FLAIR: Frequency- and Locality-Aware Implicit Neural Representations

Supplementary Material

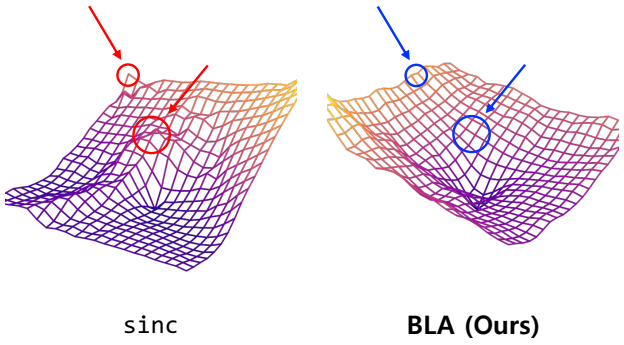


Figure 10. **Loss landscape visualizations on `sinc` and our proposed BLA.** Loss landscape visualizations show that Band-Localized Activation (BLA) produces a markedly flatter **MSE loss surface** than `sinc`. **Red arrows** indicate the sharp, irregular regions characteristic of `sinc`, whereas **blue arrows** highlight the smoother, low-curvature regions formed by BLA, leading to more stable and reliable optimization.

In this supplementary material, we first present the proposed Band-Localized Activation (BLA) in Sec. A, with ablations on its term components and initialization. Subsequently, in Sec. B, we introduce the Wavelet-Energy-Guided Encoding (WEGE) as a plug-and-play module compatible with other activations. We further perform ablations on its guided-filter score-map hyperparameters and present additional visualizations, including variants with different guided-filter settings and the no-WEGE variant. Additionally, we provide mathematical full derivations in Sec. C and per-scene results in Sec. D, including convergence efficiency analysis. Finally, Sec. E provides a comprehensive discussion of the applications and limitations of FLAIR.

A. Band-Localized Activation (BLA)

A.1. Loss Landscape Visualizations

As shown in Fig. 10, loss landscape visualizations show that Band-Localized Activation (BLA) has a flatter loss surface (MSE reconstruction) than `sinc`. A flatter landscape is widely associated with improved performance and better generalization [15, 23, 27]. Thus, BLA improves not only accuracy but also robustness in large data regimes.

To obtain the loss landscapes, we follow the perturbation-based analysis of Li *et al.* [27]. Specifically, given the optimized parameters θ, we sample two random filter-normalized directions d_1 and d_2 in the weight space. We then evaluate the perturbed models $\theta' = \theta + \alpha d_1 + \beta d_2$ over a 2D grid of (α, β) values, where both variables are uniformly sampled within a fixed range $[-1, 1]$ to form an $n \times n$ lattice of perturbations. For each lattice point, we compute the MSE reconstruction loss, yielding a faithful 2D slice of the loss landscape that enables a clear comparison of sharpness and stability.

Table 6. **Ablation of the proposed BLA activation and the WEGE module.** Columns progressively activate the four components of BLA: band-limiting (BL), transition-smoothing (TS), localization (LOC), and frequency-shifting (FS), forming the full BLA (BL × TS × LOC × FS). The final column additionally applies WEGE on top of BLA, corresponding to the full FLAIR model. All values are evaluated on DIV2K [2] Image 00.

	BL (base)	+ TS	+ LOC	BLA (full)	BLA + WEGE
PSNR ↑	27.31	28.18	30.22	32.96	**34.75**
SSIM ↑	0.8243	0.8581	0.9069	0.9380	**0.9663**
LPIPS ↓	0.101	0.087	0.065	0.034	**0.013**

A.2. Theoretical Foundations of Signal Reconstruction

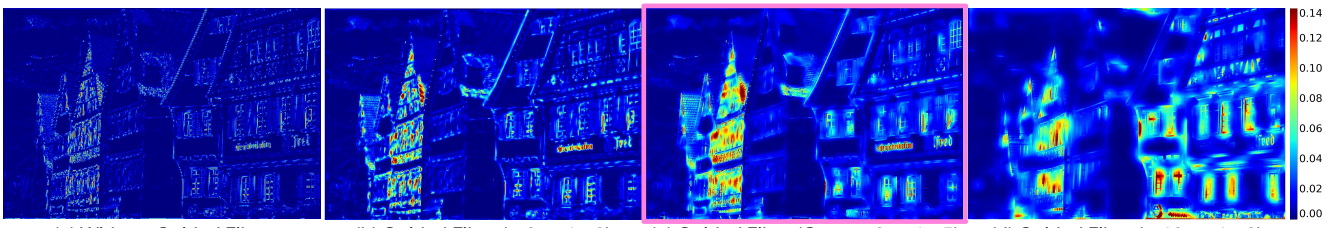
As introduced in Sec. C.2, Riesz signal reconstruction provides a general framework for stable signal representation. A Riesz basis must satisfy two conditions: the Riesz bounds and the partition-of-unity (PUC). Our BLA is constructed to meet both conditions, and therefore Riesz basis theory becomes the theoretical foundation for BLA. Furthermore, this theoretical foundation provides the basis for the BLA initialization scheme described in Sec. C.5.

A.3. Full Derivation of BLA

Within the Riesz basis framework (Sec. C.2), both the sinc function and our BLA satisfy the required Riesz bounds and partition-of-unity conditions. In deep learning optimization, however, the residual term in Eq. (2) is unavoidable, and this oscillatory component induces instability during training. This behavior is reflected in the perturbation experiments of Fig. 10, where sinc exhibits a distinctly non-flat loss landscape. The theoretical differences are formalized in the derivations of Sec. C.3 and Sec. C.4. Practically, the effect is evident in the MSE convergence of Fig. 3(a) and in Table 6, where BLA consistently improves over the pure band-limiting (BL) baseline across all metrics.

A.4. Ablation of BLA Components and WEGE Integration

In Table 6, each component yields consistent improvements. TS and LOC yield notable spatial localization gains. Moreover, the interaction between FS and WEGE provides a strong and complementary synergy, enabling more accurate modeling of the target signal.



(a) Without Guided Filter (b) Guided Filter (r=2, ϵ=1e-8) (c) Guided Filter (Ours, r=6, ϵ=1e-5) (d) Guided Filter (r=10, ϵ =1e-2)

Figure 11. **Guided filtering effects on WEGE score-map visualizations.** r is the window radius and ϵ is the regularization parameter, and larger r together with larger ϵ yields a smoother filter in (d).

(a) Ground-truth (b) Guided Filter (r=2, ϵ=1e-8) (c) Guided Filter (Ours, r=6, ϵ=1e-5) (d) Guided Filter (r=10, ϵ =1e-2)

Figure 12. **Qualitative effects of guided filtering and WEGE in the RGB domain.** (c) follows our guided-filter setting ($r = 6$, $\epsilon = 1\text{e}{-5}$), preserving fine details without noise and is closest to (a) ground truth.

B. Wavelet-Energy-Guided Encoding (WEGE)

B.1. Effect of Guided Filtering on WEGE Scores

In Fig. 11, we present score maps produced with different radius r and regularization ϵ settings to clarify our chosen hyperparameters, and we show their resulting effects on the final RGB outputs in Fig. 12.

As shown in Fig. 11, (a) presents the raw WEGE score without guided filtering, while (b)–(d) apply guided filtering with different radius r and regularization ϵ. Smaller r and ϵ preserve sharper local structures, whereas larger r and ϵ progressively smooth the score responses. The score discontinuities produced by sharp filtering lead to salt-and-pepper–like artifacts, as shown in Fig. 12(b), while overly smoothed scores result in blurry roof boundaries, as shown in Fig. 12(d). In contrast, the chosen hyperparameters in (c) yield a balanced score representation that preserves fine details while suppressing artifacts. This choice is further validated across diverse scenes and tasks by the quantitative results in Table 5 of the main paper.

B.2. Plug-and-Play Behavior of WEGE

We evaluate WEGE in a plug-and-play setting by integrating it into baseline activations like ReLU and SIREN. Fig. 13 shows that WEGE enhances fine-detail reconstruction for the image fitting task and provides consistent improvements in PSNR, SSIM [57], and LPIPS [63]. These results indicate that WEGE refines the embedding space from $f(x, y)$ to $f(x, y, \tilde{w}_b)$ through its frequency guidance, enabling more expressive and precise structural representations. Beyond fitting, WEGE extends to diverse tasks. Fig. 14 demonstrates its effectiveness on the super-resolution task, where frequency-aware refinement mitigates frequency leakage and improves high-frequency structure reconstruction.

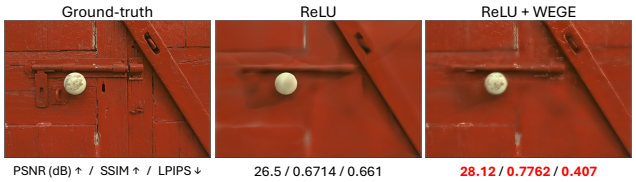


Figure 13. **Plug-and-play WEGE on ReLU.** WEGE substantially improves both the visual and quantitative quality of ReLU-based INRs.

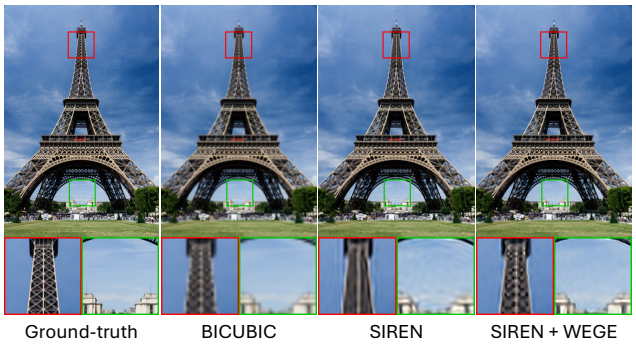


Figure 14. **Plug-and-play WEGE on SIREN for x4 super-resolution.** WEGE enhances SIREN's fine details (red box) while suppressing noise (green box) through explicit frequency guidance.

C. Theoretical Foundations for BLA

C.1. Classical Sampling Theory

The primary objective of classical sampling theory is the reconstruction of a continuous signal from its discrete samples. This objective conceptually aligns with Implicit Neural Representations (INRs), which map low-dimensional coordinates to signal values while learning a continuous representation from discrete data.

Classical sampling theory specifies when a bandlimited signal can be perfectly recovered from its discrete samples. A signal f is T-bandlimited if its Fourier transform $\hat{f}(s)$ satisfies $\hat{f}(s) = 0$ for all $|s| > T$. This definition makes the notion of perfect reconstruction well posed. In particular, the Nyquist-Shannon sampling theorem establishes that a T-bandlimited signal $f \in L^2(\mathbb{R})$ can be perfectly reconstructed as:

$$f(x) = \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2T}\right) \mathtt{sinc}\left(2Tx - n\right), \tag{8}$$

where T denotes the maximum frequency of the signal $f(x)$. The quantity $2T$ is the Nyquist sampling rate, defined as twice this maximum frequency. Its reciprocal, $1/(2T)$, specifies the sampling interval, and the discrete samples $f(n/(2T))$ serve as the coefficients used in the reconstruction formula. The term $\mathtt{sinc}(2Tx-n)$ denotes the sinc interpolation kernel scaled by the Nyquist rate $2T$ and shifted according to the sample index n.

In practice, however, classical sampling theory faces inherent limitations due to the characteristics of real-world signals. Although it guarantees theoretically perfect reconstruction under the bandlimited assumption, achieving this requires an infinite number of samples, which is infeasible in practical settings. These limitations necessitate a more general signal reconstruction framework, for which Riesz basis theory provides the appropriate mathematical framework.

C.2. Riesz Basis Theory

Riesz basis theory provides a generalized mathematical framework that overcomes the practical limitations of classical sampling theory. In this context, the Nyquist–Shannon sampling theorem, which represents a specific instance of classical sampling theory, is limited to the $\mathtt{sinc}$ function and strictly bandlimited signals. In contrast, Riesz basis theory generalizes this framework by enabling reconstruction with a much broader class of basis functions. The core concept of the Riesz basis is to specify the conditions under which the set of shifted functions $\{F(x-k)\}_{k\in\mathbb{Z}}$ forms a stable and complete representation system within a specified signal space.

Specifically, the Riesz basis framework defines the T-scaled signal space for $T > 0$ as follows:

$$V_T(F) = \left\{ s_T(x) = \sum_{k\in\mathbb{Z}} a_T(k)\, F\Big(\frac{x}{T} - k\Big) : a_T \in l^2(\mathbb{R}) \right\}, \tag{9}$$

where T determines the spatial scale of the signal space. As T decreases, the basis function F becomes spatially compressed, enabling the representation of higher-frequency components. Conversely, larger values of T expand F, capturing lower-frequency structures. Here, $s_T(x)$ denotes an arbitrary signal belonging to the space $V_T(F)$. This signal is expressed as a linear combination of the basis functions $F\left(\frac{x}{T} - k\right)$. The term $a_T(k)$ represents the k-th coefficient multiplied by each basis function, where the sequence a_T is an element of $\ell^2(\mathbb{R})$, the Hilbert space of square-summable real sequences.

In particular, for the T-scaled signal space $V_T(F)$ to operate as a signal reconstruction system, the set of shifted functions $\{F_k = F(x-k)\}_{k\in\mathbb{Z}}$ must form a Riesz basis. A Riesz basis must satisfy the following two conditions:

1. $A\|a\|_{l^2}^2 \le \left\|\sum_{k\in\mathbb{Z}} a(k)F_k\right\|^2 \le B\|a\|_{l^2}^2,\ \forall a(k) \in \ell^2(\mathbb{R}).$ (10)

2. $\sum_{k\in\mathbb{Z}} F(x+k) = 1,\ \forall x \in \mathbb{R}.$ (11)

The first condition ensures a stable representation for the linear combination of the basis functions, where A and B are positive constants satisfying $0 < A \le B < \infty$. The lower bound of this inequality guarantees the linear independence of the basis functions $\{F_k\}_{k\in\mathbb{Z}}$. The upper bound ensures that the L^2-norm of any signal $s_T \in V_T(F)$ is finite, thereby establishing $V_T(F)$ as a subspace of $L^2(\mathbb{R})$.

The second condition, known as the Partition of Unity Condition (PUC), ensures that the Riesz basis can approximate general $L^2(\mathbb{R})$ signals beyond the space $V_T(F)$. When this condition is satisfied, selecting an appropriate scale T enables the approximation of arbitrary $L^2(\mathbb{R})$ signals with arbitrary precision.

Riesz basis theory guarantees that a set of shifted basis functions $\{F(x-k)\}_{k\in\mathbb{Z}}$ that forms a Riesz basis can approximate any signal in $L^2(\mathbb{R})$ with arbitrary precision. Specifically, for any signal $s \in L^2(\mathbb{R})$ and any arbitrarily small positive value $\epsilon > 0$, there exists an appropriate scale $T > 0$ and a function $f_T \in V_T(F)$ such that the approximation error is bounded by $|s - f_T|_{L^2} < \epsilon$. This guarantee is attributed to the PUC, which performs a critical role in bridging the gap between the T-scaled space $V_T(F)$ and the general $L^2(\mathbb{R})$ space.

C.3. Gradient Derivation of Sinc-activated INR

The $\mathtt{sinc}$ function is a canonical example of a function that forms a Riesz basis, satisfying the two required conditions. This theoretical foundation suggests that the $\mathtt{sinc}$ function

is sufficient for optimal signal reconstruction. However, this theoretical guarantee establishes only the existence of a function $f_T \in V_T(F)$ that can approximate any signal in $L^2(\mathbb{R})$ with arbitrary precision, and it does not provide a specific methodology for practically obtaining the f_T. In particular, when the sinc function is employed as the activation function within the INR framework for signal reconstruction, its global oscillatory property can induce training instability. This instability can be mathematically analyzed by deriving the gradient of the sinc function.

The output of the INR network $f(x; W)$ for a single input coordinate $x \in \mathbb{R}^{d_{\text{in}}}$ is expressed as follows, omitting biases and the output layer weights for simplification:

$$f(x; W) = \sum_{i=1}^{d_{\text{out}}} \frac{1}{T} F\left(\frac{W_i x}{T}\right), \tag{12}$$

where $W \in \mathbb{R}^{d_{\text{out}} \times d_{\text{in}}}$ is the weight matrix of the hidden layer, $F(\cdot)$ is the activation function, and $T > 0$ represents the frequency scaling parameter. The loss function utilized for signal reconstruction is defined as the L^2 loss between the target signal g and the network output f:

$$\mathcal{L}(W) = \|g - f\|_2^2 = \int |g(x) - f(x; W)|^2 dx. \tag{13}$$

For notational simplicity, the continuous representation is employed in the subsequent derivation instead of the discrete representation. The gradient of the loss function with respect to the weight W_{ij} is derived by applying the chain rule, as follows:

$$\frac{\partial \mathcal{L}}{\partial W_{ij}} = -2 \int [g(x) - f(x; W)] \cdot \frac{\partial f(x; W)}{\partial W_{ij}} dx. \tag{14}$$

Thus, the gradient of the INR using the sinc activation function is derived as follows:

$$\frac{\partial \mathcal{L}}{\partial W_{ij}} = -\frac{2}{\pi} \int \frac{g(x) - f(x)}{(W_i x)^2} \left[\pi u \cos(\pi u) - \sin(\pi u)\right] x_j \, dx, \tag{15}$$

where $u = W_i x / T$, and $x_j \in \mathbb{R}$ is the value corresponding to the j-th dimension. The derived gradient expression indicates two major properties that contribute to the training instability of INRs using the sinc activation function. The derived gradient expression explains two major causes of training instability in sinc-activated INRs.

First, the term $[\pi u \cos(\pi u) - \sin(\pi u)]$ exhibits oscillatory behavior with respect to $u = W_i x / T$. This behavior causes the gradient's sign to frequently change across the input space, leading to an unstable convergence path during optimization. Second, the sinc function is defined as $\text{sinc}(x) = \sin(x)/x$ and exhibits non-zero values theoretically over an infinite range. Consequently, the gradient for each weight W_{ij} is influenced by the error $[g(x) - f(x)]$ across the entire input domain.

This analysis suggests the motivation for establishing an activation function that can ensure practical training stability while preserving the theoretical guarantees of the Riesz basis. Specifically, functions with compact support have non-zero values only within a limited spatial region, a property which allows them to effectively mitigate the issues of global oscillation and residual error propagation resulting from the global support of the sinc function.

C.4. Gradient Derivation of BLA-activated INR

To ensure training stability while maintaining the theoretical guarantees of the Riesz basis, we propose a novel activation function, the Band-Localized Activation (BLA), which incorporates a localization term. The gradient of the BLA is derived as follows, based on Eq. (14):

$$\frac{\partial \mathcal{L}}{\partial W_{ij}} = -\frac{2}{\pi} \int \frac{[g(x) - f(x)] \cdot x_j}{(W_i x)^2} \cdot \exp\left(-\frac{u^2}{\sigma^2}\right) \cdot H(u) \, dx, \tag{16}$$

$$H(u) = \pi u \cos(\pi u) - \sin(\pi u) \left(1 + \frac{2u^2}{\sigma^2}\right). \tag{17}$$

This gradient expression clearly demonstrates two key mechanisms by which the BLA function enhances training stability compared to the sinc function.

First, the exponential decay term $\exp\left(-\frac{u^2}{\sigma^2}\right)$ exponentially reduces the contribution from regions distant from the origin in the input space. This ensures that the gradient is dominated by the error in local regions, thereby effectively mitigating global residual error propagation. Second, the exponential decay term spatially restricts the overall magnitude of the gradient. This exponentially attenuates the effective contribution of the oscillatory term $H(x)$ on the gradient in regions distant from the origin. The oscillatory term itself remains, but its influence is restricted to the local domain, thus mitigating instability during the training process.

This gradient analysis shows that the BLA function ensures training stability while maintaining the theoretical signal reconstruction capability of the sinc function. To formally establish the theoretical foundation of the BLA within the Riesz basis framework, the proof of whether BLA satisfies the Riesz basis conditions is required. This proof will enable establishing the strategy to initialize the parameters of the BLA.

C.5. BLA Initialization Scheme under the Riesz Basis Framework

The BLA function must satisfy the Riesz basis condition to guarantee the theoretical foundation for signal reconstruction. Specially, this requires simultaneous adherence to two essential conditions: the stability (Eq. (10)) and the partition of unity condition (Eq. (11)).

To verify the stability condition (Eq. (10)), we reformulate it as follows using the Poisson summation formula [50]:

$$A \leq \sum_{k\in\mathbb{Z}} |\hat{F}(\xi + 2k\pi)|^2 \leq B, \tag{18}$$

where the lower bound $A > 0$ guarantees the linear independence of the basis functions. The upper bound $B < \infty$ ensures numerical stability by preventing excessive overlap in the frequency domain. Here, $\hat{F}$ denotes the Fourier Transform of the BLA function and ξ is the frequency variable. Subsequently, the Fourier transform $\hat{F}(\xi + 2k\pi)$ of the BLA function is derived as follows:

$$\hat{F}(\xi + 2k\pi) = \frac{\sigma}{\sqrt{2\pi}} \int_{-\pi/T}^{\pi/T} e^{-\sigma^2(\xi+2k\pi-\zeta)^2/2} d\zeta. \tag{19}$$

Subsequently, we use the following inequality to derive the lower bound A in Eq. (19) [50]:

$$\sum_{k\in\mathbb{Z}} \left|\hat{F}(\xi + 2k\pi)\right|^2 \geq \left|\hat{F}(\xi)\right|^2. \tag{20}$$

In this equation, $|\hat{F}(\xi)|$ serves as the lower bound of $|\hat{F}(\xi + 2k\pi)|$, which is A in Eq. (18). Furthermore, since the integrand in Eq. (19) decays exponentially for $k \neq 0$, this ensures that the upper bound B in Eq. (18) is finite. Consequently, the existence of A and B satisfying $0 < A \leq B < \infty$ verifies that the BLA function satisfies the stability condition (Eq. (18)).

We verify the Partition of Unity Condition (PUC). By the Poisson summation formula, Eq. (11) is reformulated as follows:

$$\sum_{n\in\mathbb{Z}} \hat{F}(2\pi n) e^{2\pi i n x} = 1. \tag{21}$$

This condition requires that $\hat{F}(0) = 1$ for $n = 0$ and $\hat{F}(2\pi n) = 0$ for $n \neq 0$.

We now verify that the BLA function satisfies the PUC by expanding $\hat{F}(2\pi n)$. The Fourier transform $\hat{F}(2\pi n)$ of the BLA function can be expanded as follows:

$$\hat{F}(2\pi n) = \frac{T\sigma}{\sqrt{2\pi}} \int_{-\pi/T}^{\pi/T} e^{-\sigma^2(2\pi n-\zeta)^2/2} d\zeta. \tag{22}$$

Note that $\hat{F}(2\pi n)$ is a function of T and σ. We now determine the range of T and σ satisfying $\hat{F}(0) = 1$ and $\hat{F}(2\pi n) = 0$ to prove that the BLA function satisfies the PUC. To this end, we compute $\hat{F}(0)$ from Eq. (22):

$$\hat{F}(0) = \frac{T}{\sqrt{\pi}} \int_{-\sigma\pi/(T\sqrt{2})}^{\sigma\pi/(T\sqrt{2})} e^{-u^2}\, du = T \cdot \mathtt{erf}\left(\frac{\sigma\pi}{T\sqrt{2}}\right) = 1, \tag{23}$$

where $u = \sigma\zeta/\sqrt{2}$, and $\mathtt{erf}(\cdot)$ denotes the error function, which satisfies $|\mathtt{erf}(\cdot)| < 1$.

We now derive the constraint requiring $\hat{F}(2\pi n) = 0$ for $n \neq 0$. Since the integral in Eq. (22) cannot be computed analytically for all $n \in \mathbb{Z}$, we derive its upper bound as follows:

$$\hat{F}(2\pi) = \sigma\sqrt{2\pi} \cdot e^{-\frac{1}{2}\sigma^2(2\pi-\frac{\pi}{T})^2} = 0. \tag{24}$$

The convergence of Eq. (24) to zero directly ensures that $\hat{F}(2\pi n) \to 0$ for all $n \in \mathbb{Z}$ by definition of the upper bound.

We have derived two constraints for the BLA function to satisfy the PUC: Eq. (22) and Eq. (23). Since each constraint depends on T and σ, we define the following cost function for these constraints:

$$E(T, \sigma) = |1 - \hat{F}(0)| + |\hat{F}(2\pi)|. \tag{25}$$

The theoretical optimal values of T and σ for the BLA to form a Riesz basis through Eq. (25). From Eq. (23) with $|\mathtt{erf}(\cdot)| < 1$, we obtain $T \geq 1$. To maximize the frequency bandwidth, we set $T = 1$ since smaller T values yield wider bandwidths. We then choose $\sigma = 2$, the smallest value at which the rate of decrease in Eq. (25) becomes negligible, thereby maximizing both spatial localization and frequency bandwidth.

Table 7. **Arbitrary-scale super-resolution comparison.** Quantitative results on the Kodak dataset at $\times 6$ and $\times 8$ upscaling. **Red** and blue indicate best and second-best values.

Methods	×6			×8		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
FINER (CVPR'24)	23.94	0.6495	0.425	22.57	0.5761	0.549
WIRE (CVPR'23)	22.29	0.4770	0.273	19.12	0.3191	0.392
GAUSS (ECCV'22)	24.24	0.6601	0.382	22.84	0.6051	0.492
MFN (ICLR'21)	20.64	0.3066	0.594	18.93	0.5761	0.667
SIREN (NeurIPS'20)	24.18	0.6568	0.384	22.84	0.5801	0.443
ReLU+P.E. (NeurIPS'20)	22.51	0.6073	0.398	21.85	0.5870	0.415
FLAIR (Ours)	**24.32**	**0.6786**	**0.246**	**23.92**	**0.6365**	**0.339**

D. Additional Results Across Diverse Tasks

D.1. Signal Representation

2D Image Representation. As shown in Figs. 24 and 25, FLAIR produces superior per-scene reconstruction quality, preserving fine details while modeling RGB values with higher fidelity and minimal noise. Although Table 1 in the main paper summarizes the overall performance, we additionally provide full per-scene results for DIV2K and Kodak in Tables 11 and 12 for completeness.

3D Shape Representation. As shown in Fig. 23, FLAIR accurately reconstructs diverse 3D shapes, demonstrating robustness and generality beyond data-specific biases. Together with our 2D results, FLAIR consistently outperforms prior methods across both 2D and 3D representation tasks, highlighting its strong scalability across modalities.

D.2. Image Restoration Tasks

Image Denoising. Image denoising is challenging because effective noise removal must be achieved without degrading fine structures. As shown in Fig. 22, FLAIR achieves both strong noise suppression and faithful detail preservation. FLAIR achieves both by leveraging the joint effect of band-limiting and spatial localization, rather than merely focusing on expressiveness as other baselines do.

Arbitrary-Scale Super-Resolution. Conventional SR networks rely on scale-specific upsampling modules, which constrain them to predefined output resolutions. In contrast, INRs represent signals as continuous coordinate-based mappings, inherently operating as resolution-independent interpolators. This enables super-resolution beyond fixed scales and supports seamless generation at arbitrary resolutions. Leveraging this capability, we demonstrate that FLAIR maintains strong performance even at extreme scales. As shown in Table 7, FLAIR consistently outperforms existing methods across challenging upscaling factors.

D.3. Neural Radiance Fields

In Fig. 10, we visualize additional per-scene results for the novel view synthesis task. In the first row, FLAIR recovers

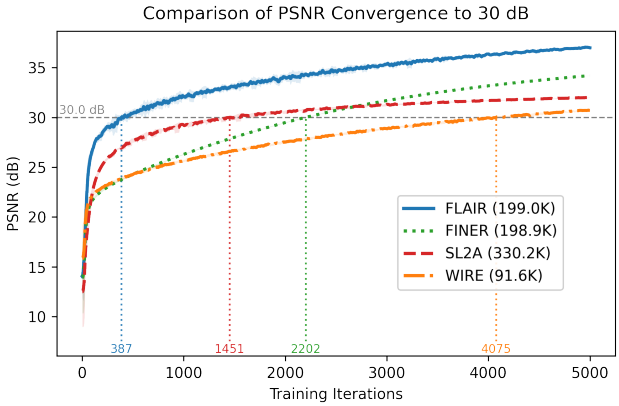


Figure 15. **FLAIR achieves the fastest convergence to 30 dB.** FLAIR reaches the target quality with significantly fewer training iterations, indicating high training efficiency, and continues to improve to outperform all baselines in the final results. This advantage arises from the synergy between WEGE's frequency-aware guidance and BLA's adaptive representation, which allows the model to capture both low and high frequencies as needed.

fine-grained structures by selectively capturing the required frequency components and, through its frequency shifting mechanism, effectively reaching the desired high-frequency details. In the second, third rows, FLAIR maintains stable rendering quality even under a challenging setting where the model trains on only 25 views and renders 100 novel views. Under this hard regime, FLAIR suppresses noise, prevents structural leakage, and produces clean textures, whereas competing methods exhibit visible degradation. The qualitative observations are in line with the quantitative results. As shown in Table 26, FLAIR generally achieves consistent performance improvements across the per-scene evaluations with stable and reliable gains.

For SL2A [42], the method adjusts polynomial degrees and low-rank linear layer ranks depending on the target task. However, SL2A does not provide NeRF-specific implementation details, and the official GitHub repository places the entry "Add Novel View Synthesis (NeRF) code" in the to-do list rather than providing an actual implementation. As of November 21, no operational NeRF implementation exists. We attempted several reasonable configurations, but training remained unstable, often resulting N/A outputs, so we do not report SL2A for the NeRF experiment. For a fair and comprehensive comparison, we also include the plug-and-play approach IGA [47]. Using IGA alone does not yield meaningful performance. Therefore, we combine IGA with a positional encoding module of comparable parameter size and report the resulting scores.

Table 8. **Quantitative comparison on the Tokyo dataset (fitting) and the LLFF dataset (novel view synthesis).** FLAIR (Small) achieves competitive performance with explicit methods while using significantly fewer parameters and a smaller model size. FLAIR (Large) achieves improved performance and attains the best results across both datasets.

	Params (M) ↓	Size (MB)↓	Steps (k)	Time↓	PSNR↑	SSIM↑	LPIPS↓
Tokyo Dataset (Task: Fitting)							
Instant-NGP	3.67 M	10.54	30k	**3.4m**	28.38	0.8240	0.208
NFFB	2.68 M	10.21	30k	30.1m	26.54	0.7675	0.234
Ours (Small)	**0.20** M	**1.51**	25k	6.2m	27.74	0.8554	0.206
Ours (Large)	3.25 M	10.19	35k	26.5m	**34.70**	**0.9525**	**0.061**
LLFF Dataset (Task: NVS)							
Instant-NGP	13.00 M	24.89	30k	**6.4m**	30.20	0.9524	**0.030**
NFFB	3.89 M	47.11	30k	42.4m	26.94	0.8852	0.258
NeRF	1.19 M	4.55	50k	35.1m	26.80	0.8794	0.256
Ours (Small)	**0.79** M	**3.03**	30k	14.0m	28.24	0.8846	0.242
Ours (Large)	2.55 M	9.74	37k	25.4m	**31.14**	**0.9566**	0.032

D.4. Extended Comparison with Explicit Methods

In Fig. 3 of the main paper, the visual improvements over recent methods [42] appear limited due to saturation in the fitting task, where SSIM approaches 1.0. To better highlight the superiority of our model, we further evaluate on a more challenging high-resolution dataset, Tokyo [31]. As shown in Fig. 20, the yellow box indicates configurations with matched layers and parameter budgets compared to prior INR methods, where our approach demonstrates strong performance. We further compare against explicit methods [35, 58], achieving competitive results with 10× lower model capacity and computational cost, as summarized in Table 8.

We further extend our evaluation on Neural Radiance Fields beyond standard synthetic benchmarks to the more realistic LLFF dataset [33], as shown in Fig. 21. As illustrated in the figure, our method reconstructs finer details compared to both implicit and explicit methods (green arrows), demonstrating its broad applicability.

D.5. Convergence Efficiency and Representation Compactness

Convergence Efficiency. Fig. 15 evaluates how rapidly each method reaches a target quality of 30 dB. FLAIR converges 4× to 11× faster than competing methods, and also achieves the best performance even at the final 5000 iterations. This demonstrates both its high training efficiency and its superior reconstruction accuracy.

Representation Compactness. FLAIR additionally exhibits strong representation compactness. Fig. 17 visualizes the learned neurons via fast Fourier transform. While the SOTA baseline FINER concentrates most neurons around similar low-frequency regions, revealing substantial redundancy, FLAIR learns diverse and distinct spectral patterns thanks to its band-limiting and frequency-shifting mechanisms. This diversity directly translates into robustness under reduced model capacity. As shown in Fig. 18, FINER degrades significantly when the hidden dimensionality is reduced, since its neurons do not cover sufficiently diverse frequency bands. In contrast, FLAIR continues to maintain comparable accuracy even when its model capacity is reduced to nearly half of that of FINER, as shown in Fig. 18. This indicates that FLAIR achieves a similar level of reconstruction quality with substantially fewer parameters and FLOPs.

NFFB (CVPR'23) — *Tokyo dataset*			
Method	PSNR↑	SSIM↑	LPIPS↓
SIREN (Baseline)	26.54	0.7675	0.234
SIREN → BLA	27.05 (+0.51)	0.7824 (+0.0149)	0.221 (-0.013)
Scaffold-GS (CVPR'24) — *Mip-NeRF360 bicycle*			
Method	PSNR↑	SSIM↑	LPIPS↓
ReLU (Baseline)	24.44	0.6979	0.304
ReLU → BLA	24.78 (+0.34)	0.7144 (+0.0165)	0.285 (-0.019)

Table 9. **BLA integrates into arbitrary MLP architectures.** Replacing SIREN in NFFB and ReLU in the two-layer MLP of Scaffold-GS with BLA consistently improves performance.

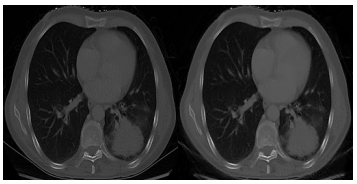


Figure 16. **Computed tomography reconstruction.** FLAIR achieves high-quality reconstruction in sparse-view CT, effectively recovering continuous structures in medical imaging.

E. Applications and Limitations of FLAIR

E.1. Applications

FLAIR exhibits broad applicability across diverse architectures and tasks. Rather than relying on complex designs, FLAIR achieves simultaneous localization and frequency selectivity using a single nonlinear activation, BLA. This allows BLA to integrate into arbitrary MLP architectures, such as NFFB and Scaffold-GS, and improves performance, as shown in Table 9. Moreover, WEGE acts as a plug-and-play module compatible with various activation functions and enhances fine details, as illustrated in Fig. 13 and Fig. 14. By jointly using BLA and WEGE within an INR framework, FLAIR models continuous signals and is well-suited for domains involving continuous data, such as medical imaging (Fig. 16).

E.2. Limitations

While INRs provide a compact and versatile representation that extends across diverse tasks, they also exhibit intrinsic limitations that FLAIR inherits. First, although FLAIR improves frequency selectivity and achieves strong quantitative results, as shown in Fig. 19, it still struggles under extreme degradation scenarios in super-resolution. In such challenging regimes, diffusion-based approaches [20]

demonstrate markedly higher stability, successfully reconstructing up to $\times 16$ magnification, where FLAIR leaves room for improvement. Second, despite being lightweight in structure, INRs are constrained to per-scene optimization, meaning that each new scene demands a full training procedure. This stands in contrast to feed-forward architectures [8, 36], which generalize across scenes and, once trained, require only a single forward pass for inference. Future work should focus on developing INR formulations that retain compactness while achieving stronger generalization and robustness across tasks and degradation levels.

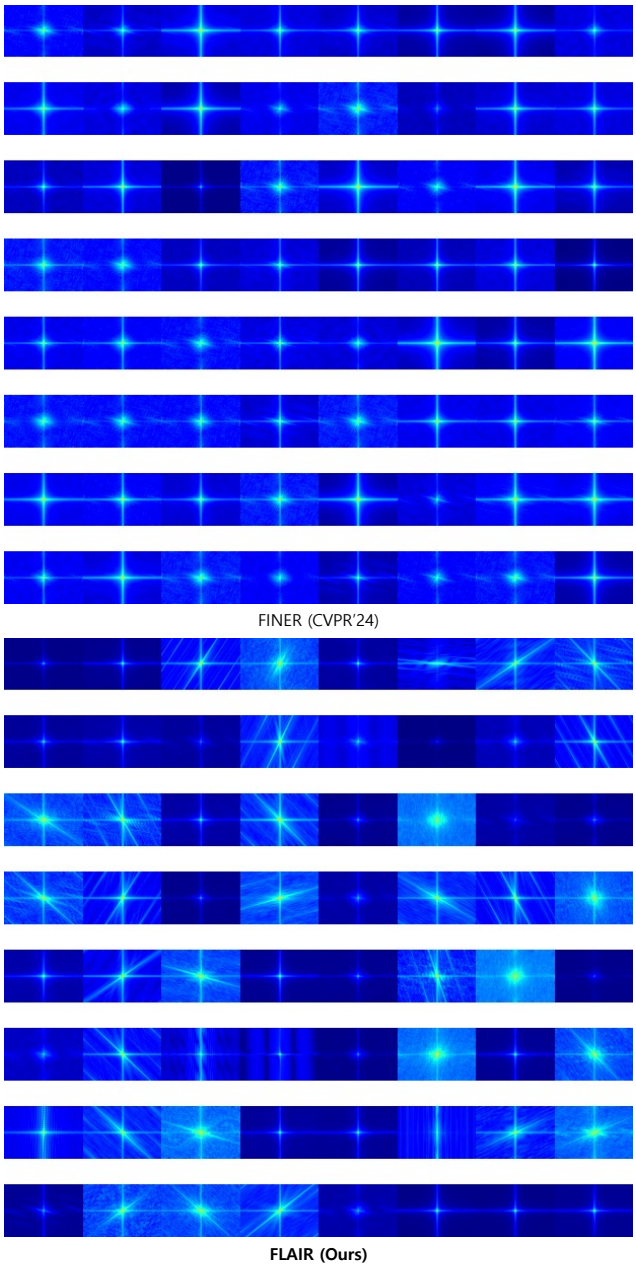


Figure 17. **Fast fourier transform on learned neurons.** Compared to the SOTA baseline FINER, whose frequency responses remain highly concentrated around the low-frequency center, FLAIR (ours) exhibits notably more diverse spectra. This indicates that FLAIR enables each neuron to learn the frequency components that are genuinely required.

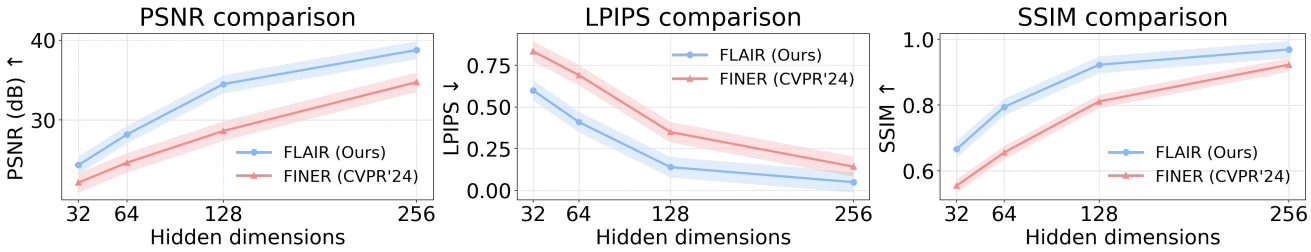


Figure 18. **Comparison with the SOTA sinusoidal-based model FINER across multiple hidden dimensions.** FLAIR achieves comparable performance to FINER while using approximately half the hidden dimensionality for the same evaluation scores. Since hidden dimension directly correlates with parameter count and FLOPs, this demonstrates that FLAIR operates significantly better under constrained model budgets.

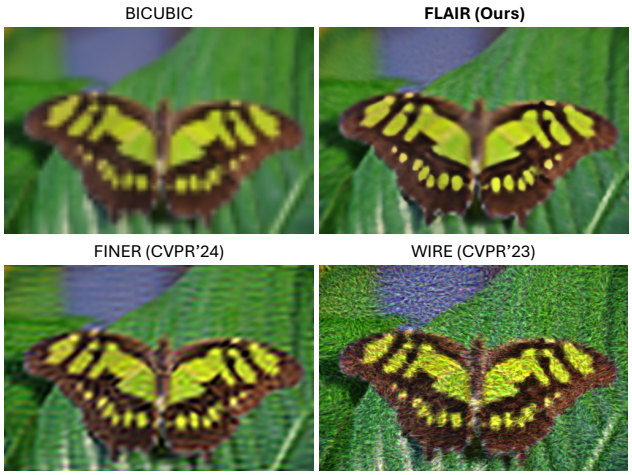


Figure 19. **Super-Resolution under Extreme Degradation.** Qualitative results on a challenging $\times 8$ setting, showing that FLAIR suppresses noise more effectively than other methods but still struggles under severe degradation.

Figure 20. **Visual comparison on a challenging gigapixel image (6144×2324).** FLAIR (Small) outperforms other implicit methods (yellow), and FLAIR (Large) produces more faithful results than explicit methods (green) at comparable computational cost.

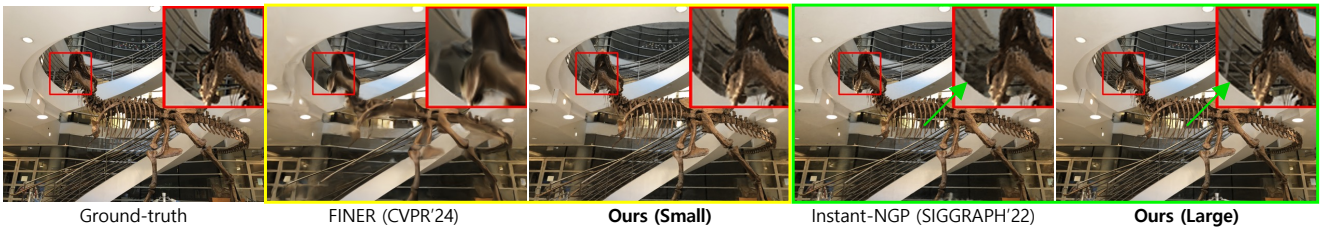


Figure 21. **Per-scene qualitative results on the LLFF dataset for neural radiance field reconstruction.** Beyond standard synthetic INR benchmarks, FLAIR reconstructs novel views while preserving fine details.

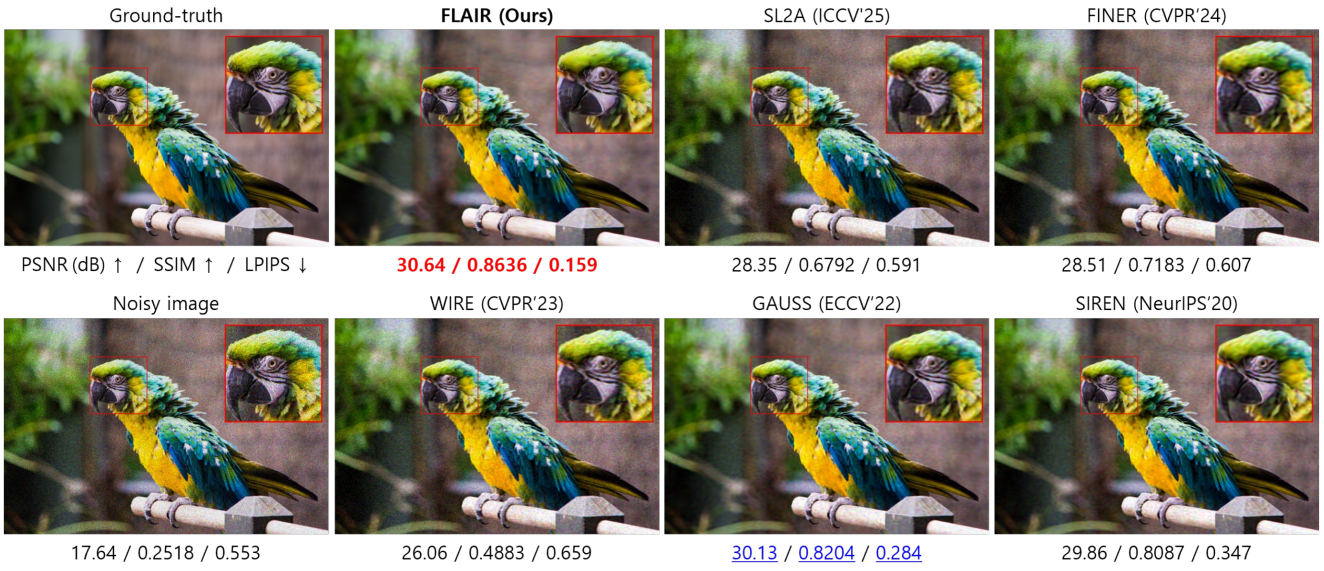


Figure 22. **Visual comparisons between FLAIR and baselines on the denoising task.** FLAIR outperforms the baselines by effectively suppressing noise while preserving fine details.

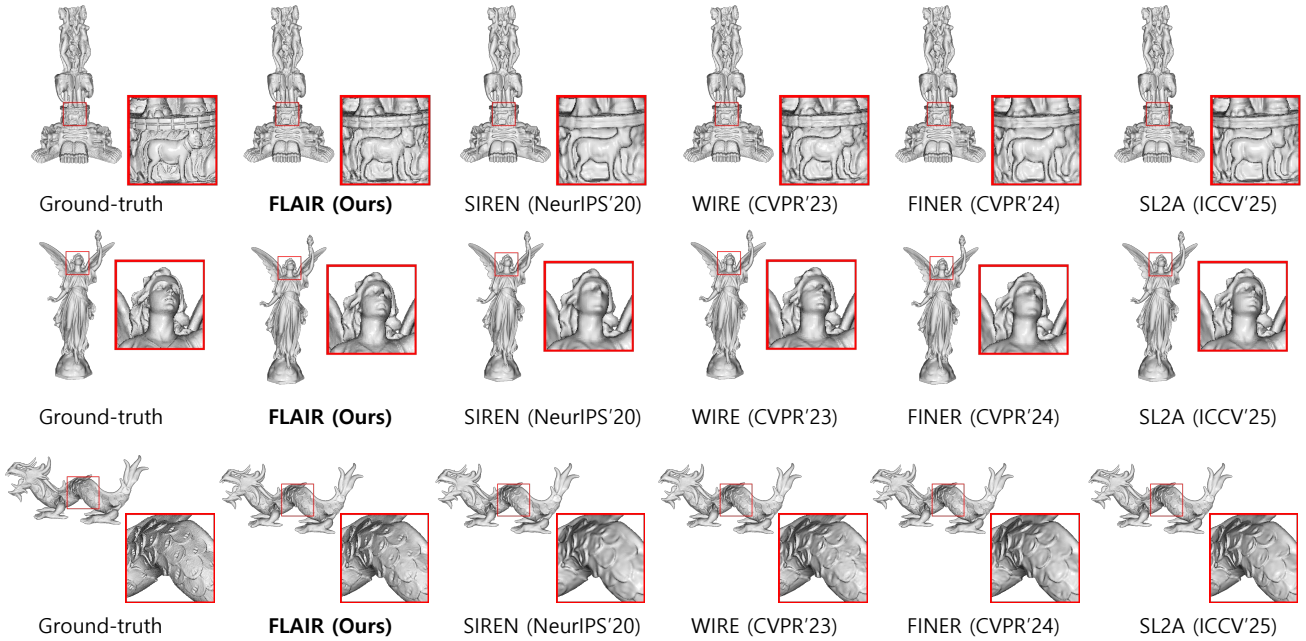


Figure 23. **Per-scene visual comparisons on signed distance field reconstruction.** FLAIR consistently outperforms other methods across all scenes, recovering finer geometry and sharper surface details.

Figure 24. **Image fitting results on DIV2K image 13.**

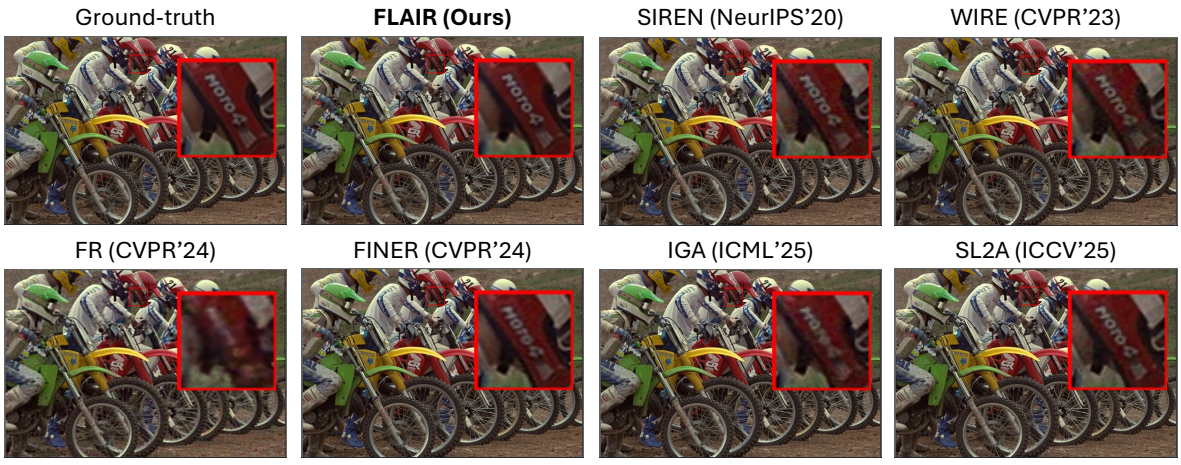


Figure 25. **Image fitting results on Kodak image 05.**

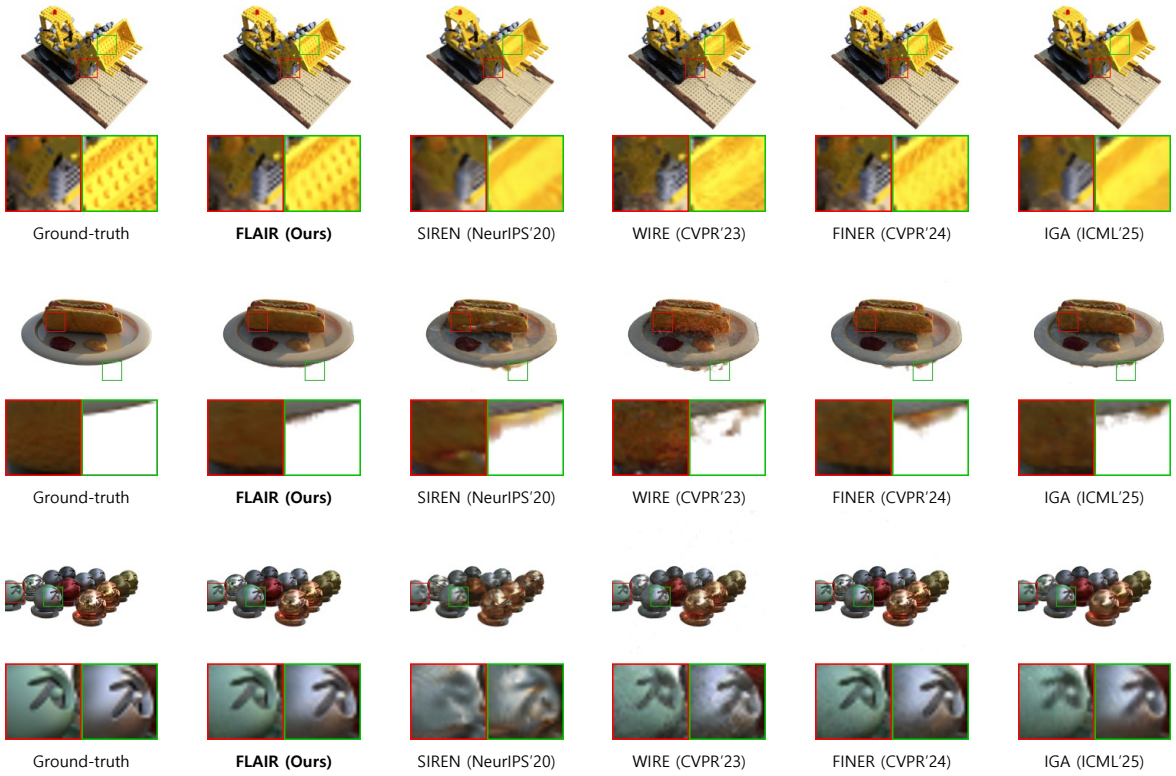


Figure 26. **Per-scene qualitative results on neural radiance field reconstruction.** Despite using only 25 input views instead of the default 100, FLAIR reconstructs unseen views without artifacts and preserves fine details.

Table 10. **Per-scene quantitative NeRF reconstruction results (8 scenes).** **Red** and blue indicate best and second-best per column.

Methods	#Params (K)	NeRF Scenes								Avg.
		Drums	Chair	Ficus	Hotdog	Lego	Materials	Mic	Ship	
PSNR ↑										
IGA (ICML'25)	205.1	24.25	32.95	26.98	32.06	28.95	27.38	33.74	21.80	28.52
FINER (CVPR'24)	198.9	24.91	33.38	27.28	32.71	29.77	27.64	34.42	21.63	28.97
WIRE (CVPR'23)	91.6	23.87	30.68	25.63	31.15	27.58	25.50	32.23	20.95	27.20
SIREN (NeurIPS'20)	198.9	21.90	25.81	24.21	31.02	27.20	24.21	29.62	20.83	25.60
FLAIR (Ours)	199.0	**25.03**	**34.03**	**27.33**	**33.12**	**30.00**	**28.29**	**34.75**	**21.93**	**29.31**
SSIM ↑										
IGA (ICML'25)	205.1	0.8858	0.9648	0.9417	0.9580	0.9349	0.9343	0.9771	0.7684	0.9217
FINER (CVPR'24)	198.9	0.8995	0.9698	0.9466	0.9601	0.9479	0.9380	0.9796	0.7671	0.9261
WIRE (CVPR'23)	91.6	0.8818	0.9482	0.9248	0.9410	0.9087	0.9047	0.9729	0.7249	0.9009
SIREN (NeurIPS'20)	198.9	0.8459	0.8955	0.9034	0.9475	0.9021	0.8967	0.9622	0.7320	0.8857
FLAIR (Ours)	199.0	**0.9098**	**0.9752**	**0.9506**	**0.9666**	**0.9509**	**0.9520**	**0.9811**	**0.7734**	**0.9325**
LPIPS ↓										
IGA (ICML'25)	205.1	0.101	0.024	0.055	0.029	0.057	0.042	0.021	0.188	0.063
FINER (CVPR'24)	198.9	0.055	0.017	0.037	0.032	**0.027**	0.033	**0.012**	0.139	0.044
WIRE (CVPR'23)	91.6	0.071	0.031	0.059	0.047	0.049	0.066	0.020	0.178	0.065
SIREN (NeurIPS'20)	198.9	0.147	0.113	0.142	0.038	0.084	0.071	0.041	0.200	0.105
FLAIR (Ours)	199.0	**0.053**	**0.015**	**0.036**	**0.022**	0.033	**0.026**	0.013	**0.126**	**0.041**

Table 11. **Per-scene quantitative image-fitting results on DIV2K (00–15).** **Red** and blue indicate best and second-best values per column.

Methods	#Params (K)	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	**Avg.**
PSNR ↑																		
SL2A (ICCV'25)	330.2	32.94	38.43	34.78	38.60	35.30	33.60	36.24	35.67	31.37	33.83	40.56	**45.20**	35.48	35.20	37.07	35.92	36.26
IGA (ICML'25)	205.1	28.72	35.12	32.59	35.09	32.47	29.61	32.32	33.65	28.10	31.58	38.26	43.92	35.07	32.15	34.15	33.39	33.51
FR (CVPR'24)	6299.9	23.13	26.64	25.09	28.38	24.64	23.84	25.67	25.99	22.69	23.84	31.17	32.75	25.79	26.24	27.67	25.96	26.22
FINER (CVPR'24)	198.9	32.18	37.28	34.20	37.06	32.29	33.48	34.95	34.38	32.03	33.08	37.32	42.03	36.30	33.94	36.31	33.35	35.01
WIRE (CVPR'23)	91.6	28.15	33.83	30.81	35.62	30.28	29.01	31.41	31.37	28.21	28.74	36.90	40.57	34.24	31.25	33.89	30.40	32.17
SIREN (NeurIPS'20)	198.9	31.15	36.74	33.50	37.44	33.28	31.84	34.18	34.31	30.24	31.56	36.83	41.39	36.17	32.59	35.43	33.09	34.36
FLAIR (Ours)	199.0	**34.75**	**38.88**	**37.19**	**40.45**	**38.12**	**35.09**	**37.64**	**37.34**	**35.22**	**36.29**	**41.35**	44.77	**38.36**	**37.15**	**39.92**	**38.51**	**38.19**
SSIM ↑																		
SL2A (ICCV'25)	330.2	0.9271	0.9649	0.9361	0.9641	0.9480	0.9437	0.9555	0.9392	0.9319	0.9263	0.9747	0.9824	0.9345	0.9358	0.9613	0.9440	0.9481
IGA (ICML'25)	205.1	0.9563	0.9736	0.9708	0.9737	0.9796	0.9665	0.9668	0.9679	0.9749	0.9732	**0.9833**	0.9839	**0.9852**	0.9746	0.9780	0.9786	0.9742
FR (CVPR'24)	6299.9	0.8816	0.8999	0.9088	0.9325	0.8960	0.8915	0.8843	0.8996	0.9140	0.8801	0.9627	0.9050	0.9558	0.9121	0.9016	0.8628	0.9100
FINER (CVPR'24)	198.9	0.9187	0.9542	0.9452	0.9564	0.9469	0.9442	0.9460	0.9385	0.9422	0.9331	0.9572	0.9747	0.9590	0.9330	0.9510	0.9302	0.9457
WIRE (CVPR'23)	91.6	0.8446	0.9207	0.8831	0.9257	0.8825	0.8565	0.8612	0.8670	0.8647	0.8500	0.9226	0.9650	0.9277	0.8620	0.9015	0.8630	0.8874
SIREN (NeurIPS'20)	198.9	0.9182	0.9577	0.9410	0.9584	0.9422	0.9221	0.9334	0.9384	0.9180	0.9249	0.9519	0.9714	0.9580	0.9151	0.9395	0.9310	0.9408
FLAIR (Ours)	199.0	**0.9663**	**0.9762**	**0.9737**	**0.9772**	**0.9809**	**0.9669**	**0.9730**	**0.9687**	**0.9768**	**0.9739**	0.9821	**0.9850**	0.9708	**0.9753**	**0.9804**	**0.9798**	**0.9754**
LPIPS ↓																		
SL2A (ICCV'25)	330.2	0.020	**0.013**	0.065	0.014	0.062	0.018	0.031	0.043	0.019	0.066	0.024	**0.008**	0.013	0.049	0.029	0.076	0.034
IGA (ICML'25)	205.1	0.036	0.019	**0.019**	0.013	**0.014**	0.038	0.042	**0.011**	0.041	0.018	0.017	0.015	0.005	0.024	**0.016**	0.026	0.022
FR (CVPR'24)	6299.9	0.176	0.103	0.145	0.046	0.168	0.156	0.170	0.141	0.184	0.122	0.038	0.064	0.033	0.111	0.103	0.226	0.124
FINER (CVPR'24)	198.9	0.019	0.027	0.050	0.047	0.048	0.016	0.038	0.036	0.012	0.047	0.070	0.032	0.005	0.046	0.045	0.085	0.039
WIRE (CVPR'23)	91.6	0.128	0.077	0.115	0.079	0.093	0.124	0.160	0.127	0.109	0.123	0.068	0.053	0.023	0.138	0.120	0.185	0.107
SIREN (NeurIPS'20)	198.9	0.034	0.031	0.047	0.047	0.039	0.032	0.077	0.051	0.042	0.044	0.083	0.041	0.013	0.082	0.067	0.070	0.050
FLAIR (Ours)	199.0	**0.013**	0.016	**0.019**	**0.012**	**0.014**	**0.014**	**0.028**	0.026	**0.011**	**0.017**	**0.016**	0.013	**0.004**	**0.022**	0.019	**0.016**	**0.016**

Table 12. **Per-scene quantitative image-fitting results on Kodak (24 scenes).** **Red** and blue indicate best and second-best values per column.

Methods	#Params (K)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	**Avg.**
PSNR ↑																										
SL2A (ICCV'25)	330.2	33.01	38.40	39.79	37.33	32.79	34.90	**39.73**	30.49	38.60	38.04	36.33	38.17	29.83	34.36	36.80	**38.85**	**39.01**	**34.36**	35.91	36.42	35.67	35.01	**40.73**	32.74	36.14
IGA (ICML'25)	205.1	31.20	36.71	38.97	36.25	30.30	31.95	39.31	28.58	**38.80**	37.88	34.03	37.17	27.29	31.95	35.28	36.31	36.83	31.53	34.32	35.75	33.54	33.92	39.23	30.73	34.49
FR (CVPR'24)	6299.9	24.68	29.57	30.69	29.25	24.01	26.23	29.40	22.88	29.59	28.85	27.24	30.37	22.42	26.09	30.03	28.73	29.57	25.78	27.04	29.61	27.05	27.53	31.81	24.51	27.62
FINER (CVPR'24)	198.9	33.67	37.49	38.44	36.19	34.05	34.03	37.57	30.80	36.92	36.46	35.49	37.81	30.81	34.47	37.37	36.52	36.71	33.63	34.99	36.52	34.69	34.86	38.79	32.65	35.46
WIRE (CVPR'23)	91.6	26.30	31.22	33.04	30.74	26.69	28.00	30.43	23.28	31.64	30.24	29.20	30.25	25.80	27.03	30.63	30.49	29.23	25.45	27.33	28.83	28.45	29.37	30.83	26.60	28.80
SIREN (NeurIPS'20)	198.9	29.23	30.84	32.40	30.41	25.18	27.06	31.31	23.00	31.71	30.95	28.05	31.53	22.71	26.98	30.49	29.28	30.41	26.03	27.20	29.73	27.22	28.59	32.47	24.54	28.64
FLAIR (Ours)	199.0	**36.11**	**40.34**	**40.93**	**39.89**	**35.01**	**36.20**	39.32	**32.89**	38.63	**38.10**	**36.75**	**39.10**	**33.67**	**35.85**	**38.06**	38.76	37.29	33.63	**35.95**	**37.22**	**36.32**	**36.45**	39.90	**34.71**	**37.12**
SSIM ↑																										
SL2A (ICCV'25)	330.2	0.9255	0.9320	0.9558	0.9307	0.9260	0.9298	0.9658	0.8980	0.9430	0.9399	0.9330	0.9382	0.8828	0.9264	0.9222	0.9579	0.9563	0.9177	0.9278	0.9207	0.9305	0.9145	0.9576	0.8985	0.9304
IGA (ICML'25)	205.1	0.9631	0.9474	0.9754	0.9690	0.9686	0.9321	0.9743	0.9501	0.9646	**0.9781**	0.9461	0.9632	0.9336	0.9599	**0.9828**	0.9636	**0.9722**	0.9311	0.9424	0.9822	0.9630	0.9459	0.9630	0.9566	0.9596
FR (CVPR'24)	6299.9	0.8404	0.8029	0.9478	0.8822	0.8778	0.8008	0.9514	0.9031	0.8945	0.8981	0.8373	0.8911	0.7742	0.8843	0.9509	0.8299	0.9285	0.8369	0.8450	0.9120	0.8800	0.8529	**0.9678**	0.8855	0.8781
FINER (CVPR'24)	198.9	0.9188	0.9155	0.9345	0.9102	0.9413	0.9135	0.9590	0.9157	0.9195	0.9257	0.9150	0.9294	0.9173	0.9213	0.9282	0.9271	0.9384	0.9207	0.9129	0.9254	0.9239	0.9161	0.9434	0.9185	0.9246
WIRE (CVPR'23)	91.6	0.6672	0.7730	0.8742	0.7921	0.8121	0.7823	0.8654	0.7013	0.8432	0.8310	0.8017	0.8194	0.7542	0.7308	0.8222	0.7892	0.8212	0.6992	0.7668	0.8188	0.8275	0.7680	0.8669	0.8001	0.7928
SIREN (NeurIPS'20)	198.9	0.8454	0.7627	0.8591	0.7829	0.7362	0.7551	0.8975	0.6887	0.8669	0.8420	0.7424	0.8206	0.5735	0.7114	0.8084	0.7489	0.8479	0.7032	0.7405	0.8336	0.8035	0.7300	0.8967	0.7027	0.7792
FLAIR (Ours)	199.0	**0.9795**	**0.9695**	**0.9785**	**0.9693**	**0.9703**	**0.9667**	**0.9792**	**0.9591**	**0.9770**	0.9605	**0.9556**	**0.9786**	**0.9559**	**0.9601**	0.9569	**0.9689**	0.9593	**0.9314**	**0.9428**	**0.9909**	**0.9645**	**0.9471**	0.9631	**0.9597**	**0.9644**
LPIPS ↓																										
SL2A (ICCV'25)	330.2	0.056	0.047	0.037	0.071	0.039	0.064	**0.022**	0.064	0.051	**0.046**	**0.062**	0.068	0.108	**0.054**	0.063	0.050	**0.030**	**0.052**	**0.083**	0.080	0.080	0.072	0.045	0.092	0.060
IGA (ICML'25)	205.1	0.045	0.043	0.027	0.055	0.033	0.059	0.030	0.051	**0.018**	0.069	0.054	0.050	0.099	0.086	0.066	0.050	0.082	0.086	0.114	**0.035**	0.066	**0.032**	0.050	0.064	0.056
FR (CVPR'24)	6299.9	0.206	0.201	0.087	0.111	0.202	0.186	0.092	0.170	0.062	0.106	0.208	0.086	0.385	0.165	0.085	0.183	0.096	0.175	0.127	0.104	0.157	0.200	0.051	0.213	0.151
FINER (CVPR'24)	198.9	0.051	0.039	0.032	0.073	0.029	0.060	0.036	0.045	0.076	0.053	0.081	0.061	0.086	0.064	0.088	0.060	0.074	0.072	0.104	0.081	0.083	0.072	0.084	0.078	0.065
WIRE (CVPR'23)	91.6	0.445	0.383	0.182	0.332	0.224	0.289	0.202	0.383	0.298	0.331	0.350	0.379	0.373	0.442	0.288	0.362	0.302	0.440	0.467	0.272	0.413	0.321	0.223	0.280	0.333
SIREN (NeurIPS'20)	198.9	0.167	0.423	0.230	0.360	0.309	0.353	0.188	0.441	0.290	0.324	0.432	0.386	0.618	0.475	0.332	0.456	0.280	0.436	0.500	0.252	0.446	0.446	0.222	0.491	0.369
FLAIR (Ours)	199.0	**0.018**	**0.021**	**0.024**	**0.052**	**0.019**	**0.052**	**0.022**	**0.034**	0.051	0.062	0.065	**0.046**	**0.071**	0.059	**0.062**	**0.048**	0.072	0.098	0.101	0.061	**0.063**	0.070	**0.043**	**0.061**	**0.054**