DRIFT: Difficulty-aware Rectified Flows for Through-plane MRI Super-Resolution

Yoonseok Choi[1], Eun-Gyu Ha[1], Daniel Kim[1], Mohammed A. Al-masni[2], Ming-Hsuan Yang[3], and Dong-Hyun Kim[1]*

[1] Department of Electrical and Electronic Engineering, Yonsei University, Seoul, Republic of Korea
[2] Department of Bioengineering, King Fahd University of Petroleum & Minerals (KFUPM), Dhahran, Saudi Arabia
[3] Electrical Engineering and Computer Science, University of California at Merced, Merced, USA
{yoonseokchoi, eungyuha9722, danny4159, donghyunkim}@yonsei.ac.kr
m.almasani@sejong.ac.kr
mhyang@ucmerced.edu

Abstract. Magnetic Resonance Imaging (MRI) is often acquired with anisotropic resolution to reduce scan time, producing stair-step artifacts along the through-plane direction. In through-plane MRI super-resolution, an efficiency–fidelity trade-off arises: feed-forward regressors are fast but oversmooth at large slice-thicknesses, while sampling-based methods improve fidelity at high inference cost. We propose DRIFT, a two-stage thickness-conditioned rectified flow framework for through-plane MRI super-resolution with continuous input slice-thickness. Stage 1 employs an Anatomical Projection Network (APN) to map low-resolution patches to a coarse high-resolution manifold, providing a deterministic anatomical initialization that shortens the residual transport of Stage 2 and stabilizes slice-wise refinement. Stage 2 refines details via rectified flow and introduces a Physics-Aware Difficulty (PAD) metric derived from slice-thickness induced through-plane bandwidth deficit to guide an Adaptive Integration Scheduler (AIS), allocating ODE steps by thickness. A Consistent Endpoint Trajectory Alignment (CETA) loss enforces thickness-consistent reconstructions. Experiments show that DRIFT outperforms super-resolution baselines while reducing inference cost. Code, models, and interactive demos are available at `https://yoonseokchoi-ai.github.io/drift-eccv2026/`.

Keywords: Super-Resolution · Rectified Flow · MRI

1 Introduction

High-resolution (HR) isotropic Magnetic Resonance Imaging (MRI) provides crucial anatomical details for precise clinical diagnosis and quantitative neuroanatomical analysis. However, fundamental physical constraints such as the

* Corresponding author.

signal-to-noise ratio (SNR) and acquisition time often necessitate anisotropic acquisition, resulting in thick-slice volumes with low through-plane resolution. To bridge this gap, through-plane super-resolution (SR) approaches have emerged as a vital task to reconstruct isotropic volumes from anisotropic inputs.

Various approaches address through-plane SR under anisotropic acquisition, including self-supervision from in-plane and through-plane asymmetry [38], alias-aware regression for thick-slice artifacts [25], and synthesis-driven training for heterogeneous clinical protocols [10]. While these pipelines mitigate stair-step artifacts and improve reconstruction quality, regression-style estimators often favor conservative predictions and tend to be optimized for discrete target scales or protocol settings, which can limit generalization under the continuous slice-thickness variation encountered in practice. More broadly, continuously varying slice-thickness motivates SR formulations that represent images as continuous functions rather than relying on discrete grid-to-grid regression.

Implicit Neural Representations (INR) methods [6, 12] represent images as continuous functions of spatial coordinates, with medical adaptations such as ArSSR [34] and SA-INR [32] extending this paradigm to 3D brain MRI. Despite their flexibility, these models face two challenges that matter for through-plane SR in MRI. First, they suffer from spectral bias [23], where Multi-Layer Perceptron (MLP)-based continuous functions prioritize low-frequency structural components, which tends to under-recover high-frequency textures essential for clinical fidelity. Second, most existing arbitrary-scale models are trained using isotropic downsampling, which may not reflect the unique physics of MRI slice acquisition, where degradation is dominated by slice-profile governed through-plane integration [21, 22].

Generative approaches, particularly diffusion-based frameworks, have recently demonstrated superior performance in capturing complex textures. For instance, TPDM [14] utilizes pre-trained perpendicular 2D diffusion models to improve 3D imaging quality. However, the prohibitive inference cost remains a significant bottleneck for real-time clinical use. While recent advancements such as ResShift [36] have reduced sampling steps by initiating the sampling process from the low-resolution (LR) image instead of noise, they utilize a fixed-step scheduler regardless of the degradation severity of the input. In contrast, AdaDiffSR [8] introduces region-aware dynamic acceleration, but its image-dependent perception modules introduce non-negligible computational overhead and remain fundamentally detached from the underlying physical degradation (*i.e.*, slice-thickness) inherent in MRI acquisition.

We propose **D**ifficulty-aware **R**ectif**I**ed **F**low for **T**hrough-plane SR (DRIFT), a two-stage framework for reconstructing a specified target resolution from inputs with continuous slice-thicknesses (Fig. 1). DRIFT conditions both stages on the input and target slice thicknesses, allowing a single model to adapt its reconstruction process across different through-plane degradation levels. In Stage 1, DRIFT uses an Anatomical Projection Network (APN) to map LR patches to a coarse HR manifold, providing a structured initialization that shortens the subsequent refinement trajectory. In Stage 2, DRIFT refines high-frequency details using

rectified flow, enabling deterministic inference with a reduced number of function evaluations (NFEs) [17, 18]. During training, DRIFT simulates thickness-diverse anisotropic inputs using a slice-profile model grounded in the Shinnar-Le Roux (SLR) algorithm [21] which mirrors the slice-selection process in MRI acquisition. We sample slice-thicknesses from common clinical ranges [10, 11] and degradation axis, and apply an SLR-derived slice profile along the chosen axis to model through-plane signal integration. We then resample to the thick-slice grid and map the result back to the HR grid with nearest-neighbor interpolation to reproduce stair-step artifact morphology. To encourage thickness-consistent outputs across conditions, we introduce the Consistent Endpoint Trajectory Alignment (CETA) loss using proximal thickness pairs (Fig. 2).

At test time, DRIFT uses a Physics-Aware Difficulty (PAD) metric computed from slice-thickness metadata to drive an Adaptive Integration Scheduler (AIS), which allocates ordinary differential equation (ODE) steps without image-dependent decision modules [8] (Fig. 1). Our contributions are summarized as follows:

- We formulate input slice-thickness-continuous through-plane MRI SR as a two-stage anatomical projection-to-transport problem, where a thickness-conditioned APN provides deterministic and spatially correlated slice-wise initialization that shortens the residual rectified-flow trajectory.
- We introduce an MRI protocol-aware conditioning and inference mechanism: both stages are conditioned on input and target inverse thickness, while PAD/AIS maps slice-thickness-induced bandwidth deficit to adaptive ODE step budgets without auxiliary image-difficulty networks.
- We propose CETA, a proximal-thickness endpoint alignment loss that regularizes rectified-flow trajectories across neighboring slice-thickness conditions and improves over naive random-pair consistency.
- We validate DRIFT on public and real thick-slice MRI datasets with comprehensive reconstruction, ablation, standardized efficiency, zero-shot clinical transfer, and through-plane continuity analyses.

2 Related Works

Through-plane MRI SR. Through-plane MRI SR focuses on reconstructing isotropic volumes from anisotropic thick-slice scans, where the resolution is significantly lower along the slice-selection direction. Self-supervised frameworks such as SMORE [38] and SIMPLE [1] exploit the resolution asymmetry between in-plane and through-plane directions to learn reconstruction from the high-resolution in-plane data itself. To address aliasing artifacts inherent in thick-slice acquisition, AFCM [25] introduces an alias-free formulation using co-modulated generators. For clinically heterogeneous data, SynthSR [10] and the joint optimization framework by Iglesias *et al.* [11] utilize synthetic data from segmentation labels to ensure robustness. More recent methods explore generative priors. TPDM [14] leverages perpendicular 2D diffusion models for 3D reconstruction. Despite these advances, most methods assume discrete scale degradation models,

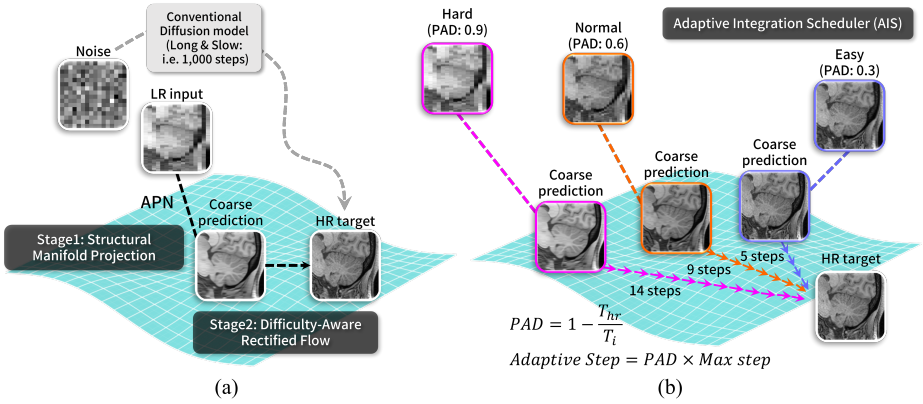

Fig. 1: Conceptual overview of DRIFT inference. (a) DRIFT initializes rectified flow from a coarse manifold estimate z rather than noise, shortening the residual transport trajectory. (b) AIS allocates ODE steps using the PAD metric, assigning more steps to thick-slice hard cases and fewer steps to thin-slice easy cases.

failing to generalize across the continuous slice-thickness manifold encountered in clinical practice.

Arbitrary-Scale and INR-based SR. Arbitrary-scale SR eliminates the requirement for scale-specific models by treating images as continuous functions. Meta-SR [9] first achieves this via dynamic weight prediction, but INRs such as LIIF [6] have since become the standard paradigm. Subsequent developments, LTE [12] and CiaoSR [3] improve this by incorporating frequency-domain texture estimation. In medical imaging, ArSSR [34] and SA-INR [32] adapt INRs to 3D medical volumes, while SAINT [22] uses voxel-spacing-conditioned weights for spatially aware interpolation. Despite their flexibility, coordinate-based MLP inherently suffers from spectral bias [23], favoring low-frequency structural components over high-frequency details. Furthermore, most existing INR models fail to explicitly incorporate the directional physics of MRI slice acquisition [21].

Generative models and adaptive inference. Diffusion-based SR methods such as SR3 [24] and SRDiff [15] improve perceptual quality but require costly iterative sampling. Acceleration methods such as ResShift [36] and AddSR [28] reduce sampling cost through residual shifting or distillation. Rectified Flow [18] and Flow Matching [17] learn straight-line ODE trajectories, with InstaFlow [19] and FlowSR [35] enabling few-step generation or restoration. Posterior-mean rectified flow (PMRF) [20] further formulates image restoration as a two-stage transport process, where a posterior-mean estimate is first obtained and then transported toward the clean image distribution by rectified flow. DRIFT shares this broad two-stage transport view, but specializes it for through-plane MRI SR by introducing acquisition-thickness conditioning, SLR-based slice-profile degradation, metadata-driven adaptive integration, and proximal-thickness trajectory regularization. Unlike image-dependent adaptive inference modules such as AdaDiffSR [8], DRIFT uses acquisition metadata to allocate computation

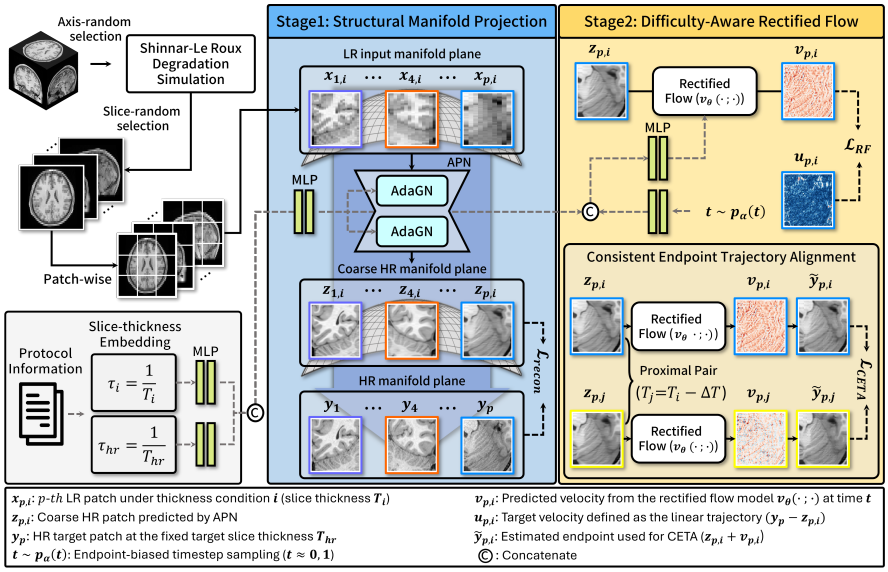

Fig. 2: Training pipeline of DRIFT. Stage 1 uses an Anatomical Projection Network (APN) to project LR patches to a coarse HR manifold conditioned on inverse thickness τ. Stage 2 refines textures via rectified flow by predicting a straight-line velocity field. We enforce Consistent Endpoint Trajectory Alignment (CETA) by regularizing trajectories between proximal thickness pairs ($T_j = T_i - \Delta T$, $\Delta T = 1\,\mathrm{mm}$). Notation is summarized in the figure legend.

without auxiliary difficulty networks, which is suitable for MRI protocols where slice thickness governs through-plane bandwidth loss.

3 Method

We propose DRIFT (Difficulty-aware RectifIed Flow for Through-plane SR), a two-stage, slice-thickness-conditioned framework for through-plane MRI SR. For the p-th training patch and a slice-thickness condition indexed by i with input thickness T_i, DRIFT reconstructs an isotropic target at thickness T_{hr} via

$$\mathbf{x}_{p,i} \xrightarrow[\text{Stage 1}]{\text{APN}} \mathbf{z}_{p,i} \xrightarrow[\text{Stage 2}]{\text{Rectified Flow}} \tilde{\mathbf{y}}_{p,i}. \tag{1}$$

Here, $\mathbf{x}_{p,i} \in \mathbb{R}^{H\times W}$ denotes the LR input patch under condition i, $\mathbf{y}_p \in \mathbb{R}^{H\times W}$ denotes the corresponding HR target patch at T_{hr}, $\mathbf{z}_{p,i}$ is a coarse anatomical estimate, and $\tilde{\mathbf{y}}_{p,i}$ is the final output. DRIFT is implemented as a 2D slice wise model. Each volume is reconstructed by applying the same thickness-conditioned operator to slices orthogonal to the degraded axis and then composing the predicted slices back into a 3D volume. Therefore, DRIFT does not rely on explicit 3D convolutions or cross-slice communication. Its volumetric continuity instead comes from deterministic APN initialization and shared deterministic rectified flow refinement across neighboring band limited slices.

3.1 Slice-profile based thick-slice simulation

Through-plane degradation in thick-slice MRI is largely governed by slice-profile induced signal integration along the slice axis. To approximate clinical slice selection, we synthesize LR inputs of thickness T_i from HR scans at thickness T_{hr} using an SLR-derived slice-profile model [21].

Volume-level simulation. Let $\mathbf{Y} \in \mathbb{R}^{H\times W\times D}$ denote an HR 3D volume sampled on an isotropic grid with spacing T_{hr}. At each iteration, we sample a degradation axis $a \in \{x, y, z\}$ and apply a thickness-specific 1D slice-profile kernel h_{T_i} by convolving $\mathbf{Y}$ only along axis a (leaving the other axes unchanged), yielding a slice-profile integrated volume $\mathbf{Y}'$. We then resample only along axis a from spacing T_{hr} to T_i (allowing non-integer factors), and represent the resulting thick-slice volume on the HR grid to form paired training data with $\mathbf{Y}$. This procedure reproduces the characteristic stair-step morphology observed in reformatted thick-slice MRI. Implementation details are provided in the supplementary material (Sec. S1).

Patch extraction. From the simulated thick-slice volume (represented on the HR grid) and the HR volume $\mathbf{Y}$, we extract 2D slices orthogonal to the degraded axis a and crop aligned patch pairs, corresponding to $\mathbf{x}_{p,i}$ and $\mathbf{y}_p$ defined above.

3.2 Slice-Thickness Conditioning

We condition DRIFT on slice-thickness using the inverse thickness $\tau = 1/T$, which proxies the effective through-plane bandwidth under slice selection [2,21]. For condition i, we set $\tau_i = 1/T_i$ for the input thickness and $\tau_{\mathrm{hr}} = 1/T_{\mathrm{hr}}$ for the target thickness. Using τ reduces the scale disparity between input and target thickness values over clinically common ranges and improves numerical conditioning. We embed τ_i and τ_{hr} using two lightweight, unshared MLP encoders, concatenate their embeddings, and map the result to a thickness-conditioning vector $\mathbf{c}_i \in \mathbb{R}^{d_c}$ (we use $d_c = 512$):

$$\mathbf{c}_i = \mathrm{MLP}_{\mathrm{cond}}([\mathrm{MLP}_{\mathrm{in}}(\tau_i) \parallel \mathrm{MLP}_{\mathrm{tgt}}(\tau_{\mathrm{hr}})]), \tag{2}$$

where $\parallel$ denotes concatenation, $\mathrm{MLP}_{\mathrm{in}}$ and $\mathrm{MLP}_{\mathrm{tgt}}$ encode the input and target inverse thickness, respectively, and $\mathrm{MLP}_{\mathrm{cond}}$ fuses the two embeddings into the final conditioning vector $\mathbf{c}_i$. In the main experiments, T_{hr} is fixed to the native isotropic resolution of each dataset, while the architecture itself accepts both input and target thickness embeddings. Target-continuous operation therefore requires sampling target thicknesses during training, which we examine in the supplementary material. We inject $\mathbf{c}_i$ into APN residual blocks via Adaptive Group Normalization (AdaGN) [7]:

$$\mathrm{AdaGN}(\mathbf{h}, \mathbf{c}) = \boldsymbol{\gamma}(\mathbf{c}) \odot \mathrm{GN}(\mathbf{h}) + \boldsymbol{\beta}(\mathbf{c}), \tag{3}$$

where $\mathbf{h} \in \mathbb{R}^{C\times H\times W}$ is a feature map, GN denotes Group Normalization, and $(\boldsymbol{\gamma}(\mathbf{c}), \boldsymbol{\beta}(\mathbf{c})) \in \mathbb{R}^{2\times C}$ are channel-wise scale and shift predicted from $\mathbf{c}$ by a learned linear projection. Architectural details of the MLPs are provided in the supplementary material (Sec. S4.1).

3.3 Stage 1: Structural Manifold Projection

Stage 1 is designed to shorten the rectified-flow transport in Stage 2. Given an LR patch $\mathbf{x}_{p,i}$ acquired under thickness condition i (input thickness T_i) and the thickness-conditioning vector $\mathbf{c}_i$, the Anatomical Projection Network (APN) f_ϕ produces a coarse HR estimate $\mathbf{z}_{p,i}$ at the target thickness T_{hr}:

$$\mathbf{z}_{p,i} = f_\phi(\mathbf{x}_{p,i}, \mathbf{c}_i). \tag{4}$$

We refer to this step as structural manifold projection because APN learns a thickness-conditioned mapping that projects samples from the thick-slice input space onto a coarse HR space anchored at T_{hr}. This projection reduces the remaining distance to the target HR distribution, allowing Stage 2 to focus on refining residual high-frequency details with fewer ODE steps. Beyond reducing the transport distance, APN also provides an image-conditioned initial state for slice-wise generative refinement. Because neighboring thick-slice inputs share anatomical content and are processed by the same deterministic APN, their coarse HR states remain spatially correlated before Stage 2. This differs from slice-wise stochastic generative reconstruction, where independent noise initialization can introduce slice-to-slice variations unless additional noise-correlation heuristics are used. Through-plane SR is ill-posed, and multiple HR solutions can explain the same thick-slice observation. Under pixel-wise regression, predictors tend to average over plausible solutions, resulting in smooth yet structurally consistent outputs. Accordingly, Stage 1 does not aim to recover the full HR texture distribution; instead, it provides a stable anatomical estimate close to the target HR manifold, which Stage 2 subsequently refines to recover sharp textures and boundaries. We train APN with a reconstruction objective that combines the Charbonnier penalty [4] and SSIM [33]:

$$\mathcal{L}_{\text{recon}} = \mathcal{L}_{\text{Char}}(\mathbf{z}_{p,i}, \mathbf{y}_p) + \lambda_{\text{ssim}} \mathcal{L}_{\text{SSIM}}(\mathbf{z}_{p,i}, \mathbf{y}_p), \tag{5}$$

where λ_{ssim} weights the SSIM term (we use $\lambda_{\text{ssim}} = 0.5$).

3.4 Stage 2: Difficulty-Aware Rectified Flow

Stage 2 refines high-frequency details that are under-recovered by Stage 1 by learning a rectified-flow velocity field [17,18]. During Stage 2 training, we freeze the Stage 1 APN so that the velocity network learns to refine a stable coarse estimate rather than a moving target. The velocity network takes the intermediate image state $\mathbf{s}_{p,i}(t)$ as input, while time t and slice-thickness information modulate intermediate features via AdaGN (Sec. 3.2). For training, we sample a normalized time $t \in [0, 1]$ and define a straight path between the Stage 1 output $\mathbf{z}_{p,i}$ and the HR target $\mathbf{y}_p$:

$$\mathbf{s}_{p,i}(t) = (1-t)\mathbf{z}_{p,i} + t\mathbf{y}_p, \qquad \mathbf{u}_{p,i} = \mathbf{y}_p - \mathbf{z}_{p,i}, \tag{6}$$

where $\mathbf{u}_{p,i}$ is the constant target velocity along the path. Given $\mathbf{s}_{p,i}(t)$, the velocity network v_θ predicts the rectified-flow velocity, with t and slice-thickness entering only through conditioning. Concretely, we compute a time embedding from

a sinusoidal time embedding [31], concatenate it with the thickness-conditioning vector $\mathbf{c}_i$ (Sec. 3.2), and map the concatenation to a joint time–thickness conditioning vector:

$$\mathbf{e}_t = \mathrm{MLP}_{\mathrm{time}}(\mathrm{SinEmb}(t)), \qquad \mathbf{c}_{t,i} = \mathrm{MLP}_{\mathrm{tt}}([\mathbf{e}_t \parallel \mathbf{c}_i]), \tag{7}$$

where $\parallel$ denotes concatenation and $\mathrm{MLP}_{\mathrm{time}}$ and $\mathrm{MLP}_{\mathrm{tt}}$ are lightweight MLPs (architectural details are provided in the supplementary material Sec. S4.1). The predicted velocity is

$$\mathbf{v}_{p,i}(t) = v_\theta(\mathbf{s}_{p,i}(t); \mathbf{c}_{t,i}), \tag{8}$$

where $\mathbf{s}_{p,i}(t)$ is the only network input, and $\mathbf{c}_{t,i}$ is injected into residual blocks via AdaGN. We train v_θ to match $\mathbf{u}_{p,i}$ using a Huber objective with threshold $\delta = 0.1$ and an endpoint-biased timestep distribution $p_\alpha(t)$, inspired by the difficulty-aware timestep reweighting principle in RF++ [13]:

$$\mathcal{L}_{\mathrm{RF}} = \mathbb{E}_{t\sim p_\alpha(t)}\left[\mathrm{Huber}_\delta(\mathbf{v}_{p,i}(t), \mathbf{u}_{p,i})\right]. \tag{9}$$

To sample t, we draw $r \sim \mathcal{U}(0,1)$ and apply

$$t = \frac{1}{2} - \frac{1}{2}\,\mathrm{sgn}\big(r - \tfrac{1}{2}\big)\;|2r-1|^{1/\alpha}, \tag{10}$$

where $\mathrm{sgn}(\cdot)$ is the sign function and α controls the degree of endpoint emphasis (we use $\alpha = 2.0$; see Sec. S8 in the supplementary material). Here, $p_\alpha(t)$ denotes the induced distribution of t under Eq. 10. This sampling allocates more probability mass to $t \approx 0$ and $t \approx 1$, where refinement is most sensitive in our residual regime. Finally, we zero-initialize the output layer of v_θ so that $\mathbf{v}_{p,i}(t) \approx \mathbf{0}$ at the beginning of training, making Stage 2 initially behave as an identity refinement that preserves $\mathbf{z}_{p,i}$ and then progressively learns residual high-frequency details.

3.5 CETA: Consistent Endpoint Trajectory Alignment Loss

CETA regularizes thickness-consistent reconstructions for the same underlying anatomy. For a proximal thickness pair (i, j) generated from the same HR patch $\mathbf{y}_p$, we set $T_j = T_i - \Delta T$ with $\Delta T = 1\,\mathrm{mm}$. Here, $\mathbf{z}_{p,k}$ denotes the Stage 1 APN output for condition k (Eq. (4)), for $k \in \{i, j\}$. We form endpoint proxies

$$\tilde{\mathbf{y}}_{p,k}(t) = \mathbf{z}_{p,k} + \mathbf{v}_{p,k}(t), \qquad k \in \{i, j\}, \tag{11}$$

where $\mathbf{v}_{p,k}(t)$ is the predicted rectified-flow velocity (Eq. (8)). CETA minimizes the squared L2 distance between proxies:

$$\mathcal{L}_{\mathrm{CETA}} = \|\tilde{\mathbf{y}}_{p,i}(t) - \tilde{\mathbf{y}}_{p,j}(t)\|_2^2. \tag{12}$$

Using a fixed proximal gap yields a chain of local constraints across the sampled thickness range, which propagates consistency beyond a single pair (*e.g.*, $T_6 \leftrightarrow T_5 \leftrightarrow \cdots \leftrightarrow T_1$). A very small ΔT yields nearly identical targets and provides weak alignment signal, while a very large ΔT forces alignment between substantially different trajectories and can destabilize training. We therefore use $\Delta T = 1\,\mathrm{mm}$ as a practical balance. The Stage 2 objective is $\mathcal{L}_{\mathrm{Stage2}} = \mathcal{L}_{\mathrm{RF}} + \lambda_{\mathrm{ceta}}\mathcal{L}_{\mathrm{CETA}}$ (we use $\lambda_{\mathrm{ceta}} = 1.0$. See Sec. S6 in the supplementary material).

3.6 PAD and AIS: Physics-Aware Adaptive Inference

Inference cost in Stage 2 is dominated by the number of velocity evaluations. Using a fixed step, therefore, wastes computation for thin-slice inputs and can under-refine thick-slice inputs. DRIFT addresses this by selecting the number of Euler steps N, equivalently the NFEs of the velocity network, directly from slice-thickness metadata.

Physics-Aware Difficulty (PAD). Given the input thickness T_i and the fixed target thickness T_{hr} (Sec. 3), thicker slices lose a larger fraction of through-plane frequency content due to slice-direction integration. We quantify this protocol difficulty by the normalized bandwidth deficit:

$$\mathrm{PAD}(T_i, T_{\mathrm{hr}}) = 1 - \frac{T_{\mathrm{hr}}}{T_i}, \tag{13}$$

where $\mathrm{PAD} \in [0, 1)$, equals 0 when $T_i = T_{\mathrm{hr}}$. PAD serves as a physics-aware degradation severity cue rather than an image difficulty score. It reflects the normalized through-plane bandwidth deficit given by $1 - T_{\mathrm{hr}}/T_i$, which increases as the input slice becomes thicker relative to the target resolution.

Adaptive Integration Scheduler (AIS). AIS maps $\mathrm{PAD}(T_i, T_{\mathrm{hr}})$ to the step budget:

$$N = \mathrm{clamp}\Big(\big\lfloor N_{\max} \cdot \mathrm{PAD}(T_i, T_{\mathrm{hr}}) \big\rceil,\ N_{\min},\ N_{\max}\Big), \tag{14}$$

where $\lfloor \cdot \rceil$ rounds to the nearest integer and $\mathrm{clamp}(x, a, b) = \min(\max(x, a), b)$. We set $N_{\min} = 0$ and $N_{\max} = 15$ (See Sec.S7 in the supplementary material). This allocation uses fewer NFEs for easier (thin-slice) cases and more NFEs for harder (thick-slice) cases, with negligible overhead since it depends only on slice-thickness metadata.

4 Experiments

4.1 Experimental Setups

Datasets. We evaluate on three public brain MRI datasets with isotropic HR volumes as ground truth, namely HCP [30], MIND [5], and IDEAS [29]. We use T1w and T2w for HCP and MIND, and T1w and FLAIR for IDEAS. All splits are subject-wise, and we use 890/223, 411/102, and 110/25 subjects for train/test on HCP, MIND, and IDEAS, respectively. We further perform zero-shot evaluation on real thick-slice MRI without isotropic ground truth using an IRB-approved in-house dataset and the public fastMRI brain dataset [37]. The in-house dataset includes T2w and FLAIR scans from 2 healthy volunteers, and the fastMRI subset includes 20 clinically acquired axial T2w volumes. We additionally use CC359 [26] only for zero-shot PAD scheduling analysis on multi-vendor T1w data. Additional dataset details are provided in the supplementary material (Sec. S2).

Evaluation. We evaluate all methods on reconstructed 3D volumes and present peak signal-to-noise ratio (PSNR) and SSIM [33]. For DRIFT and other 2D

baselines, we assemble each 3D volume by composing slice-wise predictions and compute PSNR over the reconstructed 3D volume. SSIM is computed on 2D slices and averaged over axial, coronal, and sagittal directions. For the IRB-approved in-house dataset, no isotropic ground truth is available, so we present qualitative results only.

Baselines. We compare DRIFT against fixed-scale SR methods [14, 16, 25, 36] and arbitrary-scale SR models [6, 12, 22, 32, 34]. All baselines are retrained per dataset using official codebases. Fixed-scale methods train one model per evaluated scale, while arbitrary-scale methods train a single model per dataset. To enable fair comparison under through-plane SR, we train the 2D restoration and SR models [16,36] and INR-based baselines [6,12,34] using the same SLR-based thick-slice degradation as DRIFT (Sec. 3.1). This ensures that the LR inputs exhibit the slice axis degradation patterns, including stair-step artifacts in reformatted views. For baselines whose formulations rely on dedicated specific-plane inputs [25,32], multi-view inputs [22], or inverse-problem operators [14], we follow their original input construction and preprocessing. Baseline-specific details are provided in the supplementary material (Sec. S3).

Implementation. We implement DRIFT in PyTorch and train on 2D 128×128 patches with a two-stage framework. Stage 1 (APN) is trained for 100 epochs with a batch size of 64 per GPU and lr 10^{-4}. In Stage 2, APN is frozen and v_θ is trained for 100 epochs with a batch size of 25 per GPU and lr 5×10^{-5}. We use AdamW with cosine annealing and mixed precision on 8×A6000 GPUs. During training, we synthesize thick-slice inputs using the SLR-based forward model (Sec. 3.1) and sample $T_i \sim \mathcal{U}(T_{\mathrm{hr}}, 6.0]\,\mathrm{mm}$. Training uses random slice and axis sampling, while evaluation uses fixed settings with a fixed seed for reproducibility. At inference, we apply sliding-window patch inference on each 2D slice with 32-pixel overlap and Gaussian blending, and AIS selects $N \in [0, 15]$ NFEs based on PAD (Sec. 3.6). Additional implementation details are provided in the supplementary material (Sec. S4).

4.2 Comparison against SOTA SR baselines on public datasets

Qualitative comparisons. Fig. 3 and Fig. S1 show visual comparisons on HCP, MIND, and IDEAS. For fixed-scale SR within the training range, DRIFT best preserves fine anatomical structures and textures while suppressing stair-step artifacts across datasets. SwinIR [16] tends to oversmooth details at larger thicknesses. AFCM [25] exhibits discontinuities along the slice-acquisition direction in non-axial views, which becomes more pronounced at higher scales on HCP and IDEAS. ResShift [36] loses fine structures at the largest scale. TPDM [14] reconstructs sharp details but often overestimates patterns, and it can leave residual noise in the output. The advantage of DRIFT becomes clearer for arbitrary-scale SR under continuous thickness changes, including the out-of-distribution thickness of 6.5 mm. Most arbitrary-scale baselines either blur anatomical details or fail to remove stair-step artifacts. In contrast, DRIFT produces outputs that are most consistent with the ground truth and maintains strong continuity along the slice-acquisition direction. Compared to 3D network ArSSR [34], DRIFT

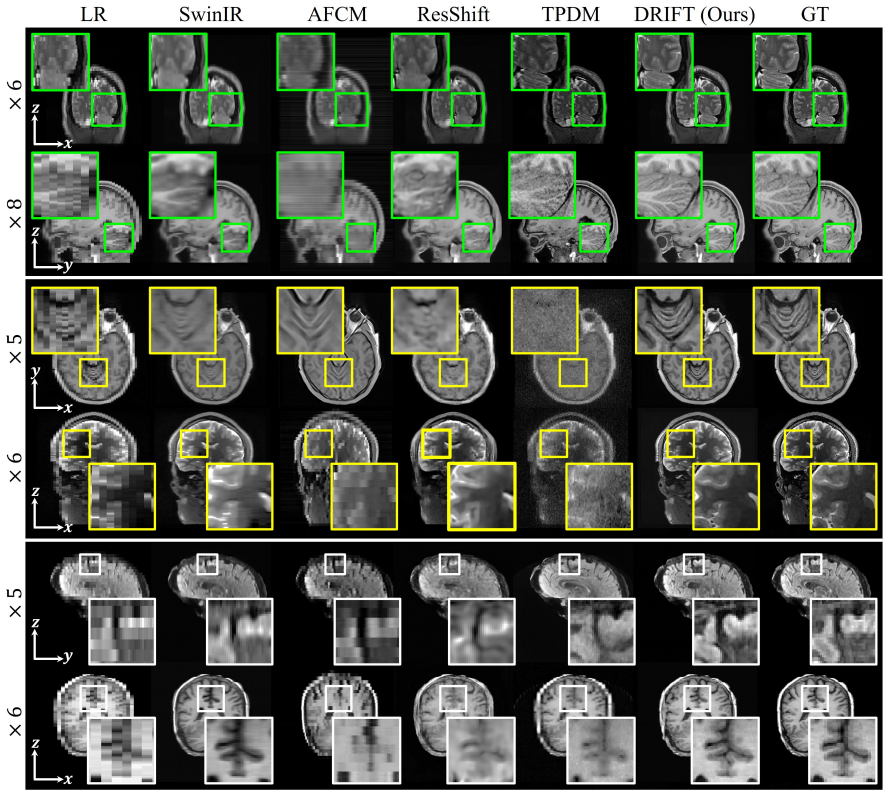

Fig. 3: Qualitative comparison with fixed-scale SR baselines on public datasets: HCP (green), MIND (yellow), and IDEAS (white). Results are shown at representative fixed scales within the training range, using HCP at ×6 (4.2 mm) and ×8 (5.6 mm), MIND at ×5 (4.5 mm) and ×6 (5.4 mm), and IDEAS at ×5 (5.0 mm) and ×6 (6.0 mm). Zoom-in boxes highlight regions of interest.

reduces inter-slice inconsistency in challenging regions where stair-step artifacts are prominent (Fig. S1 in the supplementary material). Although DRIFT is slice-wise, its deterministic APN initialization and shared RF refinement reduce visible discontinuities in reformatted views (Fig. S8). We further quantify this effect in the supplementary material using through-plane gradient RMSE, where DRIFT shows lower gradient error than TPDM, SA INR, and ArSSR on HCP ×8 (Table. S5).

Quantitative comparisons. Table 1 indicates PSNR and SSIM on HCP, MIND, and IDEAS under fixed-scale and arbitrary-scale settings, where DRIFT consistently achieves the best performance. For fixed-scale SR within the training range, DRIFT yields consistent gains across datasets, achieving 37.64/0.952 and 35.93/0.936 on HCP (×6, ×8), 32.74/0.904 and 30.02/0.886 on MIND (×5, ×6), and 32.75/0.946 and 32.01/0.938 on IDEAS (×5, ×6). Compared to the strongest fixed-scale baseline SwinIR [16], DRIFT improves PSNR by up to

Table 1: Quantitative comparison with state-of-the-art methods on HCP, MIND, and IDEAS for through-plane MRI SR. We report PSNR↑/SSIM↑ for fixed-scale and arbitrary-scale settings. The best results are highlighted in blue. Additional scales are in the supplementary.

Method	HCP		MIND		IDEAS	
Fixed-scale	×6 (4.2mm)	×8 (5.6mm)	×5 (4.5mm)	×6 (5.4mm)	×5 (5.0mm)	×6 (6.0mm)
SwinIR [16]	34.10/0.901	33.13/0.882	30.15/0.874	29.01/0.838	31.64/0.924	31.62/0.914
AFCM [25]	27.11/0.642	26.08/0.501	27.74/0.728	26.02/0.713	23.89/0.825	22.55/0.796
ResShift [36]	32.89/0.888	30.34/0.853	29.48/0.839	29.00/0.833	31.93/0.910	30.40/0.899
TPDM [14]	29.31/0.798	28.24/0.755	22.50/0.452	22.11/0.413	26.64/0.744	26.11/0.720
DRIFT (Ours)	**37.64/0.952**	**35.93/0.936**	**32.74/0.904**	**30.02/0.886**	**32.75/0.946**	**32.01/0.938**
Arbitrary-scale	×8.57 (6.0mm)	×9.28 (6.5mm)	×6.67 (6.0mm)	×7.22 (6.5mm)	×5.50 (5.5mm)	×6.50 (6.5mm)
SAINT [22]	25.09/0.686	24.77/0.666	23.39/0.721	23.05/0.701	24.63/0.649	24.33/0.627
LIIF [6]	26.60/0.801	26.53/0.782	25.27/0.774	25.11/0.773	27.24/0.850	26.98/0.842
LTE [12]	26.68/0.803	26.65/0.802	25.24/0.772	25.20/0.771	27.35/0.841	27.27/0.830
SA-INR [32]	25.86/0.694	25.18/0.665	26.22/0.810	25.12/0.774	26.18/0.745	26.10/0.735
ArSSR [34]	25.15/0.699	25.10/0.685	24.40/0.758	24.00/0.741	24.85/0.682	24.50/0.679
DRIFT (Ours)	**32.97/0.891**	**32.90/0.880**	**29.51/0.882**	**29.50/0.881**	**32.05/0.945**	**31.99/0.935**

3.54 dB (HCP), 2.59 dB (MIND), and 1.11 dB (IDEAS), with SSIM gains up to 0.054. TPDM [14] also appears sensitive to inference hyperparameters and can degrade noticeably across datasets even under the recommended setting of the authors. For arbitrary-scale SR, DRIFT shows larger advantages since many continuous-scale methods were originally developed under isotropic downsampling. Across all datasets, DRIFT outperforms the strongest arbitrary-scale baseline, typically LTE [12] or SA-INR [32], by 3.29–6.29 dB in PSNR at 6.0 mm and 4.71–4.72 dB at 6.5 mm, with large SSIM gains. Notably, DRIFT remains robust at the out-of-distribution thickness of 6.5 mm, while INR-based baselines such as LIIF [6], LTE [12], and ArSSR [34] show smaller improvements under stair-step artifacts.

4.3 Zero-Shot Evaluation on Real Thick-Slice MRI

We evaluate DRIFT on real thick-slice MRI without isotropic ground truth using an IRB-approved in-house dataset and the public fastMRI brain dataset. Because full-reference metrics are unavailable, the in-house dataset is used as a prospective qualitative case study across different sequences and acquisition grids, while fastMRI provides a larger public cohort for qualitative comparison and no-reference image-quality evaluation.

For the in-house scans, we use IDEAS-pretrained weights for FLAIR (1.0 mm in-plane) and HCP-pretrained weights for T2w (0.7 mm in-plane). As shown in Fig. 4, DRIFT reduces stair-step artifacts and preserves anatomical continuity in reformatted views across both sequences. Fixed-scale baselines cannot be evaluated on the T2w scan because models trained at the required scale are unavailable. Among arbitrary-scale baselines, several methods leave residual stair-step artifacts, and ArSSR [34] produces noticeably blurred reconstructions.

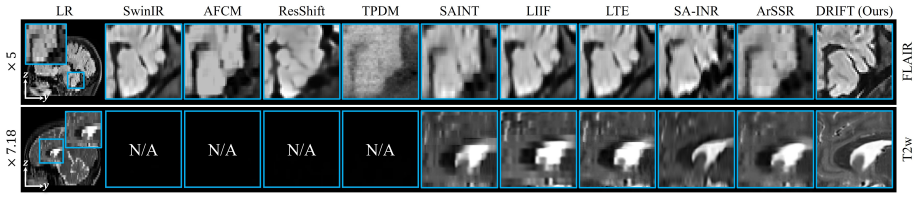

Fig. 4: Qualitative comparison on the in-house dataset (zero-shot). IRB-approved thick-slice scans (5 mm) for FLAIR and T2w are shown without isotropic ground truth. Models are applied without retraining using IDEAS-pretrained weights for FLAIR and HCP-pretrained weights for T2w. Zoom-in boxes highlight regions of interest.

TPDM [14] shows weaker cross-dataset generalization and fails to recover coherent structures on the in-house scans.

For fastMRI, we apply the HCP-pretrained DRIFT model to clinically acquired T2w scans with 5 mm slice thickness and compare against foundation model BME-X [27]. Qualitative comparisons show that DRIFT produces sharper coronal reformats with fewer through-plane artifacts than BME-X [27]. DRIFT also achieves higher sharpness (0.291 vs. 0.238) and lower NIQE and BRISQUE (5.30 and 21.16 vs. 7.22 and 62.46), supporting its zero-shot use on real thick-slice MRI without retraining or dataset-specific tuning (Fig. S9 and Table S6 in the supplementary material).

4.4 Efficiency–Fidelity Trade-off

In this section, we analyze an efficiency–fidelity trade-off between reconstruction fidelity and inference cost from two complementary perspectives and further report standardized model cost in the supplementary material.

Perceptual efficiency–fidelity (runtime vs. LPIPS). Fig. 5 indicates per-volume runtime against LPIPS on HCP ×8 (5.6 mm→0.7 mm). Most feed-forward baselines [6,12,16,25] are fast but show higher LPIPS due to oversmoothing. Several volumetric or coordinate-based feed-forward models, including SAINT [22], SA-INR [32], and ArSSR [34], can be slower than DRIFT in this setting. DRIFT achieves the lowest LPIPS (0.043) at about 20 s per volume, improving over ArSSR [34] (0.112) by ×3.3 while being ×14 faster. TPDM [14] is about ×450 slower yet yields ×6 worse LPIPS than DRIFT.

Distortion efficiency–fidelity (NFEs vs. PSNR). Fig. 6 shows PSNR versus NFEs against multi-step diffusion baselines on HCP. ResShift [36] runs with a small fixed step count (15 NFEs), which does not adapt to thickness-dependent difficulty. TPDM [14] requires 2,000 steps while producing lower PSNR. In contrast, DRIFT achieves higher PSNR with fewer evaluations, using 8, 12, and 13 NFEs at ×2, ×6, and ×8, respectively. Overall, DRIFT is slice-thickness-aware and adaptively allocates NFEs, achieving higher PSNR than fixed-step diffusion baselines while using fewer function evaluations.

Standardized model cost. We further report standardized cost and compact variant results in the supplementary material (Sec. S4.2). A smaller DRIFT-

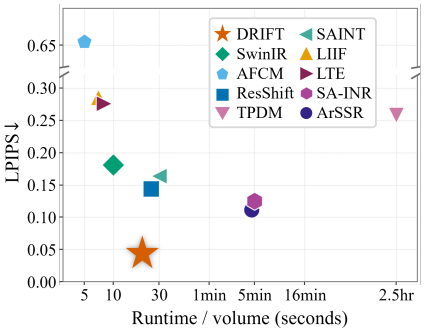

Fig. 5: Perceptual efficiency-fidelity trade-off on HCP ×8 (5.6 mm). Per-volume runtime versus LPIPS.

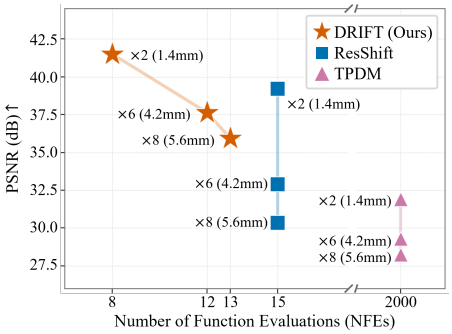

Fig. 6: Distortion efficiency-fidelity trade-off against multi-step diffusion baselines on HCP. PSNR versus NFEs at ×2 (1.4 mm), ×6 (4.2 mm), and ×8 (5.6 mm).

Small variant keeps the same two-stage design and adaptive integration mechanism while reducing parameters, FLOPs, and peak VRAM by ×3.3, ×4.0, and ×5.3 relative to DRIFT-Large. Its performance remains competitive across HCP, MIND, and IDEAS, indicating that the efficiency and fidelity behavior of DRIFT is not explained by model size alone.

4.5 Ablation Study

PAD versus Image-Dependent Difficulty. On multi-vendor CC359 [26], an AdaDiffSR-style image score shows negligible correlation with HF-PSNR within fixed input-thickness groups. Image-only and hybrid scheduling also underperform PAD, supporting slice-thickness-based scheduling in this controlled through-plane SR setting (Sec. S11).

Ablation on the Two-Stage Design. We compare Stage 1 only, Stage 2 only, and full DRIFT. Stage 1 removes coarse stair-step artifacts but oversmooths fine detail, while Stage 2 without APN improves local sharpness but lacks stable anatomical initialization at larger slice thicknesses. Full DRIFT combines anatomical initialization with residual rectified-flow refinement and gives the best quantitative and qualitative results (Fig. S2).

Ablation on Each Stage 2 Component. Table 2 quantifies the contribution of each Stage 2 component on HCP. Removing slice-thickness conditioning causes the largest degradation, showing that acquisition conditioning is essential for heterogeneous thick-slice inputs. CETA and U-shaped timestep sampling further improve the result, indicating complementary roles of trajectory regularization and timestep prioritization.

Ablation on the CETA Proximal Gap. Table 3 evaluates the gap ΔT used to form CETA loss pairs. Random CETA pairs improve over the no CETA setting in Table 2, but remain below the fixed 1.0 mm gap despite moderate coverage, indicating that pair coverage alone is not sufficient. Increasing ΔT can provide a stronger cross-thickness constraint, but it also reduces coverage

Table 2: Stage 2 component ablation on HCP. We remove one component while keeping the Stage 1 APN fixed. Metrics are averaged over $T_{\mathrm{i}} \in \{1.5, 3.0, 5.0\}$ mm with $T_{\mathrm{hr}} = 0.7$ mm. CETA is N/A without thickness conditioning.

Thickness Cond.	U-shaped Sampling	CETA loss ($\lambda_{\mathrm{CETA}} = 1$)	PSNR↑	SSIM↑	LPIPS↓
✓	✓	✗	36.92	0.9416	0.0308
✗	✓	N/A	34.82	0.8986	0.0491
✓	✗	✓	38.12	0.9458	0.0342
✓	✓	✓	**40.85**	**0.9634**	**0.0229**

Table 3: CETA gap ablation on HCP. Random denotes naive random CETA pairs, and Coverage denotes the supervised fraction of the thickness range. Metrics are averaged over $T_{\mathrm{i}} \in \{1.5, 3.0, 5.0\}$ mm with $T_{\mathrm{hr}} = 0.7$ mm.

Gap ΔT	Coverage	PSNR↑	SSIM↑	LPIPS↓
Random	65%	38.23	0.9460	0.0290
0.5 mm	91%	37.24	0.9430	0.0301
1.0 mm	**81%**	**40.85**	**0.9634**	**0.0229**
2.0 mm	62%	39.10	0.9540	0.0265
4.0 mm	25%	37.58	0.9445	0.0338

within the training range and aligns more dissimilar conditions, which may make optimization less stable. Conversely, a small gap such as 0.5 mm yields high coverage but provides a weaker learning signal because paired conditions become nearly indistinguishable, making the alignment close to an identity constraint. Across random and fixed-gap settings, $\Delta T = 1.0$ mm provides the best trade-off between constraint strength, locality, and coverage, achieving the best PSNR, SSIM, and LPIPS.

Additional evaluations. Additional downstream segmentation results and multi-planar volumetric visualizations are provided in the supplementary material to assess anatomical utility and volumetric reformat quality. (Sec. S14 and Sec. S15).

5 Conclusion

We propose DRIFT, a two-stage difficulty-aware rectified flow framework for through-plane MRI SR, with thickness conditioning for continuous input slice thicknesses. APN projects thick-slice inputs onto a coarse HR manifold, reducing the residual transport solved by Stage 2. PAD-guided AIS allocates ODE steps according to slice-thickness difficulty, while CETA regularizes proximal thickness trajectories for consistent reconstructions. Across public datasets and real thick-slice MRI, DRIFT improves fidelity and sampling efficiency over regression, INR, and diffusion baselines. A current limitation is that the main experiments use fixed native target thicknesses. Extending target-thickness sampling across resolutions and anatomies remains an important direction for cross-protocol generalization.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2025-00561616).

References

1. Benisty, R., Shteynman, Y., Porat, M., Ilivitzki, A., Freiman, M.: Simple: Simultaneous multi-plane self-supervised learning for isotropic mri restoration from anisotropic data. In: Int. Conf. Medical Image Computing and Computer-Assisted Intervention (2025)
2. Bernstein, M.A., King, K.F., Zhou, X.J.: Handbook of MRI pulse sequences. Elsevier (2004)
3. Cao, J., Wang, Q., Xian, Y., Li, Y., Ni, B., Pi, Z., Zhang, K., Zhang, Y., Timofte, R., Van Gool, L.: Ciaosr: Continuous implicit attention-in-attention network for arbitrary-scale image super-resolution. In: IEEE Conf. Comput. Vis. Pattern Recog. (2023)
4. Charbonnier, P., Blanc-Feraud, L., Aubert, G., Barlaud, M.: Two deterministic half-quadratic regularization algorithms for computed imaging. In: IEEE Int. Conf. Image Process. (1994)
5. Chen, X.J., Salvadore, M., Blanco-Elorrieta, E.: MIND: Multilingual imaging neuro dataset. OpenNeuro dataset (2025). `https://doi.org/10.18112/openneuro.ds006391.v2.0.0`, version 2.0.0, Accessed 1 July 2026
6. Chen, Y., Liu, S., Wang, X.: Learning continuous image representation with local implicit image function. In: IEEE Conf. Comput. Vis. Pattern Recog. (2021)
7. Dhariwal, P., Nichol, A.: Diffusion models beat gans on image synthesis. Advances in neural information processing systems (2021)
8. Fan, Y., Liu, C., Yin, N., Gao, C., Qian, X.: Adadiffsr: Adaptive region-aware dynamic acceleration diffusion model for real-world image super-resolution. In: Eur. Conf. Comput. Vis. (2024)
9. Hu, X., Mu, H., Zhang, X., Wang, Z., Tan, T., Sun, J.: Meta-sr: A magnification-arbitrary network for super-resolution. In: IEEE Conf. Comput. Vis. Pattern Recog. (2019)
10. Iglesias, J.E., Billot, B., Balbastre, Y., Magdamo, C., Arnold, S.E., Das, S., Edlow, B.L., Alexander, D.C., Golland, P., Fischl, B.: Synthsr: A public ai tool to turn heterogeneous clinical brain scans into high-resolution t1-weighted images for 3d morphometry. Science advances (2023)
11. Iglesias, J.E., Billot, B., Balbastre, Y., Tabari, A., Conklin, J., González, R.G., Alexander, D.C., Golland, P., Edlow, B.L., Fischl, B., et al.: Joint super-resolution and synthesis of 1 mm isotropic mp-rage volumes from clinical mri exams with scans of different orientation, resolution and contrast. Neuroimage (2021)
12. Lee, J., Jin, K.H.: Local texture estimator for implicit representation function. In: IEEE Conf. Comput. Vis. Pattern Recog. (2022)
13. Lee, S., Lin, Z., Fanti, G.: Improving the training of rectified flows. Adv. Neural Inform. Process. Syst. (2024)
14. Lee, S., Chung, H., Park, M., Park, J., Ryu, W.S., Ye, J.C.: Improving 3d imaging with pre-trained perpendicular 2d diffusion models. In: Int. Conf. Comput. Vis. pp. 10710–10720 (2023)
15. Li, H., Yang, Y., Chang, M., Chen, S., Feng, H., Xu, Z., Li, Q., Chen, Y.: Srdiff: Single image super-resolution with diffusion probabilistic models. Neurocomputing (2022)
16. Liang, J., Cao, J., Sun, G., Zhang, K., Van Gool, L., Timofte, R.: Swinir: Image restoration using swin transformer. In: Int. Conf. Comput. Vis. (2021)
17. Lipman, Y., Chen, R.T.Q., Ben-Hamu, H., Nickel, M., Le, M.: Flow matching for generative modeling. In: Int. Conf. Learn. Represent. (2023)

18. Liu, X., Gong, C., qiang liu: Flow straight and fast: Learning to generate and transfer data with rectified flow. In: Int. Conf. Learn. Represent. (2023)
19. Liu, X., Zhang, X., Ma, J., Peng, J., et al.: Instaflow: One step is enough for high-quality diffusion-based text-to-image generation. In: Int. Conf. Learn. Represent. (2023)
20. Ohayon, G., Michaeli, T., Elad, M.: Posterior-mean rectified flow: Towards minimum MSE photo-realistic image restoration. In: Int. Conf. Learn. Represent. (2025)
21. Pauly, J., Le Roux, P., Nishimura, D., Macovski, A.: Parameter relations for the shinnar-le roux selective excitation pulse design algorithm (nmr imaging). IEEE Transactions on Medical Imaging (1991)
22. Peng, C., Lin, W.A., Liao, H., Chellappa, R., Zhou, S.K.: Saint: Spatially aware interpolation network for medical slice synthesis. In: IEEE Conf. Comput. Vis. Pattern Recog. pp. 7750–7759 (2020)
23. Rahaman, N., Baratin, A., Arpit, D., Draxler, F., Lin, M., Hamprecht, F., Bengio, Y., Courville, A.: On the spectral bias of neural networks. In: Int. Conf. Learn. Represent. (2019)
24. Saharia, C., Ho, J., Chan, W., Salimans, T., Fleet, D.J., Norouzi, M.: Image super-resolution via iterative refinement. IEEE Trans. Pattern Anal. Mach. Intell. (2022)
25. Song, Z., Wang, X., Zhao, X., Wang, S., Shen, Z., Zhuang, Z., Liu, M., Wang, Q., Zhang, L.: Alias-free co-modulated network for cross-modality synthesis and super-resolution of mr images. In: Int. Conf. Medical Image Computing and Computer-Assisted Intervention (2023)
26. Souza, R., Lucena, O., Garrafa, J., Gobbi, D., Saluzzi, M., Appenzeller, S., Rittner, L., Frayne, R., Lotufo, R.: An open, multi-vendor, multi-field-strength brain mr dataset and analysis of publicly available skull stripping methods agreement. NeuroImage (2018)
27. Sun, Y., Wang, L., Li, G., Lin, W., Wang, L.: A foundation model for enhancing magnetic resonance images and downstream segmentation, registration and diagnostic tasks. Nature Biomedical Engineering (2025)
28. Tai, Y., Xie, R., Zhao, C., Zhang, K., Zhang, Z., Zhou, J., Yang, J.: Addsr: Accelerating diffusion-based blind super-resolution with adversarial diffusion distillation. Pattern Recognition (2026)
29. Taylor, P.N., Wang, Y., Simpson, C., Janiukstyte, V., Horsley, J., Leiberg, K., Little, B., Clifford, H., Adler, S., Vos, S.B., et al.: The imaging database for epilepsy and surgery (ideas). Epilepsia (2025)
30. Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., Consortium, W.M.H., et al.: The wu-minn human connectome project: an overview. Neuroimage (2013)
31. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. Adv. Neural Inform. Process. Syst. (2017)
32. Wang, X., Wang, S., Xiong, H., Xuan, K., Zhuang, Z., Liu, M., Shen, Z., Zhao, X., Zhang, L., Wang, Q.: Spatial attention-based implicit neural representation for arbitrary reduction of mri slice spacing. Medical Image Analysis (2024)
33. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. (2004)
34. Wu, Q., Li, Y., Sun, Y., Zhou, Y., Wei, H., Yu, J., Zhang, Y.: An arbitrary scale super-resolution approach for 3d mr images via implicit neural representation. IEEE Journal of Biomedical and Health Informatics (2022)

35. Xu, J., Li, W., Sun, H., Li, F., Wang, Z., Peng, L., Ren, J., Yang, H., Hu, X., Pei, R., Heng, P.A.: Fast image super-resolution via consistency rectified flow. In: Int. Conf. Comput. Vis. (2025)
36. Yue, Z., Wang, J., Loy, C.C.: Resshift: Efficient diffusion model for image super-resolution by residual shifting. Adv. Neural Inform. Process. Syst. (2023)
37. Zbontar, J., Knoll, F., Sriram, A., Murrell, T., Huang, Z., Muckley, M.J., Defazio, A., Stern, R., Johnson, P., Bruno, M., Parente, M., Geras, K.J., Katsnelson, J., Chandarana, H., Zhang, Z., Drozdzal, M., Romero, A., Rabbat, M., Vincent, P., Yakubova, N., Pinkerton, J., Wang, D., Owens, E., Zitnick, C.L., Recht, M.P., Sodickson, D.K., Lui, Y.W.: fastMRI: An open dataset and benchmarks for accelerated MRI (2018)
38. Zhao, C., Dewey, B.E., Pham, D.L., Calabresi, P.A., Reich, D.S., Prince, J.L.: Smore: a self-supervised anti-aliasing and super-resolution algorithm for mri using deep learning. IEEE Transactions on Medical Imaging (2020)