MArFE: Multi-Contrast MRI Arbitrary Scale Super-Resolution with Fourier Enhancement

Zhiwen Shi[1]

Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China zhiwenshi1025@gmail.com

Abstract. High-resolution multi-contrast MRI is essential for capturing rich anatomical structures and supports more accurate disease diagnosis through the reference images from another contrast. It often suffers from limited spatial resolution due to acquisition constraints, and requests exploring arbitrary-scale reconstruction beyond fixed scale factors. However, related research exists two following issues: 1) Implicit neural representations (INR) as mainstream methods are prone to spectral bias, which limits their ability to recover high-frequency details; 2) Multi-contrast MRI is often used as the effective prior, but lacks the targeted network design to further merge INR positional information. To solve these problems, we propose a Fourier-enhanced implicit framework for arbitrary-scale multi-contrast MRI super-resolution (MArFE). First, the Fourier reparameterization module (FRM) is introduced to characterize frequency features through fixed Fourier basis decomposition. Then, Fourier neural operators can alleviate frequency degradation by the non-local Galerkin-type linear attention (NLGLA) based on kernel integral mechanisms. Additionally, we design a hybrid loss that jointly supervises both spatial and frequency domains to improve high-frequency texture reconstruction. Extensive experiments on two public multi-contrast MRI datasets demonstrate that MArFE outperforms existing methods under various scale factors with superior visual performances and quantitative metrics. These results highlight the clinical potential of MArFE in enhancing MRI quality across multi-contrast sequences. The project code can be found at https://github.com/zhiwen-shi/MArFE.

Keywords: MRI Super-resolution · Implicit neural representation · Fourier spectrum · Multi-contrast · Arbitrary scale

1 Introduction

Magnetic Resonance Imaging (MRI) is a widely used medical imaging technique based on the principles of nuclear magnetic resonance (NMR). It acquires internal anatomical structures by detecting the intensity of electromagnetic signals emitted under externally applied gradient magnetic fields. MRI is particularly effective in the detection, diagnosis, and evaluation of cardiovascular diseases, cerebrovascular conditions, and disorders of thoracic and abdominal organs. It is

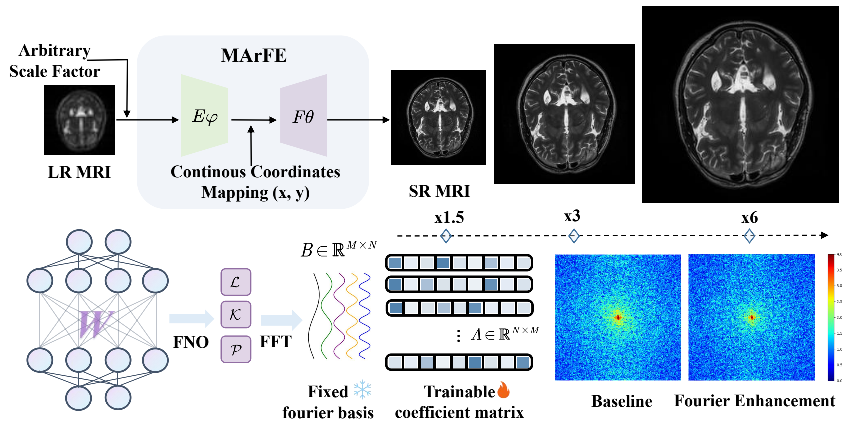

Fig. 1: A conceptual illustration of ASSR framework. The Fourier operator comprises three stages: Lifting (L), Kernel Integration (K), and Projection (P). The weight W is factorized into a fixed Fourier basis matrix B and a learnable coefficient matrix Λ. Frequency error maps under $\times 4$ MRI SR demonstrate the effectiveness of MArFE.

also extensively used in sports injury diagnostics and is especially suited for imaging the spine and joints, providing clear visualization of soft tissues near bones, such as ligaments and muscles [14]. In clinical diagnosis, conventional MRI protocols typically include multi-contrast scan sequences such as T1-weighted, T2-weighted, FLAIR, and PD-weighted images. These multi-contrast images provide complementary information.

Due to the inherent trade-off between medical image resolution, signal-to-noise ratio (SNR), and MRI scan time, acquiring high-resolution (HR) MR images remains a challenge in practical clinical settings. A common and effective strategy is to first acquire low-resolution (LR) MR images with high SNR and then apply super-resolution (SR) algorithms to reconstruct HR outputs [6,43,56]. Deep learning-based SR approaches learn the distribution of training data to model a mapping from LR images to HR images, thereby supplementing the LR inputs with high-frequency (HF) details and improving reconstruction quality. Traditional single image super-resolution (SISR) methods often suffer from a critical limitation. They are typically trained at fixed upsampling factors [9, 11, 49]. When dealing with different anatomical structures or sequence types, the model can only perform end-to-end SR at a specific scale, limiting its flexibility across varying resolutions [61].

Implicit Neural Representation (INR) methods offer a promising solution for the problem above with the arbitrary-scale super-resolution (ASSR). INR approaches typically reconstruct images by learning continuous implicit functions, which allows for flexible resolution generation and more precise modeling of complex anatomical structures in MRI data. These networks learn continuous mappings from spatial coordinates to signal intensities, enabling arbitrary-resolution inference by querying the function at any desired location [59]. Despite their advantages, as shown in Fig. 1, existing INR-based SR methods still have some drawbacks that need to be considered. Firstly, INR methods always suffer from an inherent spectrum bias, a tendency to prioritize learning low-frequency (LF) components while forgetting the high-frequency details. This often results in blurry reconstructions without rich textures [45,63]. Furthermore, MRI data are fundamentally acquired and reconstructed in the frequency domain (K-space), where the signals are represented as complex-valued data. Most existing SR methods are developed based on natural image priors in the spatial domain and largely overlook the specific properties and potential of frequency-domain modeling in MRI reconstruction. In addition, within the INR framework of continuous coordinate mapping, multi-contrast features are only operated by simply global concatenation [17, 38, 57]. The structure of the implicit decoder lacks dense fusions of multi-contrast features and relative positional coordinates from the spatial domain, thereby failing to fully leverage corresponding features of different contracts and the correlation between the location and pixel spaces. As illustrated in Fig. 1, existing ASSR methods struggle to recover high-frequency details in LR images, while our proposed MArFE significantly polishes the frequency performance, producing clearer medical textures and reduced frequency error in the process of reconstruction.

To address the above limitations of INR-based SR and fully exploit the multi-contrast MRI sequences, we propose MArFE, a novel Fourier-enhanced implicit framework for arbitrary-scale multi-contrast MRI super-resolution. Specifically, we incorporate INR into a dual-branch architecture that leverages complementary contrast information and extends the concept of ASSR into the reference image domain. Furthermore, we consider a frequency-domain modeling method tailored for complex-valued MRI signals, aiming to overcome the high-frequency deficiency of vanilla INR. In summary, our main contributions are as follows:

- We first introduce the Fourier Reparameterization Module (FRM) that enhances frequency sensitivity by transforming features in the Fourier domain, combined with sinusoidal activation to explicitly model the spectral contents.
- We first incorporate the Fourier Neural Operator in the INR-based MRI SR framework by introducing a Non-Local Galerkin-type Linear Attention (NLGLA), with coordinate-feature coupling via linear projection.
- We improve the contrast-aware dense implicit decoder structure in the multi-contrast INR by introducing efficient designs for feature recurrence and residual enhancement, boosting the reconstruction performance.
- We propose dual-domain loss functions, consisting of Global k-space Loss and High-Frequency Loss, which jointly supervise both spatial and frequency domains, enhancing the high-frequency MRI details.

2 Related Work

2.1 Single Image Super-Resolution

Single image super-resolution (SISR) is a fundamental low-level task in computer vision. Traditional SISR approaches mainly include interpolation-based and regularization-based methods. The most common interpolation techniques, such as bilinear and bicubic interpolation, are widely used in industrial applications due to their simplicity and efficiency. However, they often introduce severe blocking artifacts [2, 64]. Regularization-based methods aim to recover HR images by solving inverse problems with physical priors, typically relying on handcrafted or idealized constraints, which significantly limit the performance of such models [47, 49].

Deep learning has shown remarkable potential in SISR. The seminal work of SRCNN pioneered the use of end-to-end convolutional neural networks (CNNs) to learn the mapping between LR and HR images from large-scale datasets, greatly outperforming traditional interpolation and regularization methods [10]. Subsequent studies further advanced the field with more complex network structures and diversified training methods [22–25, 30, 34, 35, 51, 69]. VDSR explored deeper network architectures [22], and SRGAN introduced perceptual loss by combining adversarial and content loss, achieving more realistic visual quality at $\times 4$ upsampling [25]. These developments opened up broader research directions for SISR.

With the success of deep learning in natural image SR, many models have been extended to medical imaging. Given the diversity of MRI sequences, recent studies have focused on multi-contrast MRI super-resolution with deep CNNs [32, 52, 67]. Generative adversarial network (GAN) have also been attempted, and multi-contrast images can be concatenated and fed into GAN-based frameworks to reconstruct HR MRI [8, 18, 31, 33]. Multi-scale fusion techniques, such as Transformer mechanisms, multi-level feature aggregation, and contextual fusion strategies, have also been explored to better image recovery [13, 27, 28]. However, these models typically treat different scale factors as independent tasks, requiring separate training and storage, which limits their flexibility. Motivated by these limitations, we began to explore a multi-contrast MRI ASSR method for medical applications.

2.2 Implicit Neural Representation

Implicit Neural Representation (INR) represents signals in continuous domains and have found wide applications in 3D reconstruction [7, 20, 36, 41], novel view synthesis [37, 42, 44], and image&video compression [4, 53, 70]. Notably, NeRF represents 3D scenes with continuous neural functions and leverages positional encoding to map coordinates into high-dimensional space for view synthesis [37]. INR models map spatial coordinates to signal values, and their complexity depends on signal content rather than resolution. The key challenge of INR is its spectral bias, which makes it harder to capture high-frequency details [45]. To mitigate this issue, the inputs can be projected into higher-dimensional spaces via positional encoding [55], and use alternative activation functions (e.g., sinusoidal functions in SIREN). Sinusoidal Positional Encoding has also been shown to alleviate spatial bias in image generation [62]. Altering the decoder output to a higher-dimensional space can also be explored to cover distinct signal components [39]. INR has also been adapted to medical imaging tasks such as reconstruction, segmentation, registration, novel view synthesis, and compression [46,48,54]. However, INR-based MRI rarely consider the frequency domain, even though MRI data is inherently acquired in k-space. Our work addresses this gap by incorporating Fourier basis decomposition into the INR framework, enabling more effective frequency modeling for MRI arbitrary-scale SR.

2.3 Arbitrary-Scale Super-Resolution

Arbitrary-scale super-resolution (ASSR) has gained increasing attention for its scalability across various upsampling factors, with the theoretical support of INR. Recent works have demonstrated the potential of INR for this task. In the field of natural images, Meta-SR introduced a meta-upscale module to dynamically predict HR pixels from LR neighborhoods across arbitrary scales [16]. LIIF proposed a local implicit image function to model pixel-wise mappings from coordinates to RGB values, allowing HR image reconstruction at arbitrary resolutions from a single LR input [5]. Subsequent works enhanced this paradigm. LTE utilized Fourier-domain features to define implicit functions [26], SRNO adopted neural operators via kernel integration [58], and LINF modeled local textures using normalized flows [65]. LMF introduced a latent modulation function to compress features, trading off accuracy for rendering speed [15]. CiaoSR designed a scale-aware attention mechanism to expand the receptive field [3], and HIIF integrated Fourier neural operators with multi-scale coordinate mapping for local feature adaptation [21]. In medical imaging, ArSSR extended LIIF to 3D MRI by predicting voxel intensities from spatial coordinates and LR input [60]. CycleINR emphasized inter-slice consistency for volumetric MRI [12], while Dual-Arb tackled multi-contrast MRI with a dual-branch architecture and positional encoding for enhanced image recovery performance [68]. Our work further explores the integration of efficient linear neural operators and multi-contrast convolutional frameworks within the complex domain of MRI, aiming to improve the ASSR medical tasks.

3 Model Architecture

3.1 Overview

The overall architecture of MArFE is illustrated in Fig. 2. Our network simultaneously takes the paired target LR $\in H^{tar} \times W^{tar}$ and the reference LR/HR $\in H^{ref} \times W^{ref}$ of arbitrary resolution as complex-valued MRI inputs. The reference image is not constrained by size or sequence types, allowing for flexible

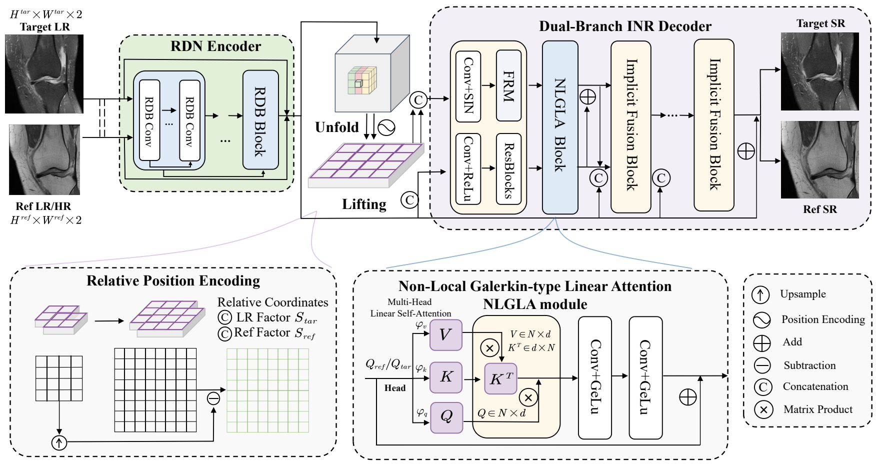

Fig. 2: The main pipeline of MArFE. MArFE propose a dual-branch super-resolution framework in the complex domain, processing arbitrary-resolution target and reference images simultaneously. In the encoder stage, both branches share the same RDN-based feature extractor. Continuous coordinate mapping is performed via relative positional encoding, and then concatenated into the INR decoder, with Fourier reparameterization and multi-head NLGLA to polish high-quality ASSR.

multi-contrast information. The Implicit Neural Representation (INR) can represent the image in the latent space $F \in R^{H \times W \times C}$, and use the Implicit Decoding Function (IDF) $F(\cdot)$ to decode the latent signal closest to itself.

We design a dual-branch super-resolution framework in the complex domain with two separate input and output channels, where the encoder shares the same weights of the Residual Dense Network (RDN) [69] for feature extraction, and then the MRI features are injected into the implicit decoder with the image and position branches. Considering that the two input images have different resolutions, we use the nearest interpolation to align the feature maps after the RDN encoder by two different scale factors S_{ref} and S_{tar}:

$$F_{z\uparrow} = Upsample\left(RDN\left(I_Z\right), S_z\right) \tag{1}$$

where $z \in \{tar, ref\}$ represents the two input images, I_{tar} and I_{ref} are the input target LR and reference image. The latent features are aligned to a uniform size, enabling their subsequent processing by the INR decoder. Coordinate mapping is performed via relative positional encoding, and then concatenated with features into the decoder. The encoded features are transformed with the Fourier Reparameterization Module (FRM) and further refined by a multi-head Non-local Galerkin-type Linear Attention (NLGLA). It is worth noting that we further enhance the reconstruction quality by joint spatial and frequency losses designed for MRI data. Through the continuous implicit fusion modules, MArFE enables high-quality arbitrary-scale SR for multi-contrast MRI.

3.2 Fourier Reparameterization

While INR-based methods have shown promise in image super-resolution, they still suffer from inherent limitations, notably the spectral bias toward low-frequency components. As a type of medical imaging, MR images typically contain more homogeneous and structured content compared to natural images. This characteristic makes frequency-domain modeling particularly effective in representing MR image information [50]. To address this, we incorporate a Fourier reparameterization module in the MRI recovery, where MR images are decomposed into multiple groups of fixed Fourier orthogonal bases, which are subsequently integrated into the high-dimensional mapping of the INR decoder. Fourier series represent functions as weighted sums of *sin* and *cos* functions, which are orthogonal and together form a complete basis capable of reconstructing any periodic function. A 2D image can be considered a bounded periodic function and decomposed using the Fourier series. For a one-dimensional discrete function $f(x)$ with period 1, the Fourier expansion is defined as:

$$f(x) = a_0 + \sum_{n \in N} a_n \sin(2\pi n x) + b_n (2\pi n x) \tag{2}$$

where N is the number of sampling points of $f(x)$. By converting it into a vector, it can be further expressed as:

$$f(x) = a_0 + WB^T \tag{3}$$

The form of Fourier reparameterization is similar to the positional encoding proposed in previous works. In fact, positional encoding can be regarded as an optimized form based on Fourier bases. A group of the fixed Fourier basis decomposition was specifically designed for image characteristics. MRI possesses a more concentrated frequency domain representation, making it easier to select relevant frequencies for representation to achieve more effective reconstruction.

3.3 Non-Local Galerkin-type Linear Attention

To further alleviate the problems of spectral bias, more Fourier methods are being considered. The Fourier Neural Operator (FNO) is an efficient NO variant, designed to learn continuous mappings between infinite-dimensional spaces [29]. It directly operates on inputs of arbitrary resolution, enabling super-resolution by mapping low-resolution inputs to high-resolution outputs. We first integrate the FNO in the INR-based MRI SR framework by introducing a Non-Local Galerkin-type Linear Attention (NLGLA), with coordinate-feature coupling via linear projection. The attention method transforms input data into the frequency domain, applies learned filters, and then transforms the data back into the spatial domain to refine the features. Due to its inherent frequency-domain operations, NLGLA with the FNO mechanism aligns well with the requirements of frequency-domain enhancement in MRI reconstruction tasks. I_{LR} represents the LR image as the input, where $r \in \mathbb{R}^2$ denotes the spatial coordinates. Our goal is to learn an operator $\mathcal{G}$ such that:

$$I_{HR}(r) = \mathcal{G}[I_{LR}(r)] \tag{4}$$

where $I_{HR}(r)$ is the output. The FNO models $\mathcal{G}$ can be achieved by stacking:

$$f_{l+1}(r) = \sigma(\mathcal{W}_l f_l(r) + \mathcal{K}_l f_l(r)) \tag{5}$$

where f_l is the feature at layer l evaluated at spatial location r, and $f_{l+1}(r)$ is its updated representation in the following layer. σ is a nonlinear activation function, and $\mathcal{W}_l$ is a linear transformation. $\mathcal{K}_l$, the integral operator, transforms the features into the Fourier domain. Fourier modes are then truncated for computational efficiency. Global convolution is performed with a pointwise multiplication between the transformed features and the learned Fourier coefficients.

$$\mathcal{K}_l f_l(r) = \mathcal{F}^{-1}(P_l(\xi) \cdot \mathcal{F}[f_l](\xi)) \tag{6}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ represent the Fourier and inverse Fourier transforms, respectively. ξ is the frequency domain variables. $\mathcal{F}[f_l](\xi)$ is the Fourier transform of f_l , evaluated at frequency ξ. $P_l(\xi)$ is a complex-valued tensor of learnable parameters denoting the Fourier domain filters. In practical INR implementations, the above FNO process can be reformulated as a linear attention mechanism based on Galerkin projection:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \phi(\mathbf{Q})\left(\phi(\mathbf{K})^\top \mathbf{V}\right) \tag{7}$$

where $\mathbf{Q} \in \mathbb{R}^{N\times d}$, $\mathbf{K} \in \mathbb{R}^{N\times d}$, and $\mathbf{V} \in \mathbb{R}^{N\times d_v}$ denote the query, key, and value matrices, respectively. N is the number of tokens (or spatial positions), d is the feature dimension of Q and K, and d_v is the value dimension. The mapping $\phi(\cdot)$ is a kernel feature function that ensures the inner products are positive and enables the linearization. By reformulating the standard quadratic attention into a kernelized linear form, the proposed mechanism reduces computational complexity from $\mathcal{O}(N^2)$ to $\mathcal{O}(N)$, while preserving global context modeling capability. This allows efficient capture of long-range spatial dependencies and non-local spectral interactions, which are critical for maintaining structural consistency and high-frequency fidelity in multi-contrast MRI reconstruction.

3.4 Loss Function

To capture the frequency-domain characteristics in K-space, we propose a Global K-space Loss (GKLoss) that replaces traditional high-pass filters, aiming to preserve both low-frequency (LF) structural information and high-frequency (HF) edge details. In K-space, the magnitude represents the energy distribution across different frequencies—higher magnitudes correspond to stronger signals that contribute significantly to image reconstruction. To leverage this property, GKLoss incorporates magnitude-aware and frequency-position-aware weighting schemes.

$$
\begin{aligned}
d(i,j) &= \sqrt{(i-h/2)^2+(j-w/2)^2} \\
w(i,j) &= e^{\alpha}\cdot\frac{d(i,j)}{\max d(i,j)}\cdot e^{\beta}\cdot\frac{|F_{SR}(i,j)|}{\max|F_{SR}(i,j)|} \\
\mathcal{L}_{\text{gkloss}} &= \sum_{i,j} w(i,j)\cdot\|F_{SR}(i,j)-F_{HR}(i,j)\|_1
\end{aligned}
\tag{8}
$$

where (i,j) denotes the frequency-domain pixel obtained via the Fourier transform. Two hyper-parameters α and β are introduced to exponentially compute the normalized magnitude and distance weights. Finally, an L1 loss is applied to the MRI in the frequency domain with weighted emphasis.

We shift the HF constraint from the frequency domain to the pixel domain and design a spatial high-frequency loss (HFLoss) to guide the network in preserving fine HF details. Instead of directly applying a high-pass filter to extract HF information, we use a Gaussian low-pass filter and obtain the HF component by subtracting the low-frequency part from the original image.

$$
\begin{aligned}
k\left(i,j\right) &= \frac{1}{2\pi\sigma^2}e^{\frac{-\left(i^2+j^2\right)}{2\sigma^2}} \\
I_{HF} &= I - I_{LF} \\
\mathcal{L}_{\text{hfloss}} &= \left\|I_{HF}^{SR}-I_{HF}^{HR}\right\|_1
\end{aligned}
\tag{9}
$$

where $I \in \mathbb{R}^{H\times W}$ represents the MRI pixels in the spatial domain, a Gaussian filter is first used as a low-pass filter with a hyperparameter $\sigma = 0.5$ to extract the LF components of the image. The total loss of the model is formulated as:

$$
\mathcal{L} = \mathcal{L}_1 + \lambda_1\mathcal{L}_{\text{gkloss}} + \lambda_2\mathcal{L}_{\text{hfloss}}, \tag{10}
$$

where λ_1, and λ_2 are the hyper-parameters of losses. We set $\lambda_1 = 0.05$, and $\lambda_2 = 0.05$ to balance the contributions of the three losses.

4 Experiments

4.1 Datasets

In this study, we evaluate MArFEN on two publicly available multi-contrast MRI datasets: the fastMRI dataset [66] (PDW as reference and FS-PDW as target) and the IXI dataset (PDW as reference and T2W as target) [19]. We utilize complex-valued MRI data, treating the real and imaginary components as two separate channels. To simulate realistic LR data, we adopt a common MRI downsampling protocol in the K-space domain. Specifically, we first apply the Fourier Transform to convert the image into K-space, retain only the central low-frequency region, and discard all high-frequency components. For a downsampling factor k, only the central $1/k^2$ frequency content of K-space

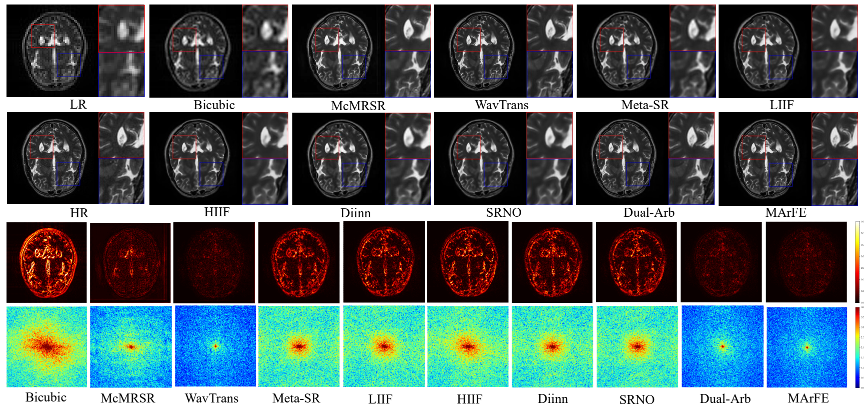

Fig. 3: Qualitative comparison for $\times 4$ SR to other super-resolution methods on IXI dataset. RDN-backbone [69] is used as an encoder for all ASSR methods. The local patches are zoomed in for better view. The last two rows are the spatial and frequency error maps of different SR methods on IXI dataset. The color bar on the right indicates the map values. Our method significantly outperforms other methods, reconstructing fewer blocking artifacts and sharper texture details.

is preserved. We then perform the Inverse Fourier Transform to obtain the final LR image in the spatial domain. We compare our method against several state-of-the-art baselines, including: Bicubic interpolation, Multi-contrast SR methods: McMRSR [27] and WavTrans [28], and Arbitrary-scale SR methods: Meta-SR [16], LIIF [5], HIIF [21], Diinn [40], SRNO [58], and Dual-Arb [68].

4.2 Implementation details

To simulate the continuous upsampling, we generate LR images using a random downsampling scale factor sampled uniformly from the range [1, 4], with stride 0.1, enabling the model to handle various degradation. We train the MArFENet using the Adam optimizer with an initial learning rate of 1e-4 for all modules. The learning rate is halved every 40 epochs, and training is conducted for 200 epochs in total. During training, we randomly crop LR images into 32×32 patches, and extract 6 patches as a batch input. The model is implemented in PyTorch and trained on a single NVIDIA GeForce RTX 3090 GPU. Following the data augmentation strategy in Meta-SR [16], we apply random horizontal or vertical flipping and 90°rotation to the image patches. To evaluate the reconstruction performance, we use two widely adopted metrics: PSNR and SSIM.

4.3 Quantitative and Qualitative Results

The quantitative comparison of different methods is summarized in Table 1, using the evaluation metrics of PSNR and SSIM. Since McMRSR [27] and WavTrans [28] are designed for fixed-scale SISR, we train them specifically for $\times 2$

Table 1: Quantitative comparison with other super-resolution methods on the fastMRI and IXI test set (PSNR(dB) and SSIM) with scaling factors ×1.5, ×2, ×3, ×4, ×6, ×8. All ASSR models use the same RDN-backbone [69] as the feature extraction encoder. We bold the best results and underline the second-best results.

Dataset	Methods	In-distribution						Out-of-distribution						Average	
		× 1.5		× 2		× 3		×4		× 6		× 8			
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
fastMRI	Bicubic	34.799	0.912	30.719	0.822	27.646	0.704	26.415	0.514	24.415	0.514	23.337	0.432	15.527	0.737
	McMRSR	37.773	0.947	34.546	0.884	31.087	0.788	30.141	0.734	27.859	0.630	26.200	0.547	17.417	0.842
	WavTrans	36.390	0.934	32.841	0.859	31.153	0.788	30.197	0.741	<u>28.370</u>	<u>0.663</u>	<u>26.722</u>	0.585	17.242	0.919
	Meta-SR	39.974	0.948	34.354	0.883	31.558	<u>0.811</u>	30.085	0.740	28.008	0.647	24.483	0.498	17.499	0.881
	LIIF	37.314	0.939	34.170	0.877	31.601	0.795	30.232	0.736	28.014	0.641	25.318	0.512	17.331	0.889
	HIIF	38.154	<u>0.950</u>	<u>34.772</u>	**0.888**	31.841	0.801	30.324	0.740	26.707	0.661	22.178	0.521	17.092	0.891
	Diinn	37.413	0.942	34.126	0.879	31.623	0.801	30.183	<u>0.744</u>	28.303	0.659	25.621	0.539	17.390	0.897
	SRNO	<u>38.348</u>	0.951	34.728	**0.888**	31.798	0.803	30.348	<u>0.744</u>	28.031	0.641	25.023	**0.625**	17.482	0.897
	DualArb	38.166	0.948	34.686	0.884	<u>31.987</u>	0.806	<u>30.578</u>	0.743	27.981	0.646	26.250	0.549	<u>17.607</u>	<u>0.963</u>
	MArFE	**38.449**	**0.951**	**34.968**	<u>0.887</u>	**32.371**	**0.812**	**31.080**	**0.757**	**28.852**	**0.670**	**26.871**	<u>0.598</u>	**17.879**	**0.964**
IXI	Bicubic	33.165	0.967	28.267	0.906	24.632	0.799	22.958	0.703	21.363	0.559	20.650	0.486	14.088	0.650
	McMRSR	37.450	0.966	37.046	0.901	34.416	0.956	33.910	0.818	29.765	0.771	27.239	0.641	18.567	0.755
	WavTrans	39.118	0.966	38.171	0.929	37.671	0.970	35.804	0.940	31.037	0.886	27.832	0.825	19.484	0.762
	Meta-SR	42.251	0.990	36.375	0.972	32.039	0.944	29.283	0.903	25.390	0.789	23.329	0.688	17.570	0.755
	LIIF	41.886	0.992	35.557	0.973	32.333	0.945	29.458	0.904	26.116	0.806	24.009	0.712	17.634	0.750
	HIIF	43.413	0.993	38.020	0.980	32.939	0.952	27.650	0.891	25.802	0.812	23.478	0.720	17.812	0.760
	Diinn	43.240	0.991	36.531	0.972	33.180	0.949	30.439	0.920	26.369	0.820	24.121	0.730	18.048	0.761
	SRNO	43.661	<u>0.994</u>	37.346	0.980	33.342	0.954	30.262	0.919	26.166	0.819	23.919	0.718	18.124	0.775
	DualArb	<u>44.691</u>	<u>0.994</u>	<u>41.187</u>	<u>0.986</u>	<u>38.478</u>	<u>0.977</u>	<u>36.863</u>	0.969	<u>32.950</u>	**0.946**	<u>28.681</u>	**0.906**	<u>20.702</u>	<u>0.763</u>
	MArFE	**46.663**	**0.996**	**42.317**	**0.988**	**39.525**	**0.979**	**38.041**	**0.973**	**34.356**	<u>0.945</u>	**30.625**	<u>0.905</u>	**21.492**	**0.779**

and ×4 scale factors. To ensure fair comparisons, we use the ×2 model weights to test on ×1.5 downsampled inputs, and the ×4 models weights to test on ×3 downsampled inputs, and on ×6 and ×8 upsampled inputs. From the results, our method consistently outperforms existing methods on average results. Notably, even in out-of-distribution situation, MArFE has strong robustness and adaptability, underscoring its great potential in real-world medical applications.

Figure 3 shows the in-distribution ×4 SR results of different methods on the IXI dataset, with spatial and frequency error maps. Brighter textures or colors indicate larger reconstruction errors. MArFE produces sharper textures and fewer blocking artifacts compared to interpolation-based methods. It preserves low-frequency components and recovers high-frequency details better. This superiority also exists in the zoomed-in patches. While ASSR methods tend to oversmooth and lose tissue details, MArFE not only retains rich textures but even surpasses fixed-scale methods like WavTrans in visual quality. Additional results for out-of distribution scales can be found in the supplementary materials.

4.4 Fully Automatic Segmentation Based Evaluation

To further validate the effectiveness of the proposed model, we evaluated the performance of MRI SR results on the downstream segmentation task. Specifically, we adopted the open-source Advanced Normalization Tools (ANTs), a medical image processing platform renowned for its robust normalization and multi-modality image segmentation [1]. Through the Atropos algorithm integrated within ANTs, fully automatic segmentation was performed on both the ground-truth (GT) images and the SISR outputs of all comparative models on the IXI dataset. Each brain MR image was segmented into three anatomical regions, including Cerebro-Spinal Fluid (CSF), Gray Matter (GM), and White Matter (WM). The Dice Similarity Coefficient (Dice) was quantitatively computed between the segmentation results of GT and SR paired images to assess the auxiliary value for clinical medical image segmentation.

Fig. 4 presents the qualitative results of the $\times 2$ SR segmentation experiments, where blue, yellow, and red labels correspond to CSF, GM, and WM, respectively. The zoomed-in local details clearly demonstrate that the segmentation results of MArFE are closest to GT standards compared with other baseline methods. Table 2 provides the quantitative results of the segmentation tests. Our model achieves the optimal segmentation performance across all anatomical regions within the tested upsampling scales. This also confirms the outstanding performance of SR results for downstream segmentation and related medical visual tasks, improving the accuracy of clinical medical diagnosis.

Table 2: Quantitative comparison of the fully automatic segmentation with SISR methods on the IXI dataset (Dice). We bold the best results and underline the second-best results of different anatomical regions for brain MRI. The model we proposed gets the best performance within the tested in-distribution and out-of-distribution scales.

Scales	Regions	Bicubic	McMRSR	WavTrans	Meta-SR	LIIF	Diinn	SRNO	HIIF	DualArb	MArFE
$\times$ 2	CSF	0.912	0.970	0.977	0.958	0.962	0.964	0.969	0.969	0.977	**0.981**
	GM	0.690	0.875	0.901	0.841	0.844	0.852	0.871	0.870	0.906	**0.920**
	WM	0.797	0.949	0.960	0.926	0.915	0.933	0.939	0.942	0.967	**0.970**
	Average	0.800	0.931	0.946	0.908	0.907	0.916	0.926	0.927	0.950	**0.957**
$\times$ 4	CSF	0.824	0.960	0.967	0.902	0.903	0.928	0.911	0.907	0.961	**0.968**
	GM	0.496	0.813	0.866	0.679	0.673	0.719	0.687	0.650	0.839	**0.867**
	WM	0.607	0.890	0.941	0.816	0.816	0.832	0.831	0.778	0.929	**0.950**
	Average	0.642	0.888	0.925	0.799	0.797	0.826	0.810	0.778	0.910	**0.928**
$\times$ 6	CSF	0.684	0.918	0.906	0.902	0.876	0.885	0.879	0.877	0.930	**0.932**
	GM	0.388	0.698	0.688	0.679	0.579	0.577	0.580	0.572	0.725	**0.752**
	WM	0.508	0.828	0.849	0.816	0.729	0.742	0.739	0.714	0.866	**0.893**
	Average	0.527	0.815	0.814	0.799	0.728	0.735	0.733	0.721	0.841	**0.859**

4.5 Ablation Study

We conduct ablations to evaluate the contributions of key components in the MArFE architecture. The final results are summarized in Table 3 (modules and decoder variants) and Table 4 (loss functions), with detailed analysis as follows.

Modules and Decoder Structure. Table 3 reports the impact of different modules on MRI reconstruction abilities. Decoder+ indicates a loop-back optimization strategy fusing the relative positional coordinates and encoder features. The results show that feedback mechanisms and residual connections significantly enhance SR performance across both in-distribution and out-of-distribution scales, leading to higher quantitative metrics. The NLGLA module

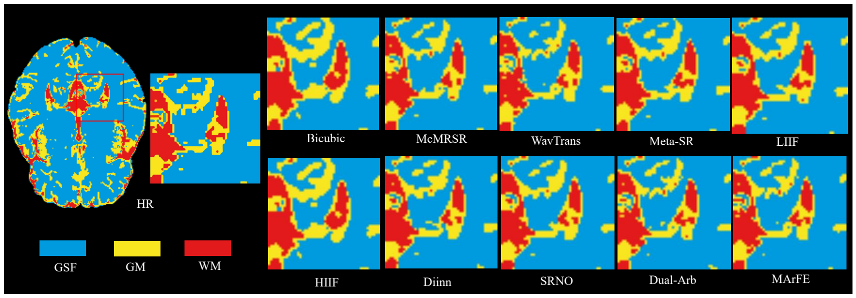

Fig. 4: Qualitative results of the fully automatic brain segmentation on a chosen sample from the ×2 SISR results of all the compared models on the IXI dataset. Different colors in the segmentation masks correspond to distinct anatomical regions of brain MRI, including Cerebro-Spinal Fluid (CSF), Gray Matter (GM), and White Matter (WM). The proposed method achieves the optimal performance in both segmentation integrity and detail preservation.

utilizes a Galerkin-type linear attention mechanism, which further refines the continuous mapping. Given the dual-stream structure for multi-contrast MRI, we perform ablations on the FRM to find its proper positions in INR decoder. Specifically, the FRM^a refers to inserting the Fourier reparameterization after the $ReLU$ activation branch, while the FRM^b applies them after the Sin activation branch. These results reveal that the strategy of Fourier basis decomposition after the higher dimensional mapping gets better performance, providing experimental support for INR-based methods.

Table 3: Ablation study on different key modules under fastMRI and IXI dataset (PSNR(dB) and SSIM). We bold the best results and underline the second-best results. Decoder+ represents the loop-back optimization with the fusion of features and coordinates. NLGLA is the Galerkin-type linear attention. FRM^a means the Fourier reparameterization after the $ReLu$ activation branch, while FRM^b means the Fourier reparameterization after the Sin activation branch, which gets the best performance.

Dataset	Module				In-distribution				Out-of-distribution		Average	
	Decoder+	NLGLA	FRM^a	FRM^b	× 1.5	× 2	× 3	×4	× 6	× 8	PSNR	SSIM
fastMRI	×	×	×	×	38.166	34.686	31.987	30.578	27.981	26.250	31.608	0.763
	✓	×	×	×	38.438	34.925	32.368	31.068	28.263	26.103	31.861	0.770
	✓	✓	×	×	**38.464**	34.962	32.329	31.002	28.755	26.812	32.054	0.774
	✓	×	×	✓	38.439	34.911	32.301	31.025	28.602	26.560	31.973	0.766
	✓	✓	✓	×	38.111	34.875	32.306	31.055	28.478	26.548	31.896	0.769
	✓	✓	×	✓	38.449	**34.968**	**32.371**	**31.080**	**28.852**	**26.871**	**32.099**	**0.779**
IXI	×	×	×	×	44.691	41.187	38.478	36.863	32.950	28.681	37.142	0.963
	✓	×	×	×	45.935	41.941	39.238	37.795	33.739	29.269	37.986	0.964
	✓	✓	×	×	46.208	42.056	39.427	38.002	34.242	30.125	38.343	0.969
	✓	×	×	✓	46.187	42.119	39.516	**38.098**	33.969	30.022	38.319	0.968
	✓	✓	✓	×	45.947	42.232	39.432	37.957	**35.909**	29.977	38.576	0.968
	✓	✓	×	✓	**46.663**	**42.317**	**39.525**	38.041	34.356	**30.625**	**38.588**	**0.970**

Table 4: Ablation study on loss functions under fastMRI and IXI (PSNR(dB) and SSIM). We bold the best results and underline the second-best results. HFloss represents the high-frequency loss in the spatial domain, and GKloss represents the global k-space loss. The best performance is achieved with the combination of losses above.

Dataset	Loss			In-distribution				Out-of-distribution		Average	
	L1	HFloss	GKloss	× 1.5	× 2	× 3	×4	× 6	× 8	PSNR	SSIM
fastMRI	✓	×	×	38.325	34.841	32.272	30.998	25.680	24.057	31.029	0.744
	✓	✓	×	38.408	<u>34.913</u>	<u>32.279</u>	30.962	28.243	26.011	31.803	0.764
	✓	×	✓	<u>38.449</u>	34.885	32.265	30.963	<u>28.499</u>	<u>26.150</u>	<u>31.868</u>	<u>0.767</u>
	✓	✓	✓	**38.449**	**34.968**	**32.371**	**31.080**	**28.852**	**26.871**	**32.099**	**0.779**
IXI	✓	×	×	46.291	42.188	39.489	<u>38.037</u>	33.691	30.441	38.356	0.966
	✓	✓	×	46.383	42.182	39.386	37.928	34.123	<u>30.462</u>	38.411	0.967
	✓	×	✓	<u>46.596</u>	<u>42.258</u>	<u>39.509</u>	38.033	<u>34.309</u>	30.133	<u>38.473</u>	<u>0.968</u>
	✓	✓	✓	**46.663**	**42.317**	**39.525**	**38.041**	**34.356**	**30.625**	**38.588**	**0.970**

United Spatial and Frequency Losses. Table 4 presents the effect of different combinations of designed loss functions on the reconstruction of high-frequency MRI details. High-frequency loss in spatial domain encourages the model to emphasize edge representations and regions with rich texture. Global k-space loss utilizes magnitude-weighted penalties in the frequency domain, which effectively guides the network to restore fine-grained details with higher fidelity. The best performance of our proposed method is achieved when both spatial and frequency losses are used jointly. Additional experiments on hyperparameter configuration of losses are included in the supplementary material.

5 Conclusion

In this paper, we propose a novel integration of multi-contrast MRI and frequency-domain signal characteristics by utilizing a Fourier neural operator with the non-local Galerkin-type linear attention. This module performs linear projection coupling between spatial coordinates and deep features, significantly enhancing the ability to capture structural patterns in the frequency domain. To further address the spectral bias issue and strengthen high-frequency reconstruction, we use Fourier reparameterization, which incorporates sinusoidal activation functions to increase the sensitivity to MRI-specific frequency features. Moreover, we optimize the dense decoder architecture by introducing the feature loop-back structure and residual enhancement strategies, effectively improving super-resolution performance. Finally, we propose specific loss functions that jointly supervise the model in both spatial and frequency domains, leveraging the complex-valued nature of MRI data to better preserve structural details and improve high-frequency fidelity. Extensive experiments conducted on multiple public datasets demonstrate that our proposed MArFE method achieves state-of-the-art performance both within and beyond the training distribution. For future research, more research can integrate richer multi-modal MRI data (e.g., fMRI, DTI) for comprehensive structural-functional SR. Moreover, extending the model to handle larger-scale volumetric data and investigating its robustness to practical clinical challenges, such as motion artifacts and acquisition variability, could further enhance its clinical applicability.

Acknowledgements

This work was partly supported by the National Natural Science Foundation of China (Nos.62311530100&62171251) and the Special Foundations for the Development of Strategic Emerging Industries of Shenzhen(No.KJZD20231023094700001). This work sincerely thanks Prof. Wenming Yang for his invaluable guidance, continuous encouragement, and unwavering support throughout this research. We also thank Prof. Xiaochen Yang and Prof. Rui Zhu for their insightful comments and valuable suggestions that significantly improved the manuscript, and Jiamian Bian, Jiamiao Zhang, and Yining Miao for their constructive discussions and generous assistance throughout this research.

References

1. Avants, B.B., Tustison, N., Song, G., et al.: Advanced normalization tools (ants). Insight j **2**(365), 1–35 (2009)
2. Blu, T., Thévenaz, P., Unser, M.: Linear interpolation revitalized. IEEE Transactions on Image Processing **13**(5), 710–719 (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: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1796–1807 (2023)
4. Chen, H., He, B., Wang, H., Ren, Y., Lim, S.N., Shrivastava, A.: Nerv: Neural representations for videos. Advances in Neural Information Processing Systems **34**, 21557–21568 (2021)
5. Chen, Y., Liu, S., Wang, X.: Learning continuous image representation with local implicit image function. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 8628–8638 (2021)
6. Chen, Z., Pawar, K., Ekanayake, M., Pain, C., Zhong, S., Egan, G.F.: Deep learning for image enhancement and correction in magnetic resonance imaging—state-of-the-art and challenges. Journal of Digital Imaging **36**(1), 204–230 (2023)
7. Chen, Z., Zhang, H.: Learning implicit fields for generative shape modeling. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 5939–5948 (2019)
8. Dar, S.U., Yurt, M., Shahdloo, M., Ildız, M.E., Tınaz, B., Çukur, T.: Prior-guided image reconstruction for accelerated multi-contrast mri via generative adversarial networks. IEEE Journal of Selected Topics in Signal Processing **14**(6), 1072–1087 (2020)
9. Delannoy, Q., Pham, C.H., Cazorla, C., Tor-Díez, C., Dollé, G., Meunier, H., Bednarek, N., Fablet, R., Passat, N., Rousseau, F.: Segsrgan: Super-resolution and segmentation using generative adversarial networks—application to neonatal brain mri. Computers in Biology and Medicine **120**, 103755 (2020)

10. Dong, C., Loy, C.C., He, K., Tang, X.: Image super-resolution using deep convolutional networks. IEEE transactions on pattern analysis and machine intelligence **38**(2), 295–307 (2015)
11. Ebner, M., Wang, G., Li, W., Aertsen, M., Patel, P.A., Aughwane, R., Melbourne, A., Doel, T., Dymarkowski, S., De Coppi, P., et al.: An automated framework for localization, segmentation and super-resolution reconstruction of fetal brain mri. NeuroImage **206**, 116324 (2020)
12. Fang, W., Tang, Y., Guo, H., Yuan, M., Mok, T.C., Yan, K., Yao, J., Chen, X., Liu, Z., Lu, L., et al.: Cycleinr: Cycle implicit neural representation for arbitrary-scale volumetric super-resolution of medical data. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11631–11641 (2024)
13. Feng, C.M., Fu, H., Yuan, S., Xu, Y.: Multi-contrast mri super-resolution via a multi-stage integration network. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 140–149. Springer (2021)
14. Feng, C.M., Wang, K., Lu, S., Xu, Y., Li, X.: Brain mri super-resolution using coupled-projection residual network. Neurocomputing **456**, 190–199 (2021)
15. He, Z., Jin, Z.: Latent modulated function for computational optimal continuous image representation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 26026–26035 (2024)
16. Hu, X., Mu, H., Zhang, X., Wang, Z., Tan, T., Sun, J.: Meta-sr: A magnification-arbitrary network for super-resolution. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 1575–1584 (2019)
17. Hu, Y., Li, J., Huang, Y., Gao, X.: Image super-resolution with self-similarity prior guided network and sample-discriminating learning. IEEE Transactions on Circuits and Systems for Video Technology **32**(4), 1966–1985 (2021)
18. Hu, Y., Wang, Y., Zhang, J.: Dear-gan: Degradation-aware face restoration with gan prior. IEEE Transactions on Circuits and Systems for Video Technology **33**(9), 4603–4615 (2023)
19. Imperial College London: Ixi dataset. `https://brain-development.org/ixi-dataset/` (2008)
20. Jiang, C., Sud, A., Makadia, A., Huang, J., Nießner, M., Funkhouser, T., et al.: Local implicit grid representations for 3d scenes. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 6001–6010 (2020)
21. Jiang, Y., Kwan, H.M., Peng, T., Gao, G., Zhang, F., Zhu, X., Sole, J., Bull, D.: Hiif: Hierarchical encoding based implicit image function for continuous super-resolution. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 2289–2299 (2025)
22. Kim, J., Lee, J.K., Lee, K.M.: Accurate image super-resolution using very deep convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1646–1654 (2016)
23. Kim, J., Lee, J.K., Lee, K.M.: Deeply-recursive convolutional network for image super-resolution. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1637–1645 (2016)
24. Lai, W.S., Huang, J.B., Ahuja, N., Yang, M.H.: Deep laplacian pyramid networks for fast and accurate super-resolution. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 624–632 (2017)
25. Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al.: Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4681–4690 (2017)

26. Lee, J., Jin, K.H.: Local texture estimator for implicit representation function. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 1929–1938 (2022)
27. Li, G., Lv, J., Tian, Y., Dou, Q., Wang, C., Xu, C., Qin, J.: Transformer-empowered multi-scale contextual matching and aggregation for multi-contrast mri super-resolution. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 20636–20645 (2022)
28. Li, G., Lyu, J., Wang, C., Dou, Q., Qin, J.: Wavtrans: Synergizing wavelet and cross-attention transformer for multi-contrast mri super-resolution. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 463–473. Springer (2022)
29. Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., Anandkumar, A.: Fourier neural operator for parametric partial differential equations. arXiv preprint arXiv:2010.08895 (2020)
30. Lim, B., Son, S., Kim, H., Nah, S., Mu Lee, K.: Enhanced deep residual networks for single image super-resolution. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops. pp. 136–144 (2017)
31. Liu, X., Li, G., Zhao, Z., Cao, Q., Zhang, Z., Yan, S., Xie, J., Tang, M.: Eafwgan: Enhanced alignment fusion-wasserstein generative adversarial network for turbulent image restoration. IEEE Transactions on Circuits and Systems for Video Technology **33**(10), 5605–5616 (2023)
32. Liu, Y., Jia, Q., Fan, X., Wang, S., Ma, S., Gao, W.: Cross-srn: Structure-preserving super-resolution network with cross convolution. IEEE Transactions on Circuits and Systems for Video Technology **32**(8), 4927–4939 (2021)
33. Lyu, Q., Shan, H., Steber, C., Helis, C., Whitlow, C., Chan, M., Wang, G.: Multi-contrast super-resolution mri through a progressive network. IEEE transactions on medical imaging **39**(9), 2738–2749 (2020)
34. Ma, C., Rao, Y., Cheng, Y., Chen, C., Lu, J., Zhou, J.: Structure-preserving super resolution with gradient guidance. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 7769–7778 (2020)
35. Mao, X., Shen, C., Yang, Y.B.: Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. Advances in neural information processing systems **29** (2016)
36. Mescheder, L., Oechsle, M., Niemeyer, M., Nowozin, S., Geiger, A.: Occupancy networks: Learning 3d reconstruction in function space. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 4460–4470 (2019)
37. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: Nerf: Representing scenes as neural radiance fields for view synthesis. Communications of the ACM **65**(1), 99–106 (2021)
38. Mo, S., Cai, M., Lin, L., Tong, R., Chen, Q., Wang, F., Hu, H., Iwamoto, Y., Han, X.H., Chen, Y.W.: Mutual information-based graph co-attention networks for multimodal prior-guided magnetic resonance imaging segmentation. IEEE Transactions on Circuits and Systems for Video Technology **32**(5), 2512–2526 (2021)
39. Molaei, A., Aminimehr, A., Tavakoli, A., Kazerouni, A., Azad, B., Azad, R., Merhof, D.: Implicit neural representation in medical imaging: A comparative survey. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 2381–2391 (2023)
40. Nguyen, Q.H., Beksi, W.J.: Single image super-resolution via a dual interactive implicit neural network. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision. pp. 4936–4945 (2023)

41. Park, J.J., Florence, P., Straub, J., Newcombe, R., Lovegrove, S.: Deepsdf: Learning continuous signed distance functions for shape representation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 165–174 (2019)
42. Park, K., Sinha, U., Barron, J.T., Bouaziz, S., Goldman, D.B., Seitz, S.M., Martin-Brualla, R.: Nerfies: Deformable neural radiance fields. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 5865–5874 (2021)
43. Plenge, E., Poot, D.H., Bernsen, M., Kotek, G., Houston, G., Wielopolski, P., Van Der Weerd, L., Niessen, W.J., Meijering, E.: Super-resolution methods in mri: can they improve the trade-off between resolution, signal-to-noise ratio, and acquisition time? Magnetic resonance in medicine **68**(6), 1983–1993 (2012)
44. Pumarola, A., Corona, E., Pons-Moll, G., Moreno-Noguer, F.: D-nerf: Neural radiance fields for dynamic scenes. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10318–10327 (2021)
45. Rahaman, N., Baratin, A., Arpit, D., Draxler, F., Lin, M., Hamprecht, F., Bengio, Y., Courville, A.: On the spectral bias of neural networks. In: International conference on machine learning. pp. 5301–5310. PMLR (2019)
46. Reed, A.W., Kim, H., Anirudh, R., Mohan, K.A., Champley, K., Kang, J., Jayasuriya, S.: Dynamic ct reconstruction from limited views with implicit neural representations and parametric motion fields. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 2258–2268 (2021)
47. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica D: nonlinear phenomena **60**(1-4), 259–268 (1992)
48. Shen, L., Pauly, J., Xing, L.: Nerp: implicit neural representation learning with prior embedding for sparsely sampled image reconstruction. IEEE Transactions on Neural Networks and Learning Systems **35**(1), 770–782 (2022)
49. Shi, F., Cheng, J., Wang, L., Yap, P.T., Shen, D.: Lrtv: Mr image super-resolution with low-rank and total variation regularizations. IEEE transactions on medical imaging **34**(12), 2459–2466 (2015)
50. Shi, K., Zhou, X., Gu, S.: Improved implicit neural representation with fourier reparameterized training. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 25985–25994 (2024)
51. Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A.P., Bishop, R., Rueckert, D., Wang, Z.: Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1874–1883 (2016)
52. Shi, Y., Xia, B., Jin, X., Wang, X., Zhao, T., Xia, X., Xiao, X., Yang, W.: Vmambair: Visual state space model for image restoration. IEEE Transactions on Circuits and Systems for Video Technology **35**(6), 5560–5574 (2025)
53. Strümpler, Y., Postels, J., Yang, R., Gool, L.V., Tombari, F.: Implicit neural representations for image compression. In: European Conference on Computer Vision. pp. 74–91. Springer (2022)
54. Sun, Y., Liu, J., Xie, M., Wohlberg, B., Kamilov, U.S.: Coil: Coordinate-based internal learning for imaging inverse problems. arXiv preprint arXiv:2102.05181 (2021)
55. Tancik, M., Srinivasan, P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J., Ng, R.: Fourier features let networks learn high frequency functions in low dimensional domains. Advances in neural information processing systems **33**, 7537–7547 (2020)

56. Wang, Z., Chen, J., Hoi, S.C.: Deep learning for image super-resolution: A survey. IEEE transactions on pattern analysis and machine intelligence **43**(10), 3365–3387 (2020)
57. Wang, Z., Ran, M., Yang, Z., Yu, H., Jing, J., Wang, T., Lu, J., Zhang, Y.: Generalizable mri motion correction via compressed sensing equivariant imaging prior. IEEE Transactions on Circuits and Systems for Video Technology **34**(12), 12550–12567 (2024)
58. Wei, M., Zhang, X.: Super-resolution neural operator. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 18247–18256 (2023)
59. 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 **27**(2), 1004–1015 (2022)
60. 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 **27**(2), 1004–1015 (2022)
61. Wu, Q., Li, Y., Xu, L., Feng, R., Wei, H., Yang, Q., Yu, B., Liu, X., Yu, J., Zhang, Y.: Irem: High-resolution magnetic resonance image reconstruction via implicit neural representation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 65–74. Springer (2021)
62. Xu, R., Wang, X., Chen, K., Zhou, B., Loy, C.C.: Positional encoding as spatial inductive bias in gans. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 13569–13578 (2021)
63. Xu, Z.J.: Understanding training and generalization in deep learning by fourier analysis. arXiv preprint arXiv:1808.04295 (2018)
64. Yang, W., Zhang, X., Tian, Y., Wang, W., Xue, J.H., Liao, Q.: Deep learning for single image super-resolution: A brief review. IEEE Transactions on Multimedia **21**(12), 3106–3121 (2019)
65. Yao, J.E., Tsao, L.Y., Lo, Y.C., Tseng, R., Chang, C.C., Lee, C.Y.: Local implicit normalizing flow for arbitrary-scale image super-resolution. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1776–1785 (2023)
66. Zbontar, J., Knoll, F., Sriram, A., Murrell, T., Huang, Z., Muckley, M.J., Defazio, A., Stern, R., Johnson, P., Bruno, M., et al.: fastmri: An open dataset and benchmarks for accelerated mri. arXiv preprint arXiv:1811.08839 (2018)
67. Zeng, K., Zheng, H., Cai, C., Yang, Y., Zhang, K., Chen, Z.: Simultaneous single- and multi-contrast super-resolution for brain mri images based on a convolutional neural network. Computers in biology and medicine **99**, 133–141 (2018)
68. Zhang, J., Chi, Y., Lyu, J., Yang, W., Tian, Y.: Dual arbitrary scale super-resolution for multi-contrast mri. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 282–292. Springer (2023)
69. Zhang, Y., Tian, Y., Kong, Y., Zhong, B., Fu, Y.: Residual dense network for image super-resolution. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2472–2481 (2018)
70. Zhang, Y., Van Rozendaal, T., Brehmer, J., Nagel, M., Cohen, T.: Implicit neural video compression. arXiv preprint arXiv:2112.11312 (2021)