg3D: resolution-robust 3D image segmentation with Fourier neural operator. In *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, pages 1–5. IEEE, 2023.

[9] DJ Husband, KA Grant, and CS Romaniuk. MRI in the diagnosis and treatment of suspected malignant spinal cord compression. *The British journal of radiology*, 74:15–23, 2001.

[10] 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, et al. SKM-TEA: A dataset for accelerated MRI reconstruction with dense image labels for quantitative clinical evaluation. *arXiv preprint arXiv:2203.06823*, 2022.

[11] Kamyar Azizzadenesheli, Nikola Kovachki, Zongyi Li, Miguel Liu-Schiaffini, Jean Kossaifi, and Anima Anandkumar. Neural operators for accelerating scientific simulations and design. *Nature Reviews Physics*, pages 1–9, 2024.

[12] Julius Berner, Miguel Liu-Schiaffini, Jean Kossaifi, Valentin Duruisseaux, Boris Bonev, Kamyar Azizzadenesheli, and Anima Anandkumar. Principled approaches for extending neural architectures to function spaces for operator learning. *arXiv preprint arXiv:2506.10973*, 2025.

[13] Jeremy Ocampo, Matthew A Price, and Jason D McEwen. Scalable and equivariant spherical CNNs by discrete-continuous (DISCO) convolutions. *arXiv preprint arXiv:2209.13603*, 2022.

[14] Miguel Liu-Schiaffini, Julius Berner, Boris Bonev, Thorsten Kurth, Kamyar Azizzadenesheli, and Anima Anandkumar. Neural operators with localized integral and differential kernels. In *Forty-first International Conference on Machine Learning*, 2024.

[15] Aujasvit Datta, Jiayun Wang, Asad Aali, Armeet Singh Jatyani, and Anima Anandkumar. Resolution-independent neural operators for multi-rate sparse-view CT. *arXiv preprint arXiv:2512.12236*, 2025.

[16] Jiayun Wang, Yousuf Aborahama, Arya Khokhar, Yang Zhang, Chuwei Wang, Karteekeya Sastry, Julius Berner, Yilin Luo, Boris Bonev, Zongyi Li, Kamyar Azizzadenesheli, Lihong V. Wang, and Anima Anandkumar. Accelerating 3D photoacoustic computed tomography with end-to-end physics-aware neural operators, 2025. URL `https://arxiv.org/abs/2509.09894`.

[17] Michael Lustig, David Donoho, and John M Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 58(6):1182–1195, 2007.

[18] Tony Chan, Selim Esedoglu, Frederick Park, and Andy Yip. Total variation image restoration: Overview and recent developments. *Handbook of mathematical models in computer vision*, pages 17–31, 2006.

[19] Weijie Gan, Shirin Shoushtari, Yuyang Hu, Jiaming Liu, Hongyu An, and Ulugbek S. Kamilov. Block coordinate plug-and-play methods for blind inverse problems. *Advances in Neural Information Processing Systems*, 36:78607–78620, 2023.

[20] Ulugbek S Kamilov, Charles A Bouman, Gregery T Buzzard, and Brendt Wohlberg. Plug-and-play methods for integrating physical and learned models in computational imaging: Theory, algorithms, and applications. *IEEE Signal Processing Magazine*, 40(1):85–97, 2023.

[21] Yan Yang, Jian Sun, Huibin Li, and Zongben Xu. Deep ADMM-Net for compressive sensing MRI. *Advances in neural information processing systems*, 29, 2016.

[22] Patricia M Johnson and Maria Drangova. Conditional generative adversarial network for 3D rigid-body motion correction in MRI. *Magnetic resonance in medicine*, 82(3):901–910, 2019.

[23] Nikola Kovachki, Samuel Lanthaler, and Siddhartha Mishra. On universal approximation and error bounds for Fourier neural operators. *J. Mach. Learn. Res.*, 22(1), 2021. ISSN 1532-4435.

[24] Zongyi Li, Hongkai Zheng, Nikola Kovachki, David Jin, Haoxuan Chen, Burigede Liu, Kamyar Azizzadenesheli, and Anima Anandkumar. Physics-informed neural operator for learning partial differential equations. *ACM/JMS Journal of Data Science*, 1(3):1–27, 2024.

[25] Ryan Y. Lin, Julius Berner, Valentin Duruisseaux, David Pitt, Daniel Leibovici, Jean Kossaifi, Kamyar Azizzadenesheli, and Anima Anandkumar. Enabling automatic differentiation with mollified graph neural operators. *Transactions on Machine Learning Research*, 2025. ISSN 2835-8856. URL `https://openreview.net/forum?id=CGoR1hFAGr`.

[26] Adarsh Ganeshram, Haydn Maust, Valentin Duruisseaux, Zongyi Li, Yixuan Wang, Daniel Leibovici, Oscar Bruno, Thomas Hou, and Anima Anandkumar. FC-PINO: High precision physics-informed neural operators via Fourier continuation, 2025.

[27] Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations, 2021.

[28] Valentin Duruisseaux, Jean Kossaifi, and Anima Anandkumar. Fourier neural operators explained: A practical perspective, 2025. URL `https://arxiv.org/abs/2512.01421`.

[29] Ujjwal Baid, Satyam Ghodasara, Suyash Mohan, Michel Bilello, Evan Calabrese, Errol Colak, Keyvan Farahani, Jayashree Kalpathy-Cramer, Felipe C Kitamura, Sarthak Pati, et al. The RSNA-ASNR-MICCAI BraTS 2021 benchmark on brain tumor segmentation and radiogenomic classification. *arXiv preprint arXiv:2107.02314*, 2021.

[30] Klaas P. Pruessmann, Markus Weiger, Markus B. Scheidegger, and Peter Boesiger. SENSE: Sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine*, 42(5):952–962, 1999.

[31] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional networks for biomedical image segmentation, 2015. URL `https://arxiv.org/abs/1505.04597`.

[32] Anselm Krainovic, Stefan Ruschke, and Reinhard Heckel. Resolution-robust 3D MRI reconstruction with 2D diffusion priors: Diverse-resolution training outperforms interpolation. *arXiv preprint arXiv:2412.18584*, 2024.

[33] Ajil Jalal, Marius Arvinte, Giannis Daras, Eric Price, Alexandros G Dimakis, and Jonathan I Tamir. Robust compressed sensing MRI with deep generative priors. *Advances in Neural Information Processing Systems*, 2021.

[34] Wei Li, Jiawei Jiang, Jie Wu, Kaihao Yu, and Jianwei Zheng. LMO: Linear Mamba operator for MRI reconstruction. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5112–5122. IEEE, 2025.

[35] Pengfei Guo, Yiqun Mei, Jinyuan Zhou, Shanshan Jiang, and Vishal M Patel. ReconFormer: Accelerated MRI reconstruction using recurrent transformer. *IEEE transactions on medical imaging*, 43(1):582–593, 2023.

[36] Jure Zbontar, Florian Knoll, Anuroop Sriram, Tullie Murrell, Zhengnan Huang, Matthew J. Muckley, Aaron Defazio, Ruben Stern, Patricia Johnson, Mary Bruno, Marc Parente, Krzysztof J. Geras, Joe Katsnelson, Hersh Chandarana, Zizhao Zhang, Michal Drozdzal, Adriana Romero, Michael Rabbat, Pascal Vincent, Nafissa Yakubova, James Pinkerton, Duo Wang, Erich Owens, C. Lawrence Zitnick, Michael P. Recht, Daniel K. Sodickson, and Yvonne W. Lui. fastMRI: An open dataset and benchmarks for accelerated MRI, 2018.

Appendix

We provide details omitted in the main text in the appendix, including:

- **Appendix A: Dataset and Setup.** Preprocessing and coarse-to-fine resolution protocol for SKM-TEA, plus the multi-coil k-space simulation pipeline used for BraTS 2021.
- **Appendix B: KIR-NO Implementation.** Architecture, training hardware, hyperparameters, and compute for KIR-NO and the CNN baselines.
- **Appendix C: Memory Footprint.** Why full-resolution training is memory-bound, and how the cost of the DISCO operators scales with resolution.
- **Appendix D: Diffusion Comparison.** Single-step and 10-step comparisons against two diffusion baselines.
- **Appendix E: Additional Ablations and Experiments.** Additional baselines; robustness to measurement noise; tumor-ROI fidelity on BraTS; cross-scanner transfer (SKM-TEA $\rightarrow$ fastMRI); the k-space head; basis function choice; kernel configuration and radius cutoff; number of cascades; and sampling-pattern robustness.
- **Appendix F: Broader Impacts.** Potential benefits, risks, and the limits of clinical use.

A Dataset and Setup

SKM-TEA. The SKM-TEA dataset [10] is a multi-coil knee MRI dataset acquired with the qDESS sequence, with two echoes per sample; we use the first echo in all experiments. Scans use a 416×512 acquisition matrix (readout $\times$ phase encode; 0.38×0.31 mm in-plane), 8 or 16 receive coils, and 80 slices. The released k-space is zero-padded by the dataset to 512×512 in-plane and to twice the number of acquired slices, giving the $512 \times 512 \times 160$ volumes we start from.

We further interpolate the last dimension from 160 to 320, yielding volumes of $512 \times 512 \times 320$. We denote this as the full-resolution dataset. Downscaling by a factor of 2 gives the half-resolution dataset, $256\times256\times160$, and a further factor of 2 gives the quarter-resolution dataset, $128\times128\times80$. This enables a coarse-to-fine evaluation protocol: models trained at quarter resolution are evaluated zero-shot at half and full resolution, while models trained at half resolution are evaluated zero-shot at quarter and full resolution. At every resolution, the k-space data are formed by resizing the volume in *image* space, followed by applying the FFT and the undersampling mask. We never interpolate in k-space, so the k-space data at each resolution are the exact transform of a clean image on that grid. At zero-shot evaluation, a learned model keeps the measurements of its training resolution; only the intermediate image is resampled to the evaluation grid (Section 4). No model sees evaluation-resolution examples during training.

BraTS 2021. The BraTS 2021 dataset [29] is a brain tumor segmentation benchmark containing pre-operative multimodal MRI volumes. We use the T1Gd modality with 100 volumes for training per epoch and 20 test volumes for evaluation. The original data are preprocessed NIfTI volumes at $240 \times 240 \times 155$ image-space resolution with isotropic $1\,\text{mm}^3$ voxel spacing.

Unlike SKM-TEA, BraTS provides only reconstructed magnitude images, so we synthesize multi-coil k-space. For each volume, we generate an 8-coil sensitivity map on a virtual receive array, multiply the magnitude image by each coil profile to obtain complex coil images, and apply an orthonormal 3D FFT to produce simulated k-space. We then apply a $4\times$ Gaussian-density Cartesian undersampling mask, as for SKM-TEA.

To match the SKM-TEA coarse-to-fine protocol, we generate full-, half-, and quarter-resolution volumes of size $240 \times 240 \times 155$, $120 \times 120 \times 78$, and $60 \times 60 \times 39$, respectively, using trilinear downsampling in image space before the FFT. KIR-NO and the baselines are trained once per regime, at half ($120\times120\times78$) or at quarter ($60\times60\times39$) resolution, and evaluated at all three resolutions (Table 2). The tumor-ROI study in Appendix E.3 uses the quarter-trained models. Because BraTS starts from coil-combined magnitude images rather than raw multi-coil acquisitions, the simulated setting lacks the acquisition diversity of SKM-TEA. As for SKM-TEA, the coil sensitivity maps are provided to all models as a known operator.

B KIR-NO for 3D MRI Reconstruction from Undersampled Measurements

KIR-NO is trained on a single NVIDIA B200 with a batch size of one volume. Each neural operator "NO" layer in Fig. 2 is a U-shaped DISCO Neural Operator [5].

We evaluate two DISCO basis functions, the 3D Morlet and 3D piecewise-linear bases (Section 3.2.2). On SKM-TEA, KIR-NO uses the piecewise-linear basis with a kernel configuration $5 \times 5 \times 5$ with a radius cutoff of 0.08 for the quarter-resolution model and 0.04 for the half-resolution model. Unless stated otherwise, KIR-NO uses a single cascade with 18 channels along with 4 encoder, decoder and pooling layers for SKM-TEA and 2 encoder, decoder and pooling layers for BraTS. Ablations over the basis function and the number of cascades are in Appendix E.6 (Table 11) and Appendix E.8 (Table 13).

The 3D CNN baseline is E2E-VarNet [4] with 2D convolutions replaced by 3D convolutions; it uses the provided coil sensitivity maps in place of the learned sensitivity module (Section 4).

All models are trained with the Adam optimizer, and hyperparameters were selected on validation performance. On SKM-TEA, all models are trained for 200 epochs at quarter resolution and 120 epochs at half resolution; on BraTS, all learned models are trained for 120 epochs. On BraTS, both KIR-NO models use a radius cutoff of 0.302 and a kernel configuration of $7 \times 7 \times 7$.

Compute requirements. Training the main SKM-TEA half-resolution KIR-NO model (120 epochs) takes approximately 43 GPU-hours on a single NVIDIA B200; the quarter-resolution model (200 epochs) takes approximately 5.4 GPU-hours.

C Memory Footprint and Coarse-to-Fine Deployability

Full-volume 3D reconstruction is memory-bound at training time: a single full-resolution ($512 \times 512 \times 320$) training step exceeded the memory of a GH200 GPU for *both* the 3D E2E-VarNet CNN and KIR-NO, and KIR-NO's full-resolution inference alone needs 72.2 GiB before any activations are stored for back-propagation. Native full-resolution training would require model parallelism, gradient checkpointing, or resolution downscaling. The coarse-to-fine protocol avoids this: models are trained where back-propagation is memory-feasible (quarter or half resolution) and deployed zero-shot at full resolution (Tables 1–2).

The learned DISCO coefficients do not depend on resolution, but each kernel is defined over a fixed physical radius r_c. On a finer grid, this radius contains more voxels, so the discrete stencil obtained by sampling the kernel has more entries. The inference cost per output voxel of the image-space operators therefore grows with the evaluation resolution.

Full-resolution *inference* is feasible on high-memory GPUs. Because the image-space DISCO operators are local, peak inference memory could in principle be reduced further by exact tiled execution with coil-streamed data consistency, trading runtime for memory.

D Diffusion Comparison

We compare KIR-NO with two diffusion-based reconstruction methods [32, 33] under matched inference budgets. Since diffusion models typically require many sampling steps at test time, we evaluate two regimes: a *single-forward* setting, in which KIR-NO is evaluated with one forward pass and each diffusion baseline with one reverse-sampling step, and a *multi-step iterative* setting, in which both methods are allowed 10 inference steps. For KIR-NO, the 10 steps are 10 unrolled image-space cascades: a separately trained 10-cascade model, evaluated in one forward pass. For the diffusion baselines, the steps are 10 score-based reverse-diffusion steps.

Table 4 reports the quantitative comparison. In the single-forward setting, KIR-NO achieves 33.13 dB at quarter resolution and 33.69 dB at half resolution, ahead of the strongest one-step diffusion baseline by $+6.14$ dB and $+1.27$ dB respectively, and ahead on SSIM and NMSE at both resolutions. The 10-step setting is less uniform. At quarter resolution KIR-NO reaches 35.94 dB and stays $+6.21$ dB ahead of the best 10-step diffusion baseline, and that is 2.81 dB above its own single-forward output and better on SSIM and NMSE as well, so the additional cascades help at this resolution. At half resolution KIR-NO improves to 34.25 dB, a gain of $+0.56$ dB over its sin-

gle forward, while the 3D physics-based baseline reaches 37.75 dB and leads on all three metrics, overtaking KIR-NO by 3.50 dB.

Table 4: Iterative reconstruction on SKM-TEA at quarter and half resolution ($4\times$ acceleration; KIR-NO trained at quarter resolution, diffusion priors at half resolution). *Top:* single-forward setting, where each method performs one inference step. *Bottom:* multi-step setting with 10 inference steps: KIR-NO is a 10-cascade model evaluated in one forward pass, while the diffusion baselines take 10 score-based reverse-diffusion steps. Results at the diffusion baselines' full sampling trajectories are given in the text below.

	Quarter-Res (128)			**Half-Res (256)**		
Method	PSNR↑	SSIM↑	NMSE↓	PSNR↑	SSIM↑	NMSE↓
Single forward						
Diffusion-3D Physics [32]	$26.99_{\pm 1.33}$	$0.716_{\pm 0.035}$	$0.0982_{\pm 0.0214}$	$32.42_{\pm 1.57}$	$0.827_{\pm 0.027}$	$0.0341_{\pm 0.0104}$
Diffusion-CSGM [33]	$26.61_{\pm 1.31}$	$0.701_{\pm 0.036}$	$0.1069_{\pm 0.0222}$	$29.89_{\pm 1.50}$	$0.799_{\pm 0.031}$	$0.0603_{\pm 0.0146}$
KIR-NO (Ours)	$\mathbf{33.13}_{\pm 1.44}$	$\mathbf{0.894}_{\pm 0.020}$	$\mathbf{0.0242}_{\pm 0.0069}$	$\mathbf{33.69}_{\pm 1.45}$	$\mathbf{0.872}_{\pm 0.021}$	$\mathbf{0.0251}_{\pm 0.0060}$
Multi-step iterative (10)						
Diffusion-3D Physics [32]	$29.73_{\pm 1.41}$	$0.787_{\pm 0.032}$	$0.0530_{\pm 0.0150}$	$\mathbf{37.75}_{\pm 1.51}$	$\mathbf{0.921}_{\pm 0.014}$	$\mathbf{0.0100}_{\pm 0.0029}$
Diffusion-CSGM [33]	$28.25_{\pm 1.43}$	$0.751_{\pm 0.034}$	$0.0742_{\pm 0.0192}$	$34.82_{\pm 1.76}$	$0.885_{\pm 0.024}$	$0.0200_{\pm 0.0072}$
KIR-NO (Ours)	$\mathbf{35.94}_{\pm 1.44}$	$\mathbf{0.933}_{\pm 0.011}$	$\mathbf{0.0127}_{\pm 0.0037}$	$34.25_{\pm 1.45}$	$0.899_{\pm 0.018}$	$0.0221_{\pm 0.0056}$

Diffusion baselines at their original sampling budgets. The comparison above equalizes the number of inference steps. We also evaluate both diffusion baselines at the sampling budgets their authors specify: 200 reverse steps for the 3D physics-based model and 1200 annealed Langevin steps for CSGM. Table 5 reports per-volume reconstruction time on BraTS 8-coil volumes at $4\times$ acceleration, on a single idle B200.

Table 5: Per-volume inference time on BraTS 8-coil volumes at $4\times$, for the diffusion baselines and KIR-NO.

Method	**QR** $60 \times 60 \times 39$	**HR** $120 \times 120 \times 78$
KIR-NO (single cascade)	0.157 s	3.22 s
Diffusion-3D Physics (single step)	0.079 s	0.096 s
Diffusion-CSGM (2D, single step, batched)	0.060 s	0.082 s
Diffusion-3D Physics (200 steps)	10.1 s	13.1 s
Diffusion-CSGM (2D, 1200 steps)	35.6 s	58.9 s
KIR-NO (10 cascades)	1.14 s	17.8 s

A single KIR-NO cascade reconstructs a quarter-resolution volume in 0.157 s (Table 5); the 3D physics-based baseline takes $64\times$ more time per volume at that resolution with full sampling. The margin is resolution-dependent and much smaller at half resolution, because the operator's kernel support scales with the grid while a sampling step's does not. All times are device-resident reconstruction.

Accuracy at the full sampling trajectory. Given enough steps, both diffusion baselines are accurate. Sampling CSGM along its dense trajectory (1200 steps) on SKM-TEA gives 36.88 dB / 0.9357 / 0.0101 at quarter resolution and 44.47 dB / 0.9829 / 0.0021 at half resolution; the 3D physics-based model reaches 35.04 dB / 0.922 and 38.46 dB / 0.933 on the same grids. The 3D physics-based model already exceeds KIR-NO at half resolution with 10 steps (Table 4); along its full trajectory it does so at quarter resolution as well, and CSGM exceeds KIR-NO at both resolutions. KIR-NO's single forward pass is 1.9–3.8 dB below these values at quarter resolution and 4.8–10.8 dB below at half resolution; its advantage is inference time (Table 5).

E Additional Ablations and Experiments

This section reports additional baselines, robustness tests (measurement noise, cross-scanner transfer, sampling pattern), tumor-ROI metrics, and ablations of the k-space head, basis function, kernel configuration and radius cutoff, and number of cascades.

E.1 Additional Baseline Comparison

Beyond the classical and CNN baselines in the main text, we compare against three additional recent learned reconstruction methods: LMO (Linear Mamba Operator) [34], the recurrent transformer ReconFormer [35], and the 2D neural-operator CS-MRI model of Jatyani et al. [5] ("2D-NO"). All learned methods are trained on quarter-resolution SKM-TEA ($128\times128\times80$, $4\times$) and evaluated across resolutions with fixed weights, so half and full resolution are out-of-domain for every method. For these 2D (slice-wise) baselines we report PSNR only. KIR-NO outperforms LMO and 2D-NO at every resolution, whereas ReconFormer attains higher PSNR than KIR-NO at all three resolutions.

Table 6: Comparison with additional baselines on SKM-TEA ($4\times$), PSNR in dB. All learned models are trained at quarter resolution ($128\times128\times80$) and evaluated at quarter (in-domain), half, and full resolution with fixed weights.

Method	**Quarter (128)**	**Half (256)**	**Full (512)**
LMO [34]	29.12	31.49	33.02
ReconFormer [35]	34.16	36.80	39.05
2D-NO [5]	28.76	29.30	31.11
KIR-NO (Ours)	33.13	33.69	35.52

E.2 Robustness to Measurement Noise

We stress-test robustness to acquisition noise by injecting complex Gaussian noise (AWGN) at the *sampled* k-space positions at a controlled measurement SNR and evaluating at half resolution ($4\times$). The noise is applied at evaluation time only: both models are the clean-trained half-resolution models of Tables 1 and 3, so the comparison is *zero-shot* for the noisy data. KIR-NO's advantage narrows monotonically as the measurement SNR falls: $+2.23$ dB with no noise, $+1.81$ dB at 20 dB SNR and $+0.79$ dB at 14 dB. On SSIM the CNN already leads at 14 dB SNR noise (0.8624 versus 0.8409), where KIR-NO still leads on PSNR. KIR-NO thus leads in PSNR at every tested SNR but degrades faster than the CNN as noise increases (-2.14 dB from clean to 14 dB SNR, against the CNN's -0.70 dB); see Table 7.

Table 7: Robustness to measurement noise on SKM-TEA half resolution ($4\times$), PSNR in dB; both models are trained on clean half-resolution evaluated zero-shot

Measurement SNR (dB)	**No noise**	**20**	**14**
3D E2E-VarNet	$34.31_{\pm1.73}$	$34.14_{\pm1.70}$	$33.61_{\pm1.62}$
KIR-NO (Ours)	$\mathbf{36.54}_{\pm1.60}$	$\mathbf{35.95}_{\pm1.51}$	$\mathbf{34.40}_{\pm1.38}$

E.3 Clinical ROI Metrics (Tumor Fidelity)

Beyond whole-image PSNR/SSIM/NMSE, we report tumor-region metrics (Table 8) at quarter resolution ($60 \times 60 \times 39$, $4\times$, 8 coils) for KIR-NO and the 3D E2E-VarNet, both trained at quarter resolution, on the 20 test volumes of Table 2. The ROI is defined volumetrically for the whole tumor: comprising the necrotic core, edema and enhancing tumor. Metrics are reported as PSNR / SSIM / NMSE.

Table 8: BraTS tumor-region fidelity at quarter resolution ($4\times$, 8 coils). Both learned models are trained at quarter resolution.

	Whole image			**Tumor ROI**		
Method	PSNR↑	SSIM↑	NMSE↓	PSNR↑	SSIM↑	NMSE↓
Zero-Filled	$30.59_{\pm 1.45}$	$0.756_{\pm 0.028}$	$0.0663_{\pm 0.0135}$	$23.95_{\pm 2.65}$	$0.750_{\pm 0.047}$	$0.0427_{\pm 0.0216}$
3D E2E-VarNet (CNN)	$37.73_{\pm 1.47}$	$\mathbf{0.964}_{\pm 0.004}$	$0.0129_{\pm 0.0031}$	$28.39_{\pm 1.64}$	$0.908_{\pm 0.024}$	$0.0142_{\pm 0.0042}$
KIR-NO (Ours)	$\mathbf{40.14}_{\pm 1.48}$	$0.958_{\pm 0.008}$	$\mathbf{0.0075}_{\pm 0.0022}$	$\mathbf{31.02}_{\pm 1.98}$	$\mathbf{0.947}_{\pm 0.014}$	$\mathbf{0.0081}_{\pm 0.0035}$

We observe that KIR-NO outperforms baseline methods for tumor ROI regions.

E.4 Cross-Scanner Transfer (SKM-TEA → fastMRI Knee)

We test a genuinely different acquisition: models trained on quarter-resolution SKM-TEA and applied to fastMRI knee [36], which differs in scanner, coil array, protocol, and SNR. In the few-shot setting both models are fine-tuned for 15 epochs on 100 fastMRI knee volumes. Evaluation is on 30 unseen fastMRI whole-knee volumes at native resolution (Table 9).

Table 9: Cross-scanner transfer to fastMRI knee (PSNR / SSIM), on 30 unseen whole-knee volumes at native resolution ($0.44 \times 0.44 \times 3$ mm, $4\times$). Both models are trained on quarter-resolution SKM-TEA; few-shot fine-tunes on 100 fastMRI volumes for 15 epochs.

Evaluation on fastMRI knee	**Zero-shot**	**Few-shot**
3D E2E-VarNet (CNN)	26.33 / 0.711	31.10 / 0.824
KIR-NO (Ours)	**29.12 / 0.769**	**36.17 / 0.929**

KIR-NO leads in both settings, and its margin over the VarNet widens with fine-tuning ($+2.79 \rightarrow +5.07$ dB).

E.5 Effect of the k-space Head (kNO)

We ablate the k-space neural-operator head by removing it, leaving only the image-space DISCO cascades (*KIR-NO w/o kNO*). Both variants are trained at quarter resolution and evaluated at quarter, half, and full resolution (Table 10). The kNO head is key to KIR-NO's high-resolution performance: without it, PSNR *drops* as evaluation resolution grows ($31.29 \rightarrow 30.22 \rightarrow 28.89$ dB), whereas the full model *improves* ($33.13 \rightarrow 33.69 \rightarrow 35.52$ dB), widening the gap to $+6.6$ dB at full resolution. The image-space cascades alone therefore do not account for KIR-NO's high-resolution gains.

Table 10: Effect of the k-space head. Both variants are trained at quarter resolution SKM-TEA.

	Quarter (128)			**Half (256)**			**Full (512)**		
Method	PSNR	SSIM	NMSE	PSNR	SSIM	NMSE	PSNR	SSIM	NMSE
KIR-NO w/o kNO	$31.29_{\pm 1.47}$	$0.865_{\pm 0.024}$	$0.0370_{\pm 0.0104}$	$30.22_{\pm 1.36}$	$0.823_{\pm 0.028}$	$0.0554_{\pm 0.0107}$	$28.89_{\pm 1.18}$	$0.832_{\pm 0.020}$	$0.0850_{\pm 0.0117}$
KIR-NO (ours)	$\mathbf{33.13}_{\pm 1.44}$	$\mathbf{0.894}_{\pm 0.020}$	$\mathbf{0.0242}_{\pm 0.0069}$	$\mathbf{33.69}_{\pm 1.45}$	$\mathbf{0.872}_{\pm 0.021}$	$\mathbf{0.0251}_{\pm 0.0060}$	$\mathbf{35.52}_{\pm 1.22}$	$\mathbf{0.908}_{\pm 0.014}$	$\mathbf{0.0187}_{\pm 0.0038}$

E.6 Basis Function Choice

We compare two DISCO bases: piecewise-linear3d and Morlet3d. Table 11 isolates the basis: both KIR-NO models are trained and evaluated at SKM-TEA quarter resolution ($128\times128\times80$) with the *identical* kernel configuration $(5, 5, 5)$, radius cutoff (0.08) and 200-epoch training protocol of Table 1, and differ only in the basis. Piecewise-linear3d outperforms Morlet3d by $+2.11\,\mathrm{dB}$ (33.13 vs $31.02\,\mathrm{dB}$). Crucially, the Morlet kernel carries *more* learnable coefficients than PL, so the piecewise-linear advantage reflects the basis's inductive bias.

Table 11: Basis-function ablation on SKM-TEA quarter-resolution reconstruction.

Model	**Basis**	**PSNR ↑**	**SSIM ↑**	**NMSE ↓**
KIR-NO	Morlet	$31.02_{\pm1.41}$	$0.8320_{\pm0.027}$	$0.0393_{\pm0.0106}$
KIR-NO	Piecewise-linear	$\mathbf{33.13}_{\pm1.44}$	$\mathbf{0.8943}_{\pm0.020}$	$\mathbf{0.0242}_{\pm0.0069}$

E.7 Kernel Configuration and Radius Cutoff

Table 12 ablates the kernel configuration and radius cutoff for KIR-NO using the piecewise-linear 3D basis. Two trends are apparent. First, a radius cutoff of 0.08 outperforms other settings like $0.05, 0.13$ across all kernel configurations, suggesting that a medium-large DISCO receptive field is important for reconstruction quality. Second, the $7 \times 7 \times 7$ kernel with cutoff 0.05 achieves the best performance, reaching 33 dB and slightly outperforming other kernel configurations. Additionally, reconstruction performance is relatively robust to slight changes in hyperparameters, like kernel configurations and convolution size $\mathbf{r_c}$.

Table 12: Kernel configuration and radius-cutoff sweep for KIR-NO on SKM-TEA Dataset

		Quarter Res (128)			**Half Res (256)**		
Kernel Config	$\mathbf{r_c}$	PSNR↑	SSIM↑	NMSE↓	PSNR↑	SSIM↑	NMSE↓
$3 \times 3 \times 3$	0.05	$32.74_{\pm1.42}$	$0.886_{\pm0.021}$	$0.0265_{\pm0.0073}$	$33.38_{\pm1.47}$	$0.859_{\pm0.023}$	$0.0270_{\pm0.0066}$
$5 \times 5 \times 5$	0.05	$32.86_{\pm1.43}$	$0.888_{\pm0.021}$	$0.0258_{\pm0.0073}$	$33.02_{\pm1.41}$	$0.851_{\pm0.023}$	$0.0292_{\pm0.0064}$
$7 \times 7 \times 7$	0.05	$32.94_{\pm1.42}$	$0.890_{\pm0.020}$	$0.0253_{\pm0.0069}$	$33.03_{\pm1.38}$	$0.856_{\pm0.022}$	$0.0289_{\pm0.0054}$
$3 \times 3 \times 3$	0.08	$31.76_{\pm1.47}$	$0.860_{\pm0.026}$	$0.0332_{\pm0.0094}$	$32.77_{\pm1.40}$	$0.853_{\pm0.023}$	$0.0308_{\pm0.0064}$
$5 \times 5 \times 5$	0.08	$33.13_{\pm1.44}$	$\mathbf{0.894}_{\pm0.020}$	$0.0242_{\pm0.0069}$	$33.69_{\pm1.45}$	$0.872_{\pm0.021}$	$0.0251_{\pm0.0060}$
$7 \times 7 \times 7$	0.08	$\mathbf{33.15}_{\pm1.45}$	$\mathbf{0.894}_{\pm0.020}$	$\mathbf{0.0241}_{\pm0.0070}$	$\mathbf{33.89}_{\pm1.42}$	$\mathbf{0.877}_{\pm0.020}$	$\mathbf{0.0239}_{\pm0.0055}$
$3 \times 3 \times 3$	0.13	$30.69_{\pm1.37}$	$0.811_{\pm0.027}$	$0.0422_{\pm0.0106}$	$32.90_{\pm1.38}$	$0.826_{\pm0.025}$	$0.0299_{\pm0.0063}$
$5 \times 5 \times 5$	0.13	$32.69_{\pm1.45}$	$0.882_{\pm0.022}$	$0.0268_{\pm0.0078}$	$33.76_{\pm1.41}$	$0.873_{\pm0.021}$	$0.0247_{\pm0.0058}$
$7 \times 7 \times 7$	0.13	$33.00_{\pm1.45}$	$0.889_{\pm0.022}$	$0.0250_{\pm0.0073}$	$33.16_{\pm1.44}$	$0.846_{\pm0.026}$	$0.0282_{\pm0.0064}$

E.8 Deeper Cascades with Gradient Checkpointing

KIR-NO uses a single image-space cascade by default for memory efficiency and for a controlled comparison, but gradient checkpointing enables deeper unrolling. Table 13 reports quarter-resolution ($4\times$) reconstruction as the number of cascades increases for both the 3D CNN and KIR-NO with checkpointing. KIR-NO remains ahead at every matched cascade count (e.g., $35.25\,\mathrm{dB}$ / 0.931 SSIM versus $31.96\,\mathrm{dB}$ / 0.873 SSIM at three cascades), confirming that the advantage is not an artifact of the single-cascade default. Both models gain most from the first increase in depth; KIR-NO saturates thereafter, while the CNN regresses slightly at five cascades.

Table 13: Cascades ablation on SKM-TEA Quarter-res ($4\times$).

	E2E-VarNet3D (CNN)			KIR-NO (Ours)		
Cascades	PSNR	SSIM	NMSE	PSNR	SSIM	NMSE
cas $= 1$	$30.65_{\pm 1.50}$	$0.8519_{\pm 0.030}$	$0.0430_{\pm 0.0124}$	$\mathbf{33.13}_{\pm 1.44}$	$\mathbf{0.8943}_{\pm 0.020}$	$\mathbf{0.0242}_{\pm 0.0069}$
cas $= 3$	$31.96_{\pm 1.58}$	$0.8732_{\pm 0.032}$	$0.0319_{\pm 0.0096}$	$\mathbf{35.25}_{\pm 1.46}$	$\mathbf{0.9311}_{\pm 0.014}$	$\mathbf{0.0149}_{\pm 0.0045}$
cas $= 5$	$31.46_{\pm 1.39}$	$0.8582_{\pm 0.026}$	$0.0354_{\pm 0.0093}$	$\mathbf{35.33}_{\pm 1.47}$	$\mathbf{0.9314}_{\pm 0.013}$	$\mathbf{0.0147}_{\pm 0.0044}$

E.9 Sampling-Pattern Robustness

We assess robustness to the k-space sampling pattern, complementing the Gaussian-density Cartesian mask used in the main experiments. Table 14 reports quarter-resolution reconstruction under four patterns — Gaussian (paper default), equispaced, radial, and 2D Poisson-disc.

All learned models use a single cascade (cas=1) and are retrained per pattern at quarter resolution ($128\times128\times80$). For each of the other three patterns, a $4\times$ acceleration mask is applied to all training and evaluation volumes. KIR-NO attains the best PSNR, SSIM and NMSE under every pattern, with a PSNR margin over the strongest CNN ranging from $+1.5$ dB (radial) to $+2.9$ dB (Poisson); the Poisson-disc pattern yields the highest KIR-NO reconstruction overall (35.30 dB / 0.922 SSIM). These results indicate that the gains reported in the main text are not specific to the Gaussian mask.

Table 14: Sampling-pattern ablation on SKM-TEA Quarter-res.

	Zero-Filled			ℓ_1-Wavelet			2D E2E-VarNet			3D E2E-VarNet			KIR-NO (Ours)		
Pattern	PSNR	SSIM	NMSE	PSNR	SSIM	NMSE	PSNR	SSIM	NMSE	PSNR	SSIM	NMSE	PSNR	SSIM	NMSE
Gaussian (paper default)	$26.59_{\pm 1.31}$	$0.700_{\pm 0.036}$	$0.1074_{\pm 0.0222}$	$26.62_{\pm 1.31}$	$0.701_{\pm 0.036}$	$0.1067_{\pm 0.0221}$	$31.45_{\pm 1.46}$	$0.874_{\pm 0.027}$	$0.0357_{\pm 0.0099}$	$30.65_{\pm 1.50}$	$0.852_{\pm 0.030}$	$0.0430_{\pm 0.0124}$	$\mathbf{33.13}_{\pm 1.44}$	$\mathbf{0.894}_{\pm 0.020}$	$\mathbf{0.0242}_{\pm 0.0069}$
Equispaced	$26.44_{\pm 1.28}$	$0.699_{\pm 0.032}$	$0.1110_{\pm 0.0223}$	$26.46_{\pm 1.29}$	$0.699_{\pm 0.032}$	$0.1105_{\pm 0.0221}$	$31.91_{\pm 1.51}$	$0.880_{\pm 0.026}$	$0.0323_{\pm 0.0100}$	$30.93_{\pm 1.58}$	$0.855_{\pm 0.030}$	$0.0407_{\pm 0.0132}$	$\mathbf{33.48}_{\pm 1.53}$	$\mathbf{0.895}_{\pm 0.019}$	$\mathbf{0.0226}_{\pm 0.0074}$
Radial	$27.71_{\pm 1.31}$	$0.747_{\pm 0.032}$	$0.0833_{\pm 0.0188}$	$27.77_{\pm 1.32}$	$0.748_{\pm 0.032}$	$0.0823_{\pm 0.0186}$	$32.34_{\pm 1.44}$	$0.889_{\pm 0.023}$	$0.0290_{\pm 0.0081}$	$31.38_{\pm 1.53}$	$0.862_{\pm 0.026}$	$0.0364_{\pm 0.0108}$	$\mathbf{33.87}_{\pm 1.43}$	$\mathbf{0.912}_{\pm 0.017}$	$\mathbf{0.0205}_{\pm 0.0059}$
Poisson	$25.48_{\pm 1.11}$	$0.684_{\pm 0.027}$	$0.1370_{\pm 0.0169}$	$25.71_{\pm 1.11}$	$0.687_{\pm 0.027}$	$0.1301_{\pm 0.0166}$	$32.40_{\pm 1.48}$	$0.889_{\pm 0.023}$	$0.0287_{\pm 0.0085}$	$31.87_{\pm 1.55}$	$0.876_{\pm 0.023}$	$0.0326_{\pm 0.0099}$	$\mathbf{35.30}_{\pm 1.44}$	$\mathbf{0.922}_{\pm 0.014}$	$\mathbf{0.0147}_{\pm 0.0043}$

F Broader Impacts

Positive impacts. Accelerated MRI reconstruction reduces scan times, lowers patient discomfort (particularly for pediatric, claustrophobic, and trauma patients), and increases scanner throughput, which can broaden access to high-quality imaging in resource-constrained healthcare settings. Because KIR-NO can be trained on memory-feasible low-resolution volumes and evaluated at higher resolution, it lowers the GPU memory needed to train 3D MRI reconstruction models.

Potential negative impacts and mitigations. Learned MRI reconstruction methods can exhibit failure modes that are visually subtle, including hallucinated structures or suppressed pathology on out-of-distribution scans (e.g., unseen scanner vendors, contrasts, anatomies, or pathology). We emphasize that KIR-NO is presented as a research contribution and is *not* validated for clinical deployment. Any clinical use would require prospective validation, regulatory review, integration of uncertainty quantification, and a fallback to conventional reconstruction in out-of-distribution regimes. Our zero-shot evaluation explicitly tests cross-resolution generalization but does not address distribution shifts beyond the two anatomies and the sampling patterns studied here.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction make main claims. These claims are directly supported by the architecture description in Methods as well as experimental results.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: A dedicated *conclusion-Limitations* paragraph at the end of the Conclusion section discusses these.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [N/A]

   Justification: The paper is mainly empirical and does not state new theorems or provide formal proofs. The mathematical content is *constructive*: it specifies the unrolled image-space update (Eq. 2), the k-space neural-operator head (Eq. 1), the DISCO local-integration approximation (Eq. 3), and the two new 3D continuous kernel bases — piecewise-linear hat functions in spherical coordinates and 3D Morlet-wavelet filters (Sec. 3.2.2). These define the model architecture rather than establish theoretical claims, and they follow the established DISCO formulation of Ocampo et al. [13], Liu-Schiaffini et al. [14].

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: The architecture is fully specified in the Methods section, including the unrolled update (Eq. 2), the k-space head (Eq. 1), and both DISCO basis functions. Dataset preprocessing and the multi-coil k-space simulation pipeline are described.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: our code is available at github.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: Details are provided in the paper.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: Error bars reported in experimental results.

8. **Experiments compute resources**

Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

Answer: [Yes]

Justification: Information provided in the experimental section and Appendix B.

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics?

   Answer: [Yes]

   Justification: The research uses publicly available, IRB-approved, de-identified medical imaging datasets (SKM-TEA, BraTS 2021, fastMRI) and conforms to the NeurIPS Code of Ethics in all respects.

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: Broader impacts are discussed in Appendix F. Positive impacts include reduced MRI scan times, improved patient comfort, and increased access to high-resolution diagnostic imaging via zero-shot resolution transfer in resource-constrained settings. Potential negative impacts include reliance on reconstructions that may exhibit dataset-specific failure modes if deployed clinically without proper validation; we explicitly position the method as a research contribution rather than a clinically validated tool.

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: The released model is a task-specific MRI reconstruction model trained on publicly available, de-identified research datasets. It cannot be used for generative content creation, identification, or synthesis of personal data, and therefore poses no high-risk-for-misuse profile.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: All datasets and prior code are cited at first use. SKM-TEA [10] is used under its research-only license, BraTS 2021 [29] is used under the BraTS challenge research-use terms, and fastMRI [36] is used under the fastMRI data-use agreement for research use. Baseline implementations and library dependencies (E2E-VarNet [4], the `neuraloperator` library, `torch-harmonics`, the diffusion baselines [32, 33]) are credited and used in accordance with their respective open-source licenses.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: No new assets are released at submission time. Upon acceptance, the KIRNO codebase will be released publicly with a README documenting installation, dataset preparation, training entry points, evaluation scripts, and configuration files for every experiment reported in the paper. New 3D DISCO layers and basis functions will be released as extensions to the `neuraloperator` and `torch-harmonics` libraries with documented APIs.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: This work does not involve crowdsourcing or research directly conducted with human subjects. All data are sourced from publicly available, de-identified medical imaging datasets that obtained their own IRB approvals at the source institutions.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The work does not involve direct research with human subjects. All data come from publicly released, de-identified medical imaging datasets (SKM-TEA, BraTS 2021, fastMRI), each of which obtained the appropriate IRB approval at the source institutions prior to public release.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research?

    Answer: [N/A]

    Justification: The core methodology (DISCO-based neural operator for 3D MRI reconstruction) does not use LLMs as any important, original, or non-standard component. LLMs were used only as writing/editing aids and have no impact on scientific contributions.