

Calibrated Harmonic Overlaid Implicit Neural Representations for Multi-Dimensional Data

Honghang Chen¹, Xiujun Zhang², Xiaoli Sun¹, and
Mingqing Xiao³

¹ Shenzhen University, Shenzhen, China

² Shenzhen Polytechnic University, Shenzhen, China

³ Southern Illinois University Carbondale, USA
honghangchenhaha@gmail.com

Abstract. Implicit neural representation (INR) has emerged as a powerful prior for multi-dimensional data (e.g., multispectral images and videos). However, most INR methods employing periodic activation functions (e.g., Sine) predominantly rely on function composition. This mechanism introduces optimization instability as network depth increases, thereby limiting their performance. Meanwhile, these methods fail to incorporate proper physical priors to effectively alleviate spectrum bias. To address these issues, inspired by the commonalities between deep periodic networks and generalized Fourier series, we propose a novel Calibrated Harmonic Overlaid Implicit Neural Representation (CHOIR). Specifically, we utilize Coordinated Harmonic Superposition (CHS) to replace the conventional function composition used in most INRs, thereby ensuring optimization stability when scaling network depth. Furthermore, we introduce a Perceptual Spectrum Calibration (PSC) to mitigate spectrum bias. This calibration embeds the ubiquitous power-law spectrum prior of natural images and adjusts the globally fixed spectrum towards a physically plausible log-uniform distribution. Extensive experiments on various multidimensional data recovery problems demonstrate that our method achieves superior performance over state-of-the-art approaches. Code is available at <https://github.com/chor10229/CHOIR>.

Keywords: Implicit Neural Representation · Multi-Dimensional Data · Calibrated Harmonic Overlaid

1 Introduction

Implicit Neural Representations (INR) have become a powerful paradigm for multi-dimensional data recovery, such as multispectral images, videos, and other forms of data, by parameterizing the signal as a continuous function of coordinates [17, 22, 33, 35, 39, 52]. Among them, methods employing periodic activation functions (e.g., sine functions) [39] enable efficient modeling of high-frequency components and exhibit prominent advantages in representing high-frequency

 Corresponding author.

details, which has made them one of the important research directions in this field. Although subsequent works have continuously improved them through position encoding [12, 28, 42] or activation function design [22, 33, 35], the deep scalability of INR has not been fully explored, and spectral bias calibration matched to natural-image statistics remains underexplored.

In response, we reveal the reasons for the limitations of these methods from the perspective of harmonic signal representation. First, most periodic INR methods generally adopt a layer-wise function composition $\Phi_\theta = \phi_L \circ \dots \circ \phi_1$, whereas the superposition principle of signals requires the components to be additive. This architectural mismatch leads to a bottleneck: gradient pathologies that destabilize deep network optimization and constrain model expressiveness. Furthermore, the energy distribution of natural signals generally follows a $1/f$ power-law statistic [9], while most INR methods constrain all neurons with a globally fixed frequency scaling ω_0 . This gives rise to another bottleneck: unconstrained frequency distributions that impede model convergence. Resolving these critical limitations to enable stable and scalable deep periodic INRs constitutes our primary contribution. Figure 1 effectively confirms this problem. As shown, the PSNR of our method increases gradually with the network depth and achieves the highest value at 15 layers. All subsequent experiments use 12 layers for performance-efficiency trade-off. In contrast, other methods, such as SIREN [39], FINER [22] and FreSh [12], reach their peak PSNR at relatively shallow layers. A deeper network brings no performance gain for these methods; instead, it causes a clear performance degradation. Meanwhile, our method exhibits stronger robustness to the learning rate than other methods.

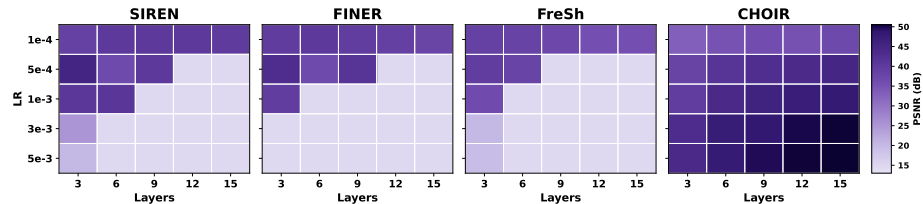


Fig. 1: PSNR heatmap vs. network depth and learning rate for sine-based INR methods under data missing completion (random missing and observation rate OR=0.30) on MSI Flowers dataset.

Inspired by the inherent equivalence between deep periodic INR and generalized Fourier series, we replace the function composition paradigm with hybrid composition-superposition architecture and propose **Calibrated Harmonic Overlaid Implicit Neural Representations (CHOIR)**. CHOIR includes two key mechanisms: Coordinated Harmonic Superposition (CHS), which eliminates optimization instability caused by deep networks via implicit curriculum learning, and Perceptual Spectrum Calibration (PSC), which embeds the $1/f$ power-law

spectrum prior into frequency allocation to systematically alleviate spectrum bias. The core contributions of this paper are summarized as follows:

- (i) We theoretically reveal the fundamental contradiction between the function composition architecture of deep periodic INR and the superposition principle of signals, clarifying the essential reason for the deep scalability limitation of most INR methods.
- (ii) We propose the Coordinated Harmonic Superposition (CHS) mechanism, which enables stable deep scaling of hybrid composition-superposition architecture through implicit curriculum learning, fully unleashing the representational potential of deep periodic INR.
- (iii) We introduce the Perceptual Spectrum Calibration (PSC) strategy, embedding the $1/f$ power-law spectrum prior as a physical calibration into frequency learning, significantly improving high-frequency detail reconstruction in multi-dimensional data recovery tasks.

2 Related Work

2.1 Implicit Neural Representations

INR parameterizes signals as continuous coordinate mappings, emerging as a powerful paradigm for multi-dimensional data representation. However, frequency principles [31, 45] reveal that ReLU-based MLPs inherently prioritize low-frequency components during training, introducing a bias against high-frequency signals. To alleviate this limitation, existing research has primarily focused on two directions: advances in position encoding and advances in activation functions.

On the position encoding front, Gaussian Fourier feature encoding [42] lifts low-dimensional coordinates into a high-dimensional frequency space via random mappings. NeRF [28] popularized sine position encoding for novel view synthesis, laying the foundation for modern coordinate encoding schemes. Subsequent works further refined this line of research. BACON [20] proposed band-limited coordinate networks to control spectrum bandwidth for multi-scale representation. PNF [46] enables controllable subband decomposition and manipulation, while BANF [36] realizes band-limited filtering for level-of-detail reconstruction. Instant-NGP [30] introduced multi-resolution hash encoding to drastically accelerate training convergence. FreSh [12] adaptively aligned frequency initialization with input data. Additionally, batch normalization has been shown to effectively mitigate spectrum bias in coordinate networks [5].

Regarding activation functions, SIREN [39] pioneered the use of sine activations for efficient high-frequency representation, establishing the theoretical framework for periodic INR. MSIREN [26] further modulates periodic activations for generalizable local representations, and INCODE [13] conditions sinusoidal activations with prior-knowledge embeddings. WIRE [35] extended this by employing continuous Gabor wavelet activations to jointly model time-frequency domains, while [32] constructed a unified framework that transcends pure periodicity using Gaussian activations. MIRE [11] matches layer-wise activation

functions to target signals through dictionary learning. Recent efforts have focused on adaptive and constrained designs. MFN [7] applies multiplicative filters directly to the input coordinates and admits a basis-expansion interpretation. However, it still relies on layer-wise nested multiplicative modulation rather than explicitly altering the functional representation through additive superposition. FINER [22] introduced variable-period activations for flexible spectrum adjustment. STAF [29] parameterized activations as trainable Fourier-series functions for adaptive spectral learning. SASNet [8] localized sinusoidal responses with spatially adaptive masks to reduce frequency leakage. [38] proposed Fourier reparameterization to impose structural constraints on weights in the frequency domain. SL2A-INR [33] leveraged a single-layer learnable activation for adaptive frequency response, and IGA-INR [37] further reduced spectrum bias through inductive gradient adjustments.

2.2 INR-based Methods in Multi-dimensional Data

Using the network structure itself as an implicit prior has emerged as an important research direction in multi-dimensional data recovery. Deep Image Prior (DIP) [43] first demonstrated the implicit regularization effect derived from the structure of untrained convolutional networks. CoIL [41] further introduced coordinate-based INR into imaging inverse problems, revealing the potential of MLP-based continuous coordinate mappings for image recovery tasks. Physics-guided imaging reconstruction further combines continuous priors with degradation modeling for inverse recovery [49]. Inspired by these advances, treating INR as a continuous prior has become a vital paradigm for multi-dimensional data recovery [25]. In this paradigm, discrete factor matrices are replaced by MLP-parameterized continuous factor functions, which unify low-rank structures and implicit smooth regularization within a unified framework.

Specifically, LRTFR [23] substitutes the discrete factor matrices in Tucker decomposition with MLP-based continuous factor functions, achieving a theoretical unification of low-rank constraints and Lipschitz smoothness. DRO-TFF [19] adopts three separate INR parameterizations for each dimensional factor under a rank-1 spatial decomposition and deep modal transformation framework, enabling space-spectral dual smoothness encoding with extremely few parameters. CRNL [24] leverages continuous functions instead of discrete block metrics to capture non-local self-similarity, and further exploits coupled low-rank function decomposition to model both cross-group shared patterns and intra-group individual variations. Furthermore, S-INR [17] incorporates superpixel semantic units with a joint-attention MLP and shared dictionaries. FA-INR [52] performs divide-and-conquer modeling of frequency components in a learnable frequency space, thereby enhancing high-frequency representation capacity. Adjacent low-level restoration studies also exploit progressive spectral decomposition and imaging-aware priors for robust detail recovery [27, 51].

Although existing methods have achieved encouraging progress, they still fall short of addressing the optimization instability and insufficient spectral calibra-

tion based on natural-signal statistics in deep periodic INR from a fundamental architectural perspective, which is the key motivation of this work.

3 Method

3.1 Our Motivations

Most INR methods based on periodic activation functions (e.g., sine [35, 39], Gabor wavelet [22]) normalize input coordinates $\mathbf{v} \in [-1, 1]^N$ (with N being the signal dimensionality) and map them to signal values layer by layer, resulting in the following network output:

$$\Phi_\theta(\mathbf{v}) = \phi_L \circ \dots \circ \phi_1(\mathbf{v}), \quad \phi_l(\mathbf{x}) \triangleq \rho^{(l)}(\mathbf{W}^{(l)}\mathbf{x} + \mathbf{b}^{(l)}) \quad (1)$$

where $\mathbf{x}$ is the input to the layer, $\mathbf{W}^{(l)}, \mathbf{b}^{(l)}$ are the weights and biases, and $\rho^{(l)}$ is the activation function. From this equation, the product of Jacobian matrices in backpropagation, $\mathbf{J}_\Phi = \prod_{l=1}^L \mathbf{J}_{\phi_l}$ [10], leads to gradient vanishing or exploding as depth increases [16], limiting network scalability.

From a harmonic perspective, the ideal signal representation should be the explicit superposition of frequency components, i.e., the generalized Fourier series:

$$\Phi_\theta(\mathbf{v}) = \sum_k \left[\mathbf{A}_k \cos(\boldsymbol{\omega}_k^\top \mathbf{v}) + \mathbf{B}_k \sin(\boldsymbol{\omega}_k^\top \mathbf{v}) \right] \quad (2)$$

where $\mathbf{A}_k, \mathbf{B}_k$ are the cosine and sine amplitudes of the k -th harmonic, and $\boldsymbol{\omega}_k$ are the angular frequencies. Notably, using the trigonometric identity $\sin(x+y) = \sin x \cos y + \cos x \sin y$, a sine unit with bias $\sin(\boldsymbol{\omega}^\top \mathbf{v} + b)$ can be decomposed as a linear combination of sine and cosine bases [2], making both representations fundamentally equivalent.

Comparing the composition structure $\phi_L \circ \dots \circ \phi_1$ with Eq. (2), we observe two major deficiencies in general INR architectures:

- (i) **Lack of a superposition paradigm.** The nested function composition fails to establish a one-to-one correspondence between learnable coefficients and basis functions under the $\sum_k$ summation operator, making it difficult to achieve term-wise pairing with $\{(\mathbf{A}_k, \mathbf{B}_k)\}$. This impedes harmonic coexistence and tends to introduce conflicting behaviors that aggravate overfitting.
- (ii) **Frequency rigidity.** Common INRs adopt a globally fixed base frequency scaling (e.g., scaling all neuron activations by a constant ω_0), resulting in a frequency support set entirely determined by initialization [48], which is inconsistent with the $1/f$ power-law statistics of natural signals [9]. Furthermore, since the magnitude of parameter gradients is proportional to their corresponding frequencies, the update step of the gradient in the high frequency part is much larger than that of the low frequency, which is easy to produce shock and lead to unstable training.

Based on these issues, although many INRs claim to integrate deep networks, experiments typically show optimal performance only in very shallow networks

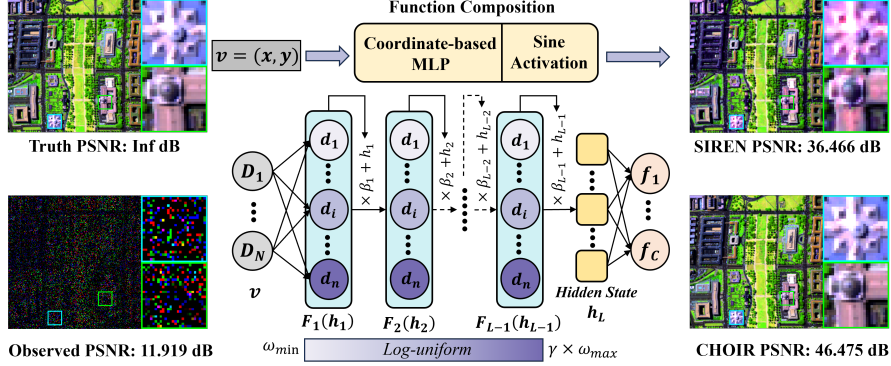


Fig. 2: Overview of the architecture of our proposed CHOIR vs. most periodic INR methods (e.g., SIREN [39]). CHOIR establishes a hybrid composition-superposition paradigm and leverages the power-law distribution of natural images for spectrum calibration. In the figure, different colors indicate different angular frequencies of neurons.

(3–5 layers), with rapid degradation as depth increases. Fourier harmonic analysis theory offers unique advantages in solving these two problems, motivating the work presented in this paper. The overview of proposed CHOIR is shown in Figure 2.

3.2 Coordinated Harmonic Superposition (CHS)

To address these issues, we propose the Coordinated Harmonic Superposition (CHS) mechanism, which restructures function composition into term-wise harmonic superposition. The network first maps the input coordinate $\mathbf{v}$ to an initial hidden state $\mathbf{h}_0(\mathbf{v}) = \mathbf{W}^{(0)}\mathbf{v} + \mathbf{b}^{(0)}$. Then, inspired by residual scaling [1], the output of each harmonic module $\mathcal{F}_l$ (detailed in Sec. 3.3) is scaled by a learnable scalar β_l (initialized to zero) and added to the input feature $\mathbf{h}_l$:

$$\mathbf{h}_{l+1} = \mathbf{h}_l + \beta_l \mathcal{F}_l(\mathbf{h}_l), \quad l = 0, 1, \dots, L-1 \quad (3)$$

Thus, at the beginning of training, the network degenerates into a simple linear mapping ($\mathbf{h}_L = \mathbf{h}_0$). This initialization ensures the overall Jacobian matrix $\mathbf{J}_\Phi$ is highly stable, providing a good starting point for optimization. The stability allows the network to first learn basic low-frequency components before handling higher-frequency details.

This structure enables the network output to be expressed as an explicit additive superposition, corresponding to the generalized Fourier series:

$$\Phi_\theta(\mathbf{v}) = \mathbf{W}_{\text{out}}\mathbf{h}_L(\mathbf{v}) = \mathbf{W}_{\text{out}}\left(\mathbf{h}_0(\mathbf{v}) + \sum_{l=0}^{L-1} \beta_l \mathcal{F}_l(\mathbf{h}_l(\mathbf{v}))\right) \quad (4)$$

where $\{\beta_l \mathbf{W}_{\text{out}} \mathcal{F}_l(\cdot)\}$ denotes a set of adaptive harmonic terms modulated by β_l and $\mathbf{W}_{\text{out}} \mathbf{h}_0(\mathbf{v})$ corresponds to the zero-frequency component. As noted in

Sec. 3.1, a sine unit with bias can be decomposed into orthogonal sine and cosine bases, supporting harmonic superposition.

As optimization proceeds, the contribution of each harmonic term is dynamically introduced through its weight β_l . The gradient of β_l is determined by the inner product between $\mathcal{F}_l(\mathbf{h}_l)$ and the downstream gradient $\frac{\partial \mathcal{L}}{\partial \mathbf{h}_{l+1}}$:

$$\frac{\partial \mathcal{L}}{\partial \beta_l} = \left\langle \mathcal{F}_l(\mathbf{h}_l), \frac{\partial \mathcal{L}}{\partial \mathbf{h}_{l+1}} \right\rangle \quad (5)$$

This follows from the chain rule $\frac{\partial \mathcal{L}}{\partial \beta_l} = \frac{\partial \mathcal{L}}{\partial \mathbf{h}_{l+1}}^\top \frac{\partial \mathbf{h}_{l+1}}{\partial \beta_l}$ with $\frac{\partial \mathbf{h}_{l+1}}{\partial \beta_l} = \mathcal{F}_l(\mathbf{h}_l)$. Eq. (5) reveals an implicit curriculum learning mechanism [3] in CHS: the weight β_l of a harmonic term increases only if its output direction $\mathcal{F}_l$ aligns with the negative gradient direction [15]. This guides the network to begin with the simplest model and progressively activate the harmonic components most effective for fitting the residual, decomposing the complex optimization into a sequence of simpler subproblems.

3.3 Perceptual Spectrum Calibration (PSC)

To resolve frequency rigidity and specify the internal structure of $\mathcal{F}_l$, we further propose the Perceptual Spectrum Calibration (PSC). This mechanism embeds physical priors into the network via frequency sampling and amplitude modulation, achieving stable harmonic allocation that matches natural signal statistics.

Log-uniform Frequency Distribution. We assign an angular frequency $\omega_{l,i}$ to each neuron in the l -th layer. The frequency range is set as $[\omega_{\min}, \omega_{\max}]$, with the lower bound $\omega_{\min} = \pi$, corresponding to the fundamental frequency on the normalized coordinate domain $[-1, 1]^N$. According to the Nyquist sampling theorem [34], the theoretical upper bound should be the Nyquist frequency $\omega_{\text{Nyq}} = \pi \cdot \min(D_1, \dots, D_N)/2$. However, when taking it as the upper bound, the spatial period corresponding to this frequency contains only two sampling points, with a phase difference of π between them. This causes the continuous sinusoid to degenerate into an alternating sequence of signs [40], leading to relatively unstable gradient optimization. Moreover, near this upper bound, the energy in natural signals is extremely low (following a $1/f$ spectrum [9]) and is sensitive to noise and sampling errors. Adopting it as the upper bound would waste capacity and increase the potential for error. Therefore we introduce a scaling factor γ (determined by ablation in Sec. 4.4) and set

$$\omega_{\max} = \gamma \cdot \omega_{\text{Nyq}} = \gamma \cdot (\pi \cdot \min(D_1, \dots, D_N)/2). \quad (6)$$

Within the interval $[\omega_{\min}, \omega_{\max}]$, frequencies for d neurons are assigned according to a geometric progression (i.e., equally spaced on a logarithmic axis):

$$\omega_{l,i} = \omega_{\min} \cdot \left(\frac{\omega_{\max}}{\omega_{\min}} \right)^{\frac{i-1}{d-1}}, \quad i = 1, \dots, d. \quad (7)$$

Since adjacent neurons satisfy $\Delta\omega \propto \omega$, the sampling density becomes $\rho(\omega) \propto 1/\omega$, which matches the power-law spectrum $P(\omega) \propto 1/\omega$ of natural signals [9].

Algorithm 1 CHOIR Algorithm for Multi-dimensional Data

Require: Degraded observations $\mathcal{Y}$, mask Ω , depth L , width d
Ensure: Network parameters θ

- 1: Pre-compute $\{\omega_{l,i}, \tilde{\omega}_l\}$ for all l via log-uniform spacing (Eq. (6) Eq. (7))
- 2: Initialize $\{\mathbf{W}^{(l)}, \mathbf{b}^{(l)}\}$ with SIREN scheme; $\beta_l \leftarrow 0$; $\alpha \leftarrow 2.0$
- 3: **for** epoch = 1 to MaxEpochs **do**
- 4: $\mathbf{h}_0 \leftarrow \mathbf{W}^{(0)}\mathbf{v} + \mathbf{b}^{(0)}$
- 5: **for** $l = 0$ to $L - 1$ **do**
- 6: Compute $\mathcal{F}_l(\mathbf{h}_l)$ via PSC amplitude modulation (Eq. (8))
- 7: Update $\mathbf{h}_{l+1}$ via CHS additive gating (Eq. (3))
- 8: **end for**
- 9: $\Phi_\theta(\mathbf{v}) \leftarrow \mathbf{W}_{\text{out}}\mathbf{h}_L + \mathbf{b}_{\text{out}}$
- 10: Update θ via Adam on $\mathcal{L}(\mathcal{P}_\Omega(\Phi_\theta), \mathcal{P}_\Omega(\mathcal{Y}))$
- 11: **end for**
- 12: **return** θ

Consequently, model capacity allocated to each frequency band is proportional to its energy, automatically assigning more parameters to energy-dense low frequencies while maintaining sufficient high-frequency coverage.

Adaptive Amplitude Modulation. To counteract the optimization instability caused by high-frequency gradients dominating in standard periodic networks, we design an adaptive amplitude modulation method. Leveraging the power-law spectrum decay $P(\omega) \propto \omega^{-\alpha}$, we constrain amplitudes to be proportional to the square root of the power spectrum density, i.e., $A \propto P^{1/2} \propto \omega^{-\alpha/2}$. Combined with log-uniform sampling, the output of the i -th neuron in $\mathcal{F}_l$ is defined as:

$$\mathcal{F}_l(\mathbf{h}_l)_i = \left(\frac{\tilde{\omega}_l}{\omega_{l,i}} \right)^{\alpha/2} \sin(\omega_{l,i}(\mathbf{w}_{l,i}^\top \mathbf{h}_l + b_{l,i})), \quad (8)$$

where α is a globally learnable spectral decay rate (initialized to 2.0), $\tilde{\omega}_l = (\prod_{j=1}^d \omega_{l,j})^{1/d}$ is the geometric mean of frequencies in layer l , computed once at initialization and kept fixed. The amplitude factor $(\tilde{\omega}_l/\omega_{l,i})^{\alpha/2}$ compresses high-frequency neuron outputs following the power law, suppressing their tendency to dominate gradients. Specifically, when $\alpha = 2$, the gradient norm with respect to $\mathbf{w}_{l,i}$ becomes proportional to $\tilde{\omega}_l \cdot |\cos(\cdot)| \cdot \|\mathbf{h}_l\|_2$, which has the same expectation for all frequencies, yielding a naturally balanced training initialization. During training, α is adaptively updated to fit the actual spectral distribution of the target signal. In summary, PSC integrates physical priors through coordinated frequency and amplitude modulation, not only constraining the solution space but also improving the geometry of the optimization landscape, leading to stable and efficient training.

3.4 Problem Formulation

To solve the inverse problem of multi-dimensional data recovery, we employ an implicit neural representation (INR) as a self-supervised prior. Concretely, we

parameterize the target tensor $\mathcal{T}$ as a continuous mapping $\Phi_\theta : \mathbb{R}^N \rightarrow \mathbb{R}^C$ implemented by an MLP, where N and C are the coordinate dimensionality and number of channels, respectively. For any coordinate $\mathbf{v} \in [-1, 1]^N$, the network outputs $\Phi_\theta(\mathbf{v}) \in \mathbb{R}^C$. Querying the full coordinate grid produces the reconstructed tensor $\mathcal{T}(\theta)$. The recovery problem is then formulated as an optimization over the network parameters θ :

$$\min_{\theta} \mathcal{L}(\theta) = \ell(\mathcal{P}_\Omega(\Phi_\theta), \mathcal{P}_\Omega(\mathcal{Y})), \quad (9)$$

where $\mathcal{Y}$ is the degraded observation, $\mathcal{P}_\Omega$ is the projection operator that retains only entries in the observable index set Ω , and $\ell(\cdot, \cdot)$ is a task-adaptive per-sample loss. The complete training procedure is summarized in Algorithm 1. More analysis and results on convergence are provided in the supplementary material.

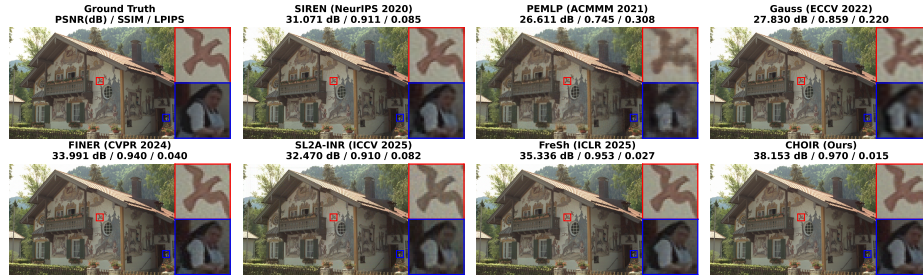
4 Experiment

4.1 Experimental Setup

We employ Peak Signal-to-Noise Ratio (PSNR, in dB), Structural Similarity Index (SSIM) [44], and Learned Perceptual Image Patch Similarity (LPIPS) [50] to assess recovery quality. Additionally, the number of parameters (Params, in M) and runtime (Time, in seconds) are used to comprehensively evaluate all methods. All reported results are the average of five independent runs. In implementation, we use the Adam optimizer [14] with a learning rate of 1×10^{-4} and weight decay. For all compared methods, we either use their open-source implementations or strictly reproduce them according to the published papers, carefully tuning hyperparameters to ensure optimal performance. All experiments are conducted on two NVIDIA RTX A6000 GPUs. In all tables of comparisons, the **best** and second-best values are highlighted.

4.2 Evaluation of Representation Capability

To assess the intrinsic representation capability of INRs, we conduct two types of experiments: (1) **2D signal fitting**, which is commonly used as a benchmark for INRs due to its rich composition of both low and high-frequency components. We employ the House image ($512 \times 768 \times 3$) from the Kodak dataset [6] as the test data. (2) **5D Novel view synthesis**, which examines the capability of modeling neural radiance fields from 5D coordinates to RGB. We utilize the classic NeRF Blender synthetic dataset [28], which comprises 8 synthetic scenes with complex geometry and fine textures. From the training set, 25 images are randomly sampled (downsampled to 200×200) for training, and the synthesis quality on unseen viewpoints is evaluated. Comparison methods include: MLP with Positional Encoding (PEMLP) [42], SIREN [39] based on sinusoidal activation functions, Gauss [32] which employs Gaussian activation functions to overcome purely periodic limitations, FINER [22] which flexibly regulates spectral bias through variable-period activations, SL2A-INR [33] which introduces

**Fig. 3:** Results of signal fitting by different methods on RGB House dataset.**Table 1:** Quantitative results on novel view synthesis in various methods.

	Methods	Chair	Drums	Ficus	Hotdog	Lego	Materials	Mic	Ship
PSNR $\uparrow$	PEMLP [28]	32.576	25.418	24.665	31.421	25.340	25.319	28.281	25.277
	SIREN [39]	32.843	24.719	24.949	35.187	23.793	27.114	32.394	22.719
	Gauss [32]	34.115	24.237	25.458	36.555	20.034	28.877	28.354	25.235
	FINER [22]	32.249	24.553	25.472	33.901	23.700	26.516	31.633	25.625
	SL2A-INR [33]	34.273	25.257	25.539	36.088	24.644	29.698	33.357	25.094
	IGA-INR [37]	34.676	25.997	24.891	36.066	23.550	28.131	31.584	27.368
	CHOIR(Ours)	35.317	26.430	27.650	37.454	28.299	29.793	35.613	27.555
SSIM $\uparrow$	PEMLP [28]	0.972	0.921	0.904	0.959	0.904	0.918	0.963	0.893
	SIREN [39]	0.973	0.915	0.908	0.977	0.862	0.933	0.980	0.894
	Gauss [32]	0.979	0.893	0.909	0.981	0.768	0.944	0.964	0.895
	FINER [22]	0.967	0.901	0.911	0.971	0.854	0.920	0.977	0.886
	SL2A-INR [33]	0.981	0.916	0.922	0.979	0.886	0.951	0.981	0.896
	IGA-INR [37]	0.984	0.930	0.903	0.980	0.884	0.937	0.973	0.917
	CHOIR(Ours)	0.986	0.939	0.952	0.985	0.933	0.955	0.987	0.926
LPIPS $\downarrow$	PEMLP [28]	0.036	0.097	0.161	0.048	0.077	0.086	0.050	0.093
	SIREN [39]	0.034	0.103	0.153	0.021	0.130	0.053	0.020	0.134
	Gauss [32]	0.031	0.174	0.209	0.021	0.258	0.045	0.045	0.129
	FINER [22]	0.040	0.112	0.175	0.030	0.150	0.055	0.019	0.130
	SL2A-INR [33]	0.025	0.120	0.110	0.024	0.110	0.034	0.015	0.084
	IGA-INR [37]	0.078	0.090	0.144	0.042	0.094	0.036	0.064	0.069
	CHOIR(Ours)	0.017	0.088	0.078	0.015	0.088	0.030	0.013	0.080

learnable single-layer activation functions, and IGA-INR [37] which mitigates spectral bias via induced gradient adjustment.

On the signal fitting task, as shown in Figure 3, CHOIR consistently outperforms all comparison methods across all evaluation metrics (PSNR, SSIM, LPIPS) on the House dataset. Moreover, the visual results reveal that our method fits higher-frequency signals, such as the clearer contour of the woman under the eaves. For novel view synthesis, Table 1 systematically demonstrates that CHOIR also achieves superior performance on the NeRF Blender dataset. From the zoomed-in details in Figure 4, it can be observed that CHOIR recovers the drums with more lustrous stands and sharper body. These two experiments validate the excellent fitting capability and generality of our proposed method.

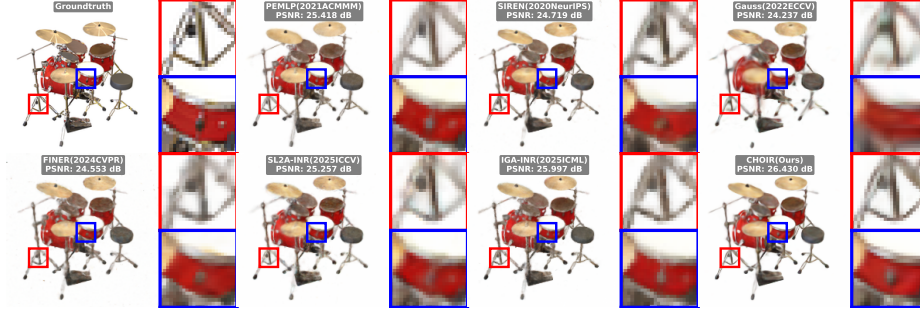


Fig. 4: Results of novel view synthesis by different methods on Drums dataset.

Table 2: Quantitative results on data missing completion in various methods.

Datasets	Methods	Random OR=0.10			Random OR=0.30			Tube OR=0.10			Tube OR=0.30		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
HSI WDC	SIREN [39]	36.466	0.901	0.032	37.149	0.905	0.029	24.670	0.661	0.260	27.832	0.783	0.134
	LRTFR [23]	32.530	0.917	0.121	34.201	0.942	0.069	23.643	0.648	0.367	27.983	0.834	0.168
	CRNL [24]	44.212	0.994	0.009	50.694	0.998	0.003	27.158	0.803	0.246	31.207	0.919	0.089
	FINER [22]	43.016	0.980	0.010	46.090	0.988	0.007	24.390	0.696	0.374	27.392	0.818	0.214
	DRO-TFF [19]	40.254	0.981	0.015	42.084	0.985	0.009	26.152	0.743	0.387	30.861	0.904	0.085
	FreSh [12]	38.762	0.916	0.020	40.145	0.922	0.017	23.310	0.607	0.428	25.883	0.713	0.265
	CHOIR	46.475	0.996	0.004	51.359	0.999	0.002	28.142	0.832	0.196	31.414	0.928	0.080
HSI Urban	SIREN [39]	39.482	0.984	0.012	40.835	0.988	0.010	21.756	0.562	0.272	25.138	0.782	0.109
	LRTFR [23]	29.979	0.978	0.135	32.479	0.933	0.063	19.478	0.315	0.499	24.617	0.700	0.218
	CRNL [24]	42.533	0.993	0.006	47.955	0.998	0.003	23.423	0.638	0.255	27.343	0.840	0.091
	FINER [22]	38.495	0.979	0.013	39.504	0.982	0.011	21.440	0.494	0.365	24.692	0.716	0.184
	DRO-TFF [19]	37.150	0.979	0.014	38.671	0.984	0.009	23.023	0.601	0.293	27.357	0.843	0.096
	FreSh [12]	38.974	0.982	0.013	39.791	0.985	0.011	20.242	0.440	0.416	23.800	0.692	0.254
	CHOIR	44.412	0.995	0.005	49.317	0.998	0.001	23.493	0.643	0.247	27.578	0.847	0.096
MSI Flowers	SIREN [39]	37.268	0.854	0.036	38.570	0.863	0.030	28.159	0.764	0.119	32.992	0.836	0.054
	LRTFR [23]	33.735	0.892	0.160	35.660	0.933	0.117	25.337	0.711	0.337	31.677	0.884	0.176
	CRNL [24]	39.975	0.991	0.031	48.604	0.998	0.005	27.333	0.882	0.182	33.097	0.964	0.053
	FINER [22]	37.995	0.940	0.051	45.436	0.971	0.006	24.833	0.743	0.237	31.814	0.869	0.078
	DRO-TFF [19]	40.912	0.986	0.017	45.754	0.996	0.005	27.991	0.871	0.146	33.277	0.964	0.051
	FreSh [12]	36.660	0.855	0.040	39.569	0.869	0.021	27.096	0.735	0.166	31.898	0.824	0.057
	CHOIR	41.316	0.992	0.014	50.395	0.998	0.001	28.726	0.894	0.138	33.465	0.973	0.048
MSI Toys	SIREN [39]	37.147	0.927	0.035	38.964	0.934	0.029	25.547	0.827	0.182	30.595	0.906	0.058
	LRTFR [23]	32.135	0.907	0.144	34.115	0.943	0.082	22.771	0.707	0.398	28.963	0.884	0.170
	CRNL [24]	39.582	0.992	0.018	49.015	0.998	0.003	24.908	0.853	0.212	30.883	0.955	0.054
	FINER [22]	37.430	0.964	0.051	46.818	0.989	0.005	22.863	0.775	0.316	28.435	0.918	0.122
	DRO-TFF [19]	39.846	0.992	0.013	43.610	0.996	0.004	25.103	0.851	0.217	30.728	0.951	0.061
	FreSh [12]	36.553	0.927	0.052	40.018	0.939	0.025	24.443	0.785	0.267	29.580	0.893	0.089
	CHOIR	41.748	0.993	0.010	49.216	0.998	0.002	25.640	0.864	0.211	30.903	0.958	0.051
Video Shop	SIREN [39]	28.146	0.854	0.123	31.551	0.933	0.061	21.087	0.577	0.485	25.704	0.861	0.168
	LRTFR [23]	25.567	0.864	0.182	27.667	0.902	0.118	20.229	0.599	0.449	24.671	0.844	0.179
	CRNL [24]	27.877	0.904	0.094	29.755	0.932	0.069	21.150	0.774	0.285	26.390	0.907	0.109
	FINER [22]	24.470	0.679	0.320	30.914	0.902	0.081	16.171	0.203	0.969	23.848	0.698	0.378
	DRO-TFF [19]	28.729	0.927	0.060	29.376	0.936	0.056	22.338	0.759	0.286	26.362	0.900	0.100
	FreSh [12]	26.564	0.788	0.188	31.451	0.924	0.068	18.762	0.376	0.712	24.433	0.767	0.338
	CHOIR	29.383	0.942	0.034	33.581	0.974	0.016	23.267	0.781	0.263	26.656	0.918	0.099
Video News	SIREN [39]	31.502	0.933	0.047	33.586	0.947	0.036	20.395	0.731	0.231	26.120	0.906	0.053
	LRTFR [23]	28.188	0.855	0.149	30.566	0.901	0.084	18.684	0.459	0.482	24.820	0.757	0.250
	CRNL [24]	31.618	0.953	0.036	36.686	0.983	0.011	21.825	0.716	0.286	26.390	0.918	0.099
	FINER [22]	30.414	0.894	0.056	36.631	0.958	0.012	19.819	0.554	0.385	25.170	0.805	0.145
	DRO-TFF [19]	32.505	0.955	0.025	35.557	0.969	0.014	22.148	0.805	0.227	27.402	0.917	0.075
	FreSh [12]	29.532	0.912	0.062	34.337	0.953	0.017	19.054	0.528	0.430	25.408	0.845	0.117
	CHOIR	33.502	0.963	0.021	38.739	0.986	0.006	22.597	0.818	0.166	27.518	0.920	0.069

4.3 Evaluation on Data Recovery Task

To assess the performance of CHOIR on multi-dimensional data recovery tasks, we conduct two categories of experiments: (1) **Data Missing Completion**, used to evaluate the capability of recovering complete data from partial observations. Experiments are performed under two observation rates ($OR \in \{10\%, 30\%\}$), considering both random missing and tube missing patterns. (2) **Mixed Degradation Restoration**, used to evaluate the recovery capability under coexisting

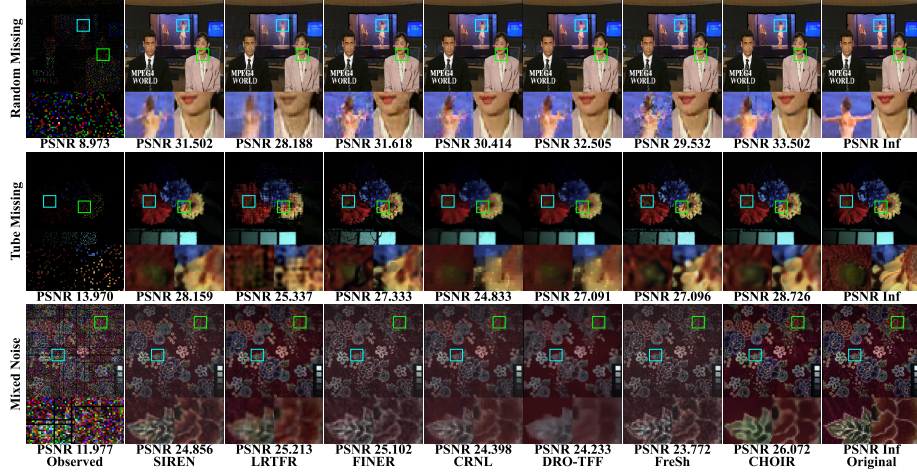


Fig. 5: Results of different methods on data recovery tasks. Top: random missing and OR=0.1 on Video News dataset. Middle: tube missing and OR=0.1 on MSI Flowers dataset. Bottom: mixed degradation restoration (Scene 3) on MSI Cloth dataset.

complex noise and non-uniform missing data. Three progressive scenarios are designed: **Scene 1** adds Gaussian noise ($\sigma = 0.20$); **Scene 2** further adds salt-and-pepper noise with a sampling rate ($SR = 10\%$); **Scene 3** additionally removes 3% of entire rows and 3% of entire columns from all channels to simulate sensor stripe failures. The datasets used in these two categories cover various modalities of multi-dimensional data: (1) **Hyperspectral Images (HSI)**: Pavia University (cropped to $200 \times 200 \times 80$) [21], Washington DC Mall (cropped to $256 \times 256 \times 191$) [21], and Urban (downsampled to $256 \times 256 \times 162$); (2) **Multispectral Images (MSI)**: Cloth, Toys, and Flowers from the CAVE dataset [47] (all downsampled to $256 \times 256 \times 31$); (3) **Videos**: Grayscale video Shop ($144 \times 192 \times 157$) [18] and RGB video News ($288 \times 352 \times 3 \times 10$) [4]; (4) **RGB Images**: Bird and Statue from the Kodak dataset [6] (downsampled to $256 \times 256 \times 3$). In addition to representative INR methods, we also include continuous representation methods that integrate structured priors with INRs as baselines. These include SIREN [39], FINER [22], FreSh [12] (which initializes frequency offsets to match the target signal spectrum), LRTFR [23] (which incorporates functional transformations into a low-rank factorization framework), CRNL [24] (which exploits non-local self-similarity priors via continuous representations), and DRO-TFF [19] (based on depth-one tensor function factorization). Notably, in mixed degradation experiments, continuous representation methods such as LRTFR employ TV regularization following their original settings to strengthen denoising and smoothing. In contrast, CHOIR inherently encodes strong implicit smoothing priors and does not require task-specific explicit regularization such as TV regularization.

Table 3: Quantitative results under mixed degradation scenario in various methods.

Datasets	Methods	Scene 1			Scene 2			Scene 3		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
HSI WDC	SIREN [39]	24.729	0.588	0.134	24.688	0.585	0.142	24.585	0.578	0.149
	LRTFR [23]	<u>28.615</u>	<u>0.610</u>	0.292	<u>28.468</u>	0.607	0.302	<u>28.255</u>	0.591	0.308
	CRNL [24]	23.373	0.435	0.574	23.602	0.464	0.538	23.529	0.459	0.489
	FINER [22]	24.766	0.585	<u>0.128</u>	24.774	0.586	<u>0.130</u>	24.669	0.575	0.146
	DRO-TFF [19]	24.188	0.518	0.340	25.532	<u>0.635</u>	0.521	25.527	<u>0.634</u>	0.499
	FreSh [12]	24.646	0.582	0.144	24.620	0.578	<u>0.146</u>	24.423	0.567	0.152
	CHOIR	31.242	0.884	0.118	31.047	0.883	0.126	30.595	0.870	0.145
HSI Pavia	SIREN [39]	27.641	0.870	0.059	27.507	<u>0.867</u>	<u>0.062</u>	27.228	0.853	<u>0.068</u>
	LRTFR [23]	<u>28.212</u>	0.824	0.182	<u>28.080</u>	0.819	0.189	<u>27.847</u>	0.812	0.194
	CRNL [24]	27.103	0.820	0.093	27.028	0.816	0.095	26.790	0.810	0.098
	FINER [22]	27.501	<u>0.871</u>	0.082	27.392	0.865	0.083	27.107	<u>0.854</u>	0.089
	DRO-TFF [19]	27.323	0.836	0.095	26.105	0.730	0.224	25.865	0.718	0.242
	FreSh [12]	27.056	0.855	0.076	26.756	0.839	0.075	26.718	0.829	0.107
	CHOIR	30.401	0.904	0.048	29.959	0.895	0.057	29.691	0.882	0.047
MSI Flowers	SIREN [39]	23.400	0.470	0.187	23.356	0.471	<u>0.195</u>	23.315	0.465	<u>0.201</u>
	LRTFR [23]	<u>29.749</u>	<u>0.601</u>	0.227	29.676	0.600	0.237	29.528	0.596	0.237
	CRNL [24]	23.822	0.465	0.297	23.766	0.464	0.205	23.624	0.455	0.370
	FINER [22]	23.345	0.462	0.224	23.287	0.456	0.236	23.204	0.452	0.245
	DRO-TFF [19]	23.861	0.469	0.154	30.151	0.701	0.214	29.805	0.696	0.221
	FreSh [12]	23.068	0.452	0.228	22.975	0.465	0.272	22.956	0.451	0.218
	CHOIR	33.094	0.856	0.122	32.919	0.852	0.129	32.605	0.847	0.130
MSI Cloth	SIREN [39]	25.435	0.774	<u>0.164</u>	25.250	<u>0.768</u>	0.174	24.856	0.753	0.182
	LRTFR [23]	<u>26.076</u>	0.733	0.216	<u>25.953</u>	0.727	0.221	<u>25.213</u>	0.716	0.229
	CRNL [24]	25.265	<u>0.783</u>	0.201	25.301	0.765	0.194	25.102	<u>0.777</u>	0.205
	FINER [22]	24.826	0.753	0.268	24.504	0.760	<u>0.161</u>	24.398	0.752	<u>0.173</u>
	DRO-TFF [19]	25.799	0.726	0.171	24.584	0.739	0.261	24.233	0.717	0.262
	FreSh [12]	24.422	0.756	0.212	24.287	0.749	0.220	23.772	0.727	0.237
	CHOIR	26.612	0.798	0.138	26.482	0.794	0.147	26.072	0.779	0.152
RGB Bird	SIREN [39]	<u>24.639</u>	<u>0.695</u>	<u>0.324</u>	<u>24.505</u>	<u>0.679</u>	<u>0.344</u>	<u>24.345</u>	<u>0.667</u>	<u>0.355</u>
	LRTFR [23]	17.217	0.357	0.808	17.089	0.336	0.822	16.986	0.323	0.824
	CRNL [24]	20.914	0.338	0.710	20.009	0.304	0.752	19.482	0.286	0.784
	FINER [22]	22.793	0.680	0.391	22.646	0.664	0.404	22.655	0.664	0.407
	DRO-TFF [19]	19.046	0.277	0.895	17.460	0.626	0.751	17.479	0.647	0.744
	FreSh [12]	23.015	0.674	0.400	22.864	0.653	0.422	22.797	0.647	0.435
	CHOIR	25.536	0.736	0.311	25.380	0.724	0.325	25.115	0.720	0.335
RGB Statue	SIREN [39]	<u>24.464</u>	<u>0.681</u>	0.481	24.350	0.673	<u>0.488</u>	24.220	0.666	<u>0.492</u>
	LRTFR [23]	22.391	0.487	0.635	22.012	0.464	0.656	21.680	0.445	0.662
	CRNL [24]	18.976	0.317	0.958	18.388	0.293	1.033	18.171	0.285	1.037
	FINER [22]	23.547	0.618	0.494	23.225	0.598	0.498	22.820	0.574	0.494
	DRO-TFF [19]	21.231	0.446	0.763	24.118	0.650	0.524	24.100	0.651	0.527
	FreSh [12]	22.990	0.584	<u>0.477</u>	22.701	0.633	0.493	22.628	0.627	0.501
	CHOIR	25.534	0.702	0.446	25.456	0.700	0.470	25.175	0.679	0.481

Tables 2 and 3 systematically demonstrate that CHOIR achieves the best performance on both categories of multi-dimensional data recovery experiments. This consistent advantage in scenarios with large-scale missing entries or dense noise with non-uniform structured missing entries strongly validates the detail preservation and recovery capabilities of CHOIR. The recovery of various multi-dimensional data (e.g., RGB, HSI, Video) also systematically reflects the generalization of CHOIR. As shown in Figure 5, the visual results indicate that our method exhibits robust performance against uniform missing, non-uniform missing, and complex noise. For instance, the lip edges in News are smoother, the petal details in Flowers are more realistic, and the pattern colors in Cloth are more lustrous. Due to space limitations, more experimental visualization results

Table 4: Ablation studies of our core components on MSI Flowers dataset.

Variants	Scale γ	Random OR=0.10			Scene 3			Params	Time/iter
		PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$	PSNR $\uparrow$	SSIM $\uparrow$	LPIPS $\downarrow$		
Sine	-	33.500	0.929	0.099	26.238	0.487	0.363	0.202M	0.046s
Sine+CHS	-	39.136	<u>0.982</u>	0.018	<u>32.296</u>	<u>0.801</u>	<u>0.135</u>	0.202M	0.054s
Sine+PSC	1.0	37.251	0.969	0.044	28.025	0.599	0.273	0.204M	0.050s
Sine+PSC	1/2	37.507	0.968	0.036	29.094	0.627	0.262	0.204M	0.050s
Sine+PSC	1/4	39.846	0.980	0.019	30.308	0.691	0.159	0.204M	0.050s
Sine+PSC	1/8	<u>40.335</u>	<u>0.982</u>	<u>0.016</u>	31.143	0.704	0.147	0.204M	0.050s
Sine+PSC	1/16	38.629	0.966	0.041	30.305	0.669	0.158	0.204M	0.050s
CHOIR	1/8	41.316	0.992	0.014	32.605	0.847	0.130	0.204M	0.064s

are detailed in the supplementary materials, including additional high-resolution experiments on the full CAVE dataset ($512 \times 512 \times 31$).

4.4 Ablation Studies

Effectiveness of Core Components of Our Method. To validate the effectiveness of the core components of our method, we conduct ablation studies on the MSI Flowers dataset. As summarized in Table 4, introducing CHS alone yields a significant PSNR improvement over the sinusoidal activation baseline, demonstrating that its additive compositional structure effectively stabilizes deep optimization. While employing PSC alone, reducing the frequency scaling factor γ from 1.0 to 1/8 brings consistent gains, highlighting the superiority of embedding the $1/f$ power-law prior and frequency scaling factor γ . However, an excessively small γ (e.g., 1/16) impedes the model from capturing necessary high-frequency components, leading to performance degradation. Therefore, the experiments use $\gamma = 1/8$. The full CHOIR model, which integrates both CHS and PSC, achieves the best performance across all evaluation metrics, verifying the effective synergy between these two core components.

Comparative Analysis of Method Complexity. As shown in Figure 6 (a), we compare various INR-based methods on the HSI WDC dataset under random missing completion (OR=0.10) in terms of PSNR, model parameters (logarithmic scale), and runtime. The horizontal axis represents ($\log_{10}$) (Params. (M)), the vertical axis denotes PSNR (dB), and the size of each circle indicates runtime. The proposed CHOIR method achieves the highest PSNR with relatively few parameters and short runtime. In contrast, although CRNL [24] attains competitive PSNR, it requires an excessively large number of parameters and long runtime. Lighter methods such as SIREN [39], Gauss [32], and WIRE [35] remain in the low-parameter regime but yield lower PSNR performance. Overall, CHOIR offers a better trade-off among reconstruction quality, model parameter count, and computational efficiency.

Neural Tangent Kernel (NTK) Analysis. The Neural Tangent Kernel (NTK) characterizes the gradient similarity between different input coordinates. Its structure at initialization reflects the inductive bias of an untrained network as well as its ability to perform decoupled updates. The degree of diagonal dom-

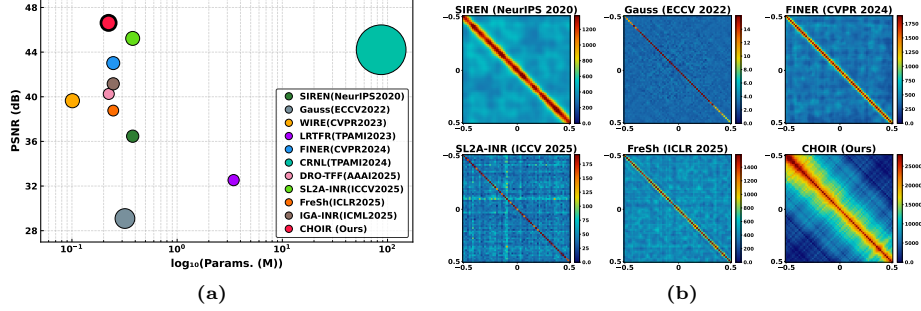


Fig. 6: (a) Comparison of various INR-based methods on HSI WDC dataset under data missing completion (random missing, $\text{OR}=0.10$). The size of each circle represents the runtime. (b) Visualization of NTK matrices at initialization for various INR methods.

inance in the NTK indicates the independence of gradient updates across spatial coordinates. A stronger diagonal with weaker off-diagonal elements (corresponding to a darker blue background in the figure) implies that gradient updates are more spatially decoupled, thereby enabling stronger fitting capacity, particularly for capturing high-frequency details. As shown in Fig. 6 (b), CHOIR exhibits the strongest diagonal dominance, accompanied by the darkest blue background. This observation theoretically validates that CHOIR possesses a more favorable inductive bias and superior high-frequency representation capability, which is consistent with the conclusions presented in Sec. 4.2. A more detailed introduction to the NTK matrix is provided in the supplementary materials.

5 Conclusion

We propose CHOIR to fundamentally address the optimization instability that has long plagued periodic INRs, which arises from their nested function composition architecture. With our novel Coordinated Harmonic Superposition (CHS) and Perceptual Spectrum Calibration (PSC) mechanisms, CHOIR enables stable training of deep periodic INRs and efficient data-adaptive spectrum learning. Comprehensive experiments demonstrate that CHOIR consistently outperforms state-of-the-art methods and sets new state-of-the-art benchmarks across a variety of multidimensional data recovery tasks. For future work, we aim to extend our experimental evaluations to unstructured or irregularly sampled data, such as point clouds, event-camera streams, and spatial transcriptomics, to further validate the generalization and effectiveness of our model.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant Nos. 62272313 and 62372302).

References

1. Bachlechner, T., Majumder, B.P., Mao, H., Cottrell, G., McAuley, J.: Rezero is all you need: Fast convergence at large depth. In: *Uncertainty in Artificial Intelligence*. pp. 1352–1361. PMLR (2021)
2. Benbarka, N., Höfer, T., Riaz, H.u.M., Zell, A.: Seeing implicit neural representations as fourier series. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. pp. 2041–2050 (2022)
3. Bengio, Y., Louradour, J., Collobert, R., Weston, J.: Curriculum learning. In: *Proceedings of the 26th annual International Conference on Machine Learning*. pp. 41–48 (2009)
4. Bengua, J.A., Phien, H.N., Tuan, H.D., Do, M.N.: Efficient tensor completion for color image and video recovery: Low-rank tensor train. *IEEE Transactions on Image Processing* **26**(5), 2466–2479 (2017)
5. Cai, Z., Zhu, H., Shen, Q., Wang, X., Cao, X.: Batch normalization alleviates the spectral bias in coordinate networks. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 25160–25171 (2024)
6. Eastman Kodak Company: Kodak lossless true color image suite. Dataset (1993), photo CD PCD0992
7. Fathony, R., Sahu, A.K., Willmott, D., Kolter, J.Z.: Multiplicative filter networks. In: *International Conference on Learning Representations* (2021)
8. Feng, H., Aldana, D., Novello, T., De Floriani, L.: Sasnet: Spatially-adaptive sinusoidal networks for inrs. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 41964–41973 (2026)
9. Field, D.J.: Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A* **4**(12), 2379–2394 (1987)
10. Glorot, X., Bengio, Y.: Understanding the difficulty of training deep feedforward neural networks. In: *Proceedings of the thirteenth international Conference on Artificial Intelligence and Statistics*. pp. 249–256. JMLR Workshop and Conference Proceedings (2010)
11. Jayasundara, D., Zhao, H., Labate, D., Patel, V.M.: Mire: Matched implicit neural representations. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 8279–8288 (2025)
12. Kania, A., Mihajlovic, M., Prokudin, S., Tabor, J., Spurek, P.: Fresh: Frequency shifting for accelerated neural representation learning. In: *The Thirteenth International Conference on Learning Representations* (2025)
13. Kazerouni, A., Azad, R., Hosseini, A., Merhof, D., Bagci, U.: Incode: Implicit neural conditioning with prior knowledge embeddings. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. pp. 1298–1307 (2024)
14. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014)
15. Kumar, M., Packer, B., Koller, D.: Self-paced learning for latent variable models. *Advances in Neural Information Processing Systems* **23** (2010)
16. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* **521**(7553), 436–444 (2015)
17. Li, J., Zhao, X., Wang, J., Wang, C., Wang, M.: Superpixel-informed implicit neural representation for multi-dimensional data. In: *European Conference on Computer Vision*. pp. 258–276. Springer (2024)

18. Li, L., Huang, W., Gu, I.Y.H., Tian, Q.: Statistical modeling of complex backgrounds for foreground object detection. *IEEE Transactions on Image Processing* **13**(11), 1459–1472 (2004)
19. Li, Y., Zhang, X., Luo, Y., Meng, D.: Deep rank-one tensor functional factorization for multi-dimensional data recovery. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. vol. 39(17), pp. 18539–18547 (2025)
20. Lindell, D.B., Van Veen, D., Park, J.J., Wetzstein, G.: Bacon: Band-limited coordinate networks for multiscale scene representation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 16252–16262 (2022)
21. Liu, J., Musialski, P., Wonka, P., Ye, J.: Tensor completion for estimating missing values in visual data. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **35**(1), 208–220 (2012)
22. Liu, Z., Zhu, H., Zhang, Q., Fu, J., Deng, W., Ma, Z., Guo, Y., Cao, X.: Finer: Flexible spectral-bias tuning in implicit neural representation by variable-periodic activation functions. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 2713–2722 (2024)
23. Luo, Y., Zhao, X., Li, Z., Ng, M.K., Meng, D.: Low-rank tensor function representation for multi-dimensional data recovery. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **46**(5), 3351–3369 (2023)
24. Luo, Y., Zhao, X., Meng, D.: Revisiting nonlocal self-similarity from continuous representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **47**(1), 450–468 (2024)
25. Luo, Y., Zhao, X., Meng, D.: Continuous representation methods, theories, and applications: An overview and perspectives. *arXiv preprint arXiv:2505.15222* (2025)
26. Mehta, I., Gharbi, M., Barnes, C., Shechtman, E., Ramamoorthi, R., Chandraker, M.: Modulated periodic activations for generalizable local functional representations. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. pp. 14214–14223 (2021)
27. Meng, Y., Jin, X., Lei, L., Guo, C.L., Li, C.: Ultraled: Learning to see everything in ultra-high dynamic range scenes. In: *Advances in Neural Information Processing Systems*. pp. 41466–41495 (2025)
28. 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)
29. Morsali, A., Vaez, M., Soltani, M., Kazerouni, A., Taati, B., Mohammad-Noori, M.: Staf: Sinusoidal trainable activation functions for implicit neural representation. *arXiv preprint arXiv:2502.00869* (2025)
30. Müller, T., Evans, A., Schied, C., Keller, A.: Instant neural graphics primitives with a multiresolution hash encoding. *ACM transactions on graphics (TOG)* **41**(4), 1–15 (2022)
31. 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)
32. Ramasinghe, S., Lucey, S.: Beyond periodicity: Towards a unifying framework for activations in coordinate-mlps. In: *European Conference on Computer Vision*. pp. 142–158. Springer (2022)
33. Rezaeian, R., Heidari, M., Azad, R., Merhof, D., Soltanian-Zadeh, H., Hachililoglu, I.: Sl2a-inr: Single-layer learnable activation for implicit neural representation. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. pp. 26065–26074 (2025)

34. Ricci, L., Perinelli, A., Prevedelli, M.: Nyquist–shannon sampling theorem. In: *The Physics Behind Electronics*, pp. 199–212. Springer (2024)
35. Saragadam, V., LeJeune, D., Tan, J., Balakrishnan, G., Veeraraghavan, A., Baraniuk, R.G.: Wire: Wavelet implicit neural representations. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 18507–18516 (2023)
36. Shabanov, A., Govindarajan, S., Reading, C., Goli, L., Rebain, D., Yi, K.M., Tagliasacchi, A.: Banf: Band-limited neural fields for levels of detail reconstruction. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 20571–20580 (2024)
37. Shi, K., Chen, H., Zhang, L., Gu, S.: Inductive gradient adjustment for spectral bias in implicit neural representations. In: *Forty-second International Conference on Machine Learning* (2025)
38. 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)
39. Sitzmann, V., Martel, J., Bergman, A., Lindell, D., Wetzstein, G.: Implicit neural representations with periodic activation functions. *Advances in Neural Information Processing Systems* **33**, 7462–7473 (2020)
40. Smith, J.O.: *Mathematics of the Discrete Fourier Transform (DFT): With Audio Applications*. Julius Smith, Stanford, 2nd ed. edn. (2007)
41. Sun, Y., Liu, J., Xie, M., Wohlberg, B., Kamilov, U.S.: Coil: Coordinate-based internal learning for tomographic imaging. *IEEE Transactions on Computational Imaging* **7**, 1400–1412 (2021)
42. 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)
43. Ulyanov, D., Vedaldi, A., Lempitsky, V.: Deep image prior. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 9446–9454 (2018)
44. Wang, Z., Simoncelli, E.P., Bovik, A.C.: Multiscale structural similarity for image quality assessment. In: *The thirty-seventh asilomar Conference on Signals, Systems & Computers*, 2003. vol. 2, pp. 1398–1402. IEEE (2003)
45. Xu, Z.Q.J., Zhang, Y., Luo, T., Xiao, Y., Ma, Z.: Frequency principle: Fourier analysis sheds light on deep neural networks. *arXiv preprint arXiv:1901.06523* (2019)
46. Yang, G., Benaim, S., Jampani, V., Genova, K., Barron, J., Funkhouser, T., Hariharan, B., Belongie, S.: Polynomial neural fields for subband decomposition and manipulation. *Advances in Neural Information Processing Systems* **35**, 4401–4415 (2022)
47. Yasuma, F., Mitsunaga, T., Iso, D., Nayar, S.K.: Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum. *IEEE Transactions on Image Processing* **19**(9), 2241–2253 (2010)
48. Yüce, G., Ortiz-Jiménez, G., Besbinar, B., Frossard, P.: A structured dictionary perspective on implicit neural representations. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 19228–19238 (2022)
49. Zhan, X., Jiang, R., Gupta, V., Swaminathan, T., Wang, Y., Zhang, G., Wang, H., Xu, M.: Microfm: Physics-guided flow matching for isotropic microscopy reconstruction. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 15639–15648 (2026)

50. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 586–595 (2018)
51. Zhao, C., Xu, Y., Chen, Z., Gu, E., Zhang, K., Liu, X., Yang, J., Tai, Y.: From zero to detail: A progressive spectral decoupling paradigm for uhd image restoration with new benchmark. *IEEE Transactions on Pattern Analysis and Machine Intelligence* pp. 1–18 (2026)
52. Zhou, T.W., Zhao, X.L., Wu, W.H., Wang, J.L., Luo, Y.S.: Frequency-aware implicit neural representation for multi-dimensional data recovery. *IEEE Transactions on Circuits and Systems for Video Technology* **35**(11), 10862–10874 (2025)