BEYOND POINT SAMPLES: ACQUISITION MODELING AS A SEPARATE AXIS FOR IMPLICIT NEURAL REPRESENTATIONS

Anonymous authors
Paper under double-blind review

ABSTRACT

Implicit neural representations (INRs) provide continuous representations of signals, yet recovering a continuous field from discrete measurements depends on both the representation and the process by which measurements constrain it. This relationship is specified by the acquisition model, which maps the continuous field to discrete observations. However, INR reconstruction has largely treated this measurement process as fixed. Most INR work has focused on representation design, while the measurement constraint itself is usually kept fixed as pointwise supervision. Here we isolate the effect of acquisition choice by holding the representation fixed and varying only the acquisition model. We find that acquisition choice materially changes the recovered continuous field, with increasingly pronounced effects as reconstruction relies increasingly on interpolation between sparse observations. Spatial acquisition improves continuous-field recovery overall, but its effect is spatially nonuniform, with larger reductions around boundaries, thin structures, and regions of rapid spatial variation, while point acquisition can remain locally preferable. Acquisition choice also changes the spectral structure of reconstruction error, particularly at intermediate and higher spatial frequencies. These results motivate treating the acquisition model as an explicit part of INR reconstruction rather than a fixed supervision choice.

1 INTRODUCTION

Implicit neural representations (INRs) model signals as continuous functions $f_\theta(\mathbf{x})$ over spatial coordinates $\mathbf{x}$, allowing images and other signals to be represented beyond a fixed discrete grid (Sitzmann et al., 2020; Mildenhall et al., 2020). This continuous formulation is particularly attractive for reconstruction, as the recovered field can be evaluated at arbitrary spatial locations, including those not directly observed. Yet the underlying field must still be inferred from discrete measurements. Standard coordinate-based fitting (Sitzmann et al., 2020; Liu et al., 2024; Wang et al., 2025) typically imposes pointwise constraints,

$$f_\theta(\mathbf{c}_i) \approx y_i. \tag{1}$$

Agreement at these coordinates does not uniquely determine the field between them. Continuous parameterization therefore provides continuous queryability, but does not by itself specify how the continuous field is recovered from the available measurements.

Much of INR research has addressed reconstruction challenges by improving representational capacity. Difficulties in recovering fine-scale and high-frequency structure, often discussed in connection with spectral bias (Rahaman et al., 2019), have motivated periodic activations (Sitzmann et al., 2020; Liu et al., 2024), Fourier feature embeddings (Tancik et al., 2020), Fourier- and Gabor-based representations (Fathony et al., 2021; Saragadam et al., 2023), multiscale formulations, and hierarchical or grid-based parameterizations (Müller et al., 2022; Wang et al., 2025). More recently, WIEN-INR introduced a hierarchical multiresolution formulation in the wavelet domain, which models signals across multiple frequency scales to better preserve high-frequency structure (Ni et al., 2026). Together, these developments have substantially expanded the range of spatial and frequency structure that INRs can represent.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

However, comparatively less attention has been paid to how discrete measurements constrain these representations. Pointwise fitting remains the dominant observation model in coordinate-based reconstruction. Improving the representation expands the class of functions that can be expressed without changing how the available observations constrain the recovered field. This leaves a complementary question: when the representation is held fixed, how much can the recovered continuous field change solely because the measurements constrain it differently?

A spatial formulation provides a controlled way to examine this question. Rather than associating each observation with a single field value, it relates the measurement to the response of the field over a finite neighborhood. Point and spatial formulations therefore impose different measurement constraints on the same underlying continuous field. Importantly, the purpose of this comparison is not to introduce spatial acquisition as a new reconstruction method, but to use the contrast between point and spatial constraints to isolate the effect of acquisition choice.

Spatial measurement formulations have previously been explored in several neural-field settings. Arbitrary-scale image models account for pixel extent (Chen et al., 2021; Becker et al., 2025), neural rendering methods integrate over finite pixel footprints (Barron et al., 2021; 2023), and application-specific reconstruction methods incorporate explicit measurement operators (Shen et al., 2024; Jyhne et al., 2026). These works, however, were not designed to study the effect of the acquisition model itself. Instead, the measurement formulation is embedded within broader changes to the architecture, rendering procedure, scale handling, or application-specific pipeline. Classical sampling theory, however, has long treated the continuous representation and the measurement process as distinct components (Unser et al., 1993; Unser & Aldroubi, 1994; Unser, 2000).

Despite its central role in classical sampling theory, acquisition modeling has received relatively little attention in INR reconstruction. We therefore isolate its effect by keeping the representation fixed and varying only the measurement constraint.

To examine how acquisition choice affects continuous-field recovery, we fix the representation and compare conventional point acquisition with a representative spatial acquisition model across SIREN (Sitzmann et al., 2020), FINER (Liu et al., 2024), SASNet (Feng et al., 2026), and MetricGrids (Wang et al., 2025), using two image datasets and three sampling densities. We characterize how these measurement constraints affect the recovered field in terms of aggregate reconstruction quality, spatial error distribution, and spectral error structure. From this controlled study, we make three main contributions:

- We identify acquisition modeling as a complementary axis of INR reconstruction and isolate its effect by holding the representation fixed. Across the evaluated INR families, changing only the acquisition model produces systematic differences in the recovered continuous field, demonstrating that recovery is determined not only by representational capacity but also by the measurement constraints imposed during fitting.
- We show that spatial acquisition improves overall reconstruction quality across all evaluated settings, with the largest gains concentrated in structurally complex regions. At the same time, point acquisition remains locally preferable in some comparatively smooth regions, which reveals that the relative effectiveness of an acquisition model within INR reconstruction is spatially and signal-dependent.
- We show that changing only the acquisition model in INR reconstruction alters the spectral structure of reconstruction error, with differences extending across a broader range of spatial frequencies as measurements become sparser.

2 Related Work

Representation design in implicit neural representations. A large body of INR research has focused on improving the representation of continuous signals. SIREN introduced sinusoidal activations to better represent fine-scale structure and signal derivatives (Sitzmann et al., 2020), while Fourier Features used sinusoidal coordinate embeddings to improve the learning of high-frequency components (Tancik et al., 2020). Subsequent approaches expanded representational capacity through Fourier- and Gabor-based filters (Fathony et al., 2021), explicitly band-limited multiscale representations (Lindell et al., 2022), Gabor-wavelet activations (Saragadam et al., 2023),

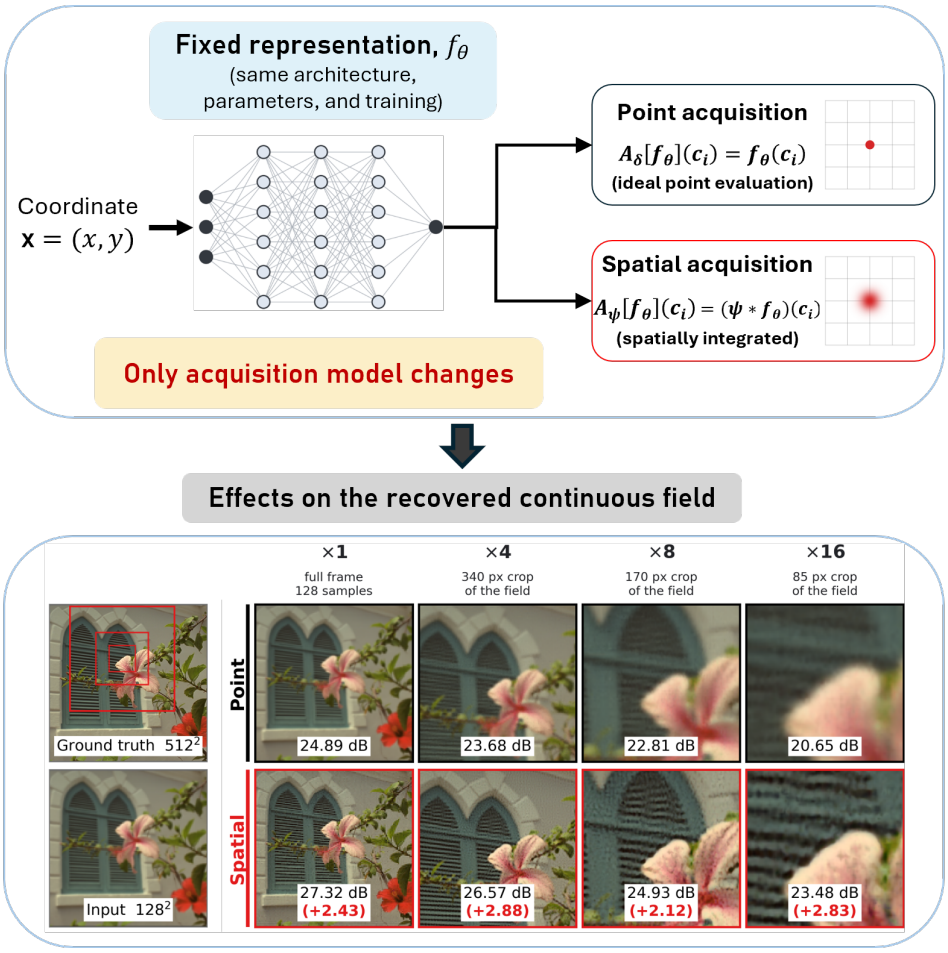

Figure 1: Effect of acquisition choice with a fixed INR representation. Only the acquisition model is varied between point and spatial constraints, resulting in different continuous-field reconstructions across sampling scales. Here, $\times k$ denotes a k-fold zoom into a field reconstructed from 128^2 measurements and is distinct from the sampling scales reported in Table 1.

and variable-periodic activations (Liu et al., 2024). Other methods employ multiresolution grids or hybrid continuous–discrete parameterizations (Müller et al., 2022; Wang et al., 2025), while recent work explores spatially adaptive frequency allocation and multiscale representations (Feng et al., 2026; Thennakoon et al., 2026; Ni et al., 2026). These developments span activations, encodings, frequency structure, and parameterization, but primarily modify the representation itself.

Sampling and reconstruction of continuous signals. Classical sampling theory distinguishes the space in which a continuous signal is represented from the measurement process that produces discrete observations. Beyond idealized point sampling, generalized formulations allow measurements to arise from filtering, local averaging, and other linear functionals (Unser, 2000; Unser & Aldroubi, 1994). Reconstruction therefore depends on both the measurement mechanism and the space in which the signal is approximated (Unser, 1999). The same distinction appears in imaging systems, where finite detector support and point-spread functions associate a recorded measurement with a spatial response rather than the signal value at an infinitesimal point (Baker & Kanade, 2002; Efrat et al., 2013).

Spatial constraints in neural fields. Finite-support and scale-aware formulations also appear in modern neural fields. LIIF incorporates output-cell extent in arbitrary-scale image decoding (Chen et al., 2021; Lee & Jin, 2022), while mip-NeRF and related methods account for finite spatial footprints in neural rendering (Barron et al., 2021; 2023). Continuous neural fields used in inverse

problems are likewise evaluated through application-specific measurement operators (Shen et al., 2024; Chen et al., 2023; Ye et al., 2023). In image reconstruction, Thera incorporates a finite spatial response when relating a continuous field to image pixels (Becker et al., 2025), SuperF combines supersampling with spatial filtering (Jyhne et al., 2026), and spectral prefiltering methods incorporate filtering directly into neural fields (Yaldiz et al., 2025).

These works show that non-point and finite-support formulations can be useful, but they are typically introduced together with changes in architecture, rendering, scale handling, or application-specific pipelines. Their effect on the recovered continuous field is therefore difficult to isolate from the surrounding system. We instead retain existing representations and vary only the acquisition model, allowing its effect on continuous reconstruction to be examined directly.

3 Acquisition Models in Continuous Reconstruction

3.1 Problem Formulation

Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ denote the continuous field to be recovered, with discrete measurements $\{(\mathbf{c}_i, y_i)\}_{i=1}^N$. An acquisition model $\mathcal{A}_\psi$ relates the field to each measurement,

$$y_i = \mathcal{A}_\psi[f](\mathbf{c}_i). \tag{2}$$

Here, A_ψ specifies how the continuous field is mapped to the discrete measurement observed at $\mathbf{c}_i$. For a representation f_θ with parameters θ, let $\tilde{\psi}$ denote the acquisition kernel assumed during reconstruction, then

$$\min_\theta \sum_{i=1}^{N} \left\| \mathcal{A}_{\tilde{\psi}}[f_\theta](\mathbf{c}_i) - y_i \right\|_2^2. \tag{3}$$

The reconstruction therefore seeks parameters θ whose predicted measurements under $A_{\tilde{\psi}}$ agree with the observed measurements y_i. Point acquisition is the special case $\mathcal{A}_\delta[g](\mathbf{c}) = g(\mathbf{c})$, which recovers the conventional coordinate-fitting objective. Varying $\tilde{\psi}$ therefore changes how the same measurements constrain the same continuous representation. Under point acquisition, each measurement constrains only the value of the reconstructed field at its corresponding coordinate. Figure 1 illustrates this controlled setting, in which the representation and reconstruction setup are held fixed while only the acquisition model is varied, leading to different recovered continuous fields.

3.2 Point and Spatial Acquisition

We compare point acquisition with a finite-support spatial condition. For the latter, we use a spatially invariant Gaussian–pixel kernel

$$\psi = k_\sigma * u_h, \tag{4}$$

where k_σ is a unit-mass Gaussian kernel and u_h is a unit-mass finite-pixel footprint (Baker & Kanade, 2002). Its response is

$$\mathcal{A}_\psi[f](\mathbf{c}_i) = (\psi * f)(\mathbf{c}_i). \tag{5}$$

For the point reconstruction condition, we set $\tilde{\psi} = \delta$, whereas for the spatial reconstruction condition, $\tilde{\psi} = \psi$. During reconstruction, the response under the assumed kernel $\tilde{\psi}$ is evaluated on the continuous INR using fixed quadrature,

$$\mathcal{A}_{\tilde{\psi}}[f_\theta](\mathbf{c}_i) \approx \sum_{k=1}^{K^2} w_k \, f_\theta(\mathbf{c}_i + \Delta_k), \qquad \sum_{k=1}^{K^2} w_k = 1. \tag{6}$$

where the nodes Δ_k are determined by $\tilde{\psi}$, while the quadrature weights w_k are fixed. The resulting weighted sum provides the numerical approximation of the spatial measurement used in the reconstruction loss. The spatial condition uses $K = 16$ samples per axis and introduces no additional trainable parameters. Sensitivity to the quadrature resolution is reported in Appendix A. We use the Gaussian–pixel kernel as one representative spatial acquisition model rather than as an optimal or universal choice. Sensitivity to the assumed Gaussian width is examined in Appendix B.

Table 1: **Effect of acquisition choice on continuous-field recovery.** Field PSNR (dB) on Kodak-24 and DIV2K-24 across sampling scales, averaged over three random seeds. Point and Spatial denote pointwise and spatial acquisition, respectively. Δ denotes Spatial $-$ Point.

		$\times 1$ (native-grid)			$\times 2$			$\times 4$		
Dataset	**Method**	**Point**	**Spatial**	Δ	**Point**	**Spatial**	Δ	**Point**	**Spatial**	Δ
Kodak-24	FINER	30.20	31.29	+**1.09**	27.05	28.91	+**1.86**	24.30	25.76	+**1.45**
	MetricGrids	30.72	32.37	+**1.65**	27.07	29.21	+**2.14**	24.17	25.56	+**1.39**
	SIREN	26.54	26.72	+**0.17**	26.01	26.69	+**0.68**	24.26	25.50	+**1.24**
	SASNet	28.78	29.19	+**0.41**	26.71	27.68	+**0.97**	24.27	25.04	+**0.77**
DIV2K-24	FINER	28.43	29.40	+**0.97**	25.48	27.44	+**1.96**	22.41	24.13	+**1.72**
	MetricGrids	28.80	30.10	+**1.29**	25.46	27.63	+**2.17**	22.28	23.92	+**1.64**
	SIREN	24.96	25.12	+**0.15**	24.41	25.12	+**0.71**	22.40	23.82	+**1.42**
	SASNet	27.04	27.48	+**0.44**	25.12	26.19	+**1.07**	22.36	23.56	+**1.20**

4 Experimental Setup

Data and sampling. We use Kodak-24 (Franzen, 1999) and the first 24 DIV2K validation images (Agustsson & Timofte, 2017). Each image defines a 512×512 reference field, extended to continuous coordinates via bilinear interpolation (Unser, 1999). We evaluate $\times 1$, $\times 2$, and $\times 4$ sampling, corresponding to measurement lattices of 512×512, 256×256, and 128×128, respectively. Measurements are synthesized directly from the reference field using the Gaussian–pixel acquisition model in Eq. (4) with $\sigma/h = 0.5$, and reconstruction is evaluated against the unfiltered 512×512 reference. A complementary acquisition control using corresponding measurement and reconstruction models is provided in Appendix E.

Implementation details. Within each Point–Spatial pair, the representation and all other hyperparameters are held fixed. Only the supervision operator changes. We use three random seeds per image, average results across seeds for each image, and treat each image as the statistical unit. Full details of the acquisition model, quadrature, and optimization are provided in Appendix C. Additional controls for model capacity and optimization budget are reported in Appendix F.

Representations and controlled comparison. We evaluate FINER (Liu et al., 2024), MetricGrids (Wang et al., 2025), SIREN (Sitzmann et al., 2020), and SASNet (Feng et al., 2026), spanning both coordinate-based and grid-based representations. Within each point–spatial comparison, the representation, initialization, optimizer, learning-rate schedule, measurements, and number of optimization steps are held fixed; only the acquisition model used during reconstruction changes. The same acquisition configuration is used consistently across all representations within each sampling setting. We use three seeds per image, average over seeds within each image, and treat the image as the statistical unit. Reconstruction quality is reported as field PSNR against the reference field.

5 Results

5.1 Acquisition Choice Changes Continuous-Field Recovery

Table 1 shows that continuous-field recovery changes systematically with the acquisition model even when the underlying INR is held fixed. Spatial acquisition improves field PSNR in all 24 representation–dataset–scale comparisons, with gains of up to 2.17 dB. This consistency across architectures and sampling densities shows that acquisition choice can materially affect the continuous field recovered from the same measurements. A complementary acquisition control using corresponding measurement and reconstruction models is reported in Appendix E.

Native-grid recovery. At $\times 1$, the measurement and evaluation coordinates coincide, providing a setting in which the effect is not attributable to increasingly sparse sampling. Spatial acquisition improves all eight representation–dataset comparisons by 0.15–1.65 dB. All four representations therefore benefit from Spatial acquisition at the native grid on both Kodak-24 and DIV2K-24. The



Figure 2: Qualitative comparison of point and spatial acquisition. Insets report field PSNR, with Spatial-minus-Point differences in parentheses.

gains are largest for FINER and MetricGrids and smaller for SIREN and SASNet, showing that the acquisition model affects recovery even at the native sampling grid.

Continuous recovery from coarser measurements. The differences become particularly informative at $\times 2$ and $\times 4$, where the full 512×512 field is recovered from 256×256 and 128×128 measurement lattices. At $\times 2$, spatial acquisition improves all eight comparisons by 0.68–2.17 dB, with the largest gains for FINER and MetricGrids. At $\times 4$, the improvement remains positive in every case, ranging from 0.77 to 1.72 dB. Across both coarser sampling settings, spatial acquisition therefore remains better than Point acquisition for every representation–dataset pair. The effect varies with both representation and sampling density, but remains substantial as recovery extends increasingly between measured locations.

Figure 2 provides qualitative evidence for how acquisition choice changes the recovered field. Across all four representations, spatial acquisition reduces reconstruction error not only in magnitude but also in its spatial distribution. The difference is already visible at $\times 1$ and becomes substantially more pronounced at $\times 2$ and $\times 4$, where point acquisition produces strong errors around edges, fine structures, and textured regions, while spatial acquisition suppresses much of this error. The same pattern appears across architectures with markedly different parameterizations, suggesting that it is not specific to one INR design. These error maps further suggest that the effect of acquisition choice is spatially nonuniform, motivating the region-level analysis that follows. Additional qualitative examples are provided in Appendix H.

5.2 Spatial Error Analysis

The qualitative results suggest that acquisition choice changes not only the magnitude of reconstruction error, but also where that error occurs. To examine this effect directly, we compute the signed

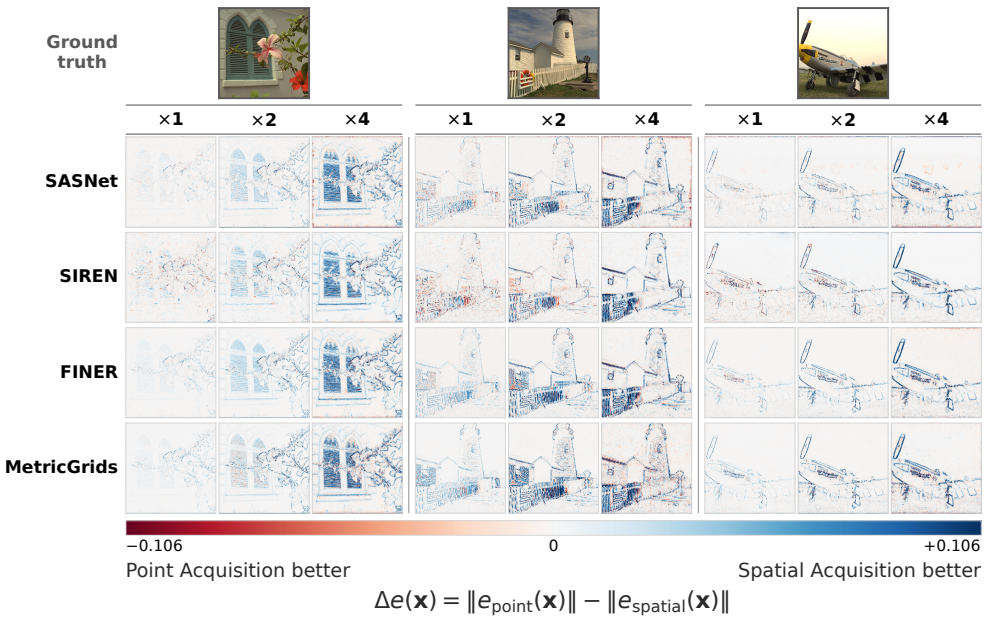

Figure 3: Spatial distribution of acquisition-dependent reconstruction error. Signed error differences across three reference fields, four INR families, and $\times 1$, $\times 2$, and $\times 4$ sampling. Positive values favor spatial acquisition and negative values favor point acquisition.

difference between point and spatial reconstruction errors at each location $\mathbf{x}$,

$$\Delta e(\mathbf{x}) = \|e_{\text{point}}(\mathbf{x})\| - \|e_{\text{spatial}}(\mathbf{x})\|, \tag{7}$$

where positive values indicate lower error under spatial acquisition and negative values indicate lower error under point acquisition.

Figure 3 reveals a clear spatial organization of these differences, with blue indicating lower error under spatial acquisition and red indicating lower error under point acquisition. Rather than producing a uniform reduction in error, spatial acquisition primarily changes recovery along object boundaries, thin structures, repeated patterns, and other regions containing rapid spatial variation. This behavior is visible across all four INR families and becomes more pronounced as the measurement lattice becomes coarser. The similarity of the spatial patterns across markedly different representations further suggests that these differences arise from the acquisition constraint rather than from a particular INR architecture. Additional spatial error maps are provided in Appendix H.

The maps also reveal an important qualification to the aggregate improvement in Table 1. Although spatial acquisition achieves higher field PSNR globally, negative regions remain in which point acquisition produces lower local reconstruction error. Therefore, its relative advantage depends on the local signal structure.

5.3 Spectral Error Analysis

We next examine how acquisition choice changes the frequency distribution of reconstruction error. We define spectral error from the radially averaged log-power spectra of the reconstruction and reference; full computation details are provided in Appendix G. Figure 4 compares the spectral error under point and spatial acquisition, with the shaded region showing the error reduction obtained by spatial acquisition at each spatial frequency. At $\times 1$, this reduction is concentrated primarily at intermediate and higher spatial frequencies. As the measurement lattice becomes coarser, the shaded region extends toward lower frequencies, indicating that the effect of acquisition choice spreads across a broader portion of the spectrum. This pattern is consistent across representations, although the magnitude and spectral extent of the reduction vary.

The spectral and spatial analysis therefore provide complementary views of the same nonuniform behavior. Under denser sampling, the effect of spatial acquisition is concentrated in finer spatial

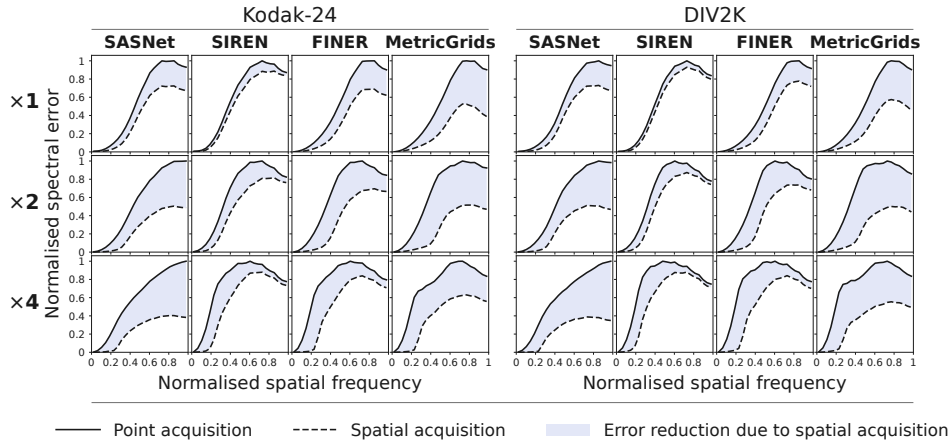

Figure 4: Acquisition choice alters the spectral structure of reconstruction error. Normalized spectral error for four INR families across Kodak-24, DIV2K, and $\times 1$, $\times 2$, and $\times 4$ sampling. Solid and dashed curves denote point and spatial acquisition, respectively. Shading indicates the error reduction under spatial acquisition. Curves are normalized within each panel for visualization.

structure and higher-frequency content. Under stronger undersampling, its influence extends across a broader range of structures and spatial frequencies.

6 Discussion

6.1 Acquisition Choice and Continuous-Field Recovery

By holding the representation fixed, we show that changing only the acquisition model produces systematically different recovered continuous fields. This difference becomes particularly important at $\times 2$ and $\times 4$, where increasingly large portions of the field lie between measured locations. Table 1 quantifies this effect across representations, datasets, and sampling scales, while Figure 2 makes the corresponding differences in the recovered fields directly visible, with acquisition-dependent reconstruction differences becoming more pronounced as the measurement lattice becomes coarser. Even with the same representation and the same discrete measurements, different measurement constraints lead to different values being recovered between observations. The ability to query an INR continuously therefore does not by itself determine the continuous field that is recovered; acquisition choice is also consequential to what is reconstructed away from the observations.

This distinction matters in inverse problems where observations sparsely constrain an underlying continuous quantity, as in tomographic or remote-sensing measurements. In such settings, values between observed locations are part of the reconstruction target itself, not merely interpolated display values.

6.2 Complementary Behavior of Acquisition Models

The spatially varying behavior of point and spatial acquisition, directly visible in Figure 3, suggests a broader implication for acquisition modeling in continuous reconstruction. If different structures within the same signal respond differently to the same measurement constraint, then imposing a single acquisition model uniformly across the field may be unnecessarily restrictive. The acquisition model can instead be viewed as potentially signal-dependent, with different forms of measurement support being advantageous in different regions. This observation suggests that a single global acquisition model may be suboptimal. A natural extension is to adapt the measurement support to local signal structure.

6.3 Acquisition Effects on High-Frequency Reconstruction

Spectral bias is commonly associated with the tendency of neural networks to favor low-frequency structure, while higher-frequency components are more difficult to learn (Rahaman et al., 2019). Much of INR research has therefore approached high-frequency recovery through representation. Our experiments show that, even when the representation is held fixed, changing the acquisition model substantially alters reconstruction error at intermediate and higher spatial frequencies, with the affected spectral range broadening as measurements become sparser. Figure 4 shows this effect directly: the error reduction under spatial acquisition is concentrated at intermediate and higher frequencies under denser sampling and extends over a broader frequency range as the measurement lattice becomes coarser. Thus, high-frequency recovery depends not only on what the representation can express, but also on how the measurements constrain it.

6.4 Acquisition Modeling as a Separate Axis

Recent work has begun to question whether increasingly sophisticated INRs necessarily lead to better continuous reconstruction. Kim & Fridovich-Keil (2025), for example, showed that simple interpolated grids can match or outperform more sophisticated INR approaches under comparable parameter budgets. Our results point to the acquisition model as another important factor: changing it produces substantial differences in reconstruction quality, spatial error structure, and spectral error without requiring a different INR. Acquisition modeling should therefore be treated as a complementary design axis alongside the INR itself. This broadens the study of continuous reconstruction from asking only how the field should be represented to also asking how discrete measurements should constrain the field being recovered.

7 Limitations and Future Work

In this study, we use a Gaussian–pixel acquisition model as a representative spatial formulation in order to keep the analysis focused. The goal is to isolate the effect of the acquisition model by comparing conventional point acquisition with this representative spatial formulation while keeping the representation fixed. Future work should examine other point-spread functions, pixel footprints, anisotropic kernels, and spatially varying acquisition models. Such extensions could help impose more appropriate constraints on INR representations during continuous-field reconstruction.

Spatial acquisition also introduces additional computational cost because each measurement depends on the response of the field over a neighborhood rather than at a single location. Training and inference costs are quantified in Appendix D. This study is intended as a controlled analysis for understanding acquisition modeling in INRs rather than as a proposal for an efficient reconstruction method. The added computational cost is therefore a limitation and a clear direction for future work. Ultimately, efficiency should be considered together with the reconstruction objective. When the goal is faithful recovery of a continuous field, computational savings are useful only insofar as that recovery quality is preserved.

8 Conclusion

In this study, we isolated the effect of acquisition modeling in INR reconstruction by holding the representation fixed and varying only the measurement constraint. Across four INR families, two datasets, and three sampling densities, this choice changed reconstruction quality as well as the spatial and spectral structure of the error. The differences were largest in structurally complex regions and under coarser sampling, although point acquisition remained locally preferable in some regions. These results suggest that the acquisition model should be treated explicitly when designing and evaluating continuous-field reconstruction systems.

AI use statement

Generative AI tools were used during the preparation of this work to assist with editing of the manuscript for clarity and presentation. The authors take full responsibility for the final content of the paper.

REFERENCES

Eirikur Agustsson and Radu Timofte. NTIRE 2017 challenge on single image super-resolution: Dataset and study. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1122–1131, 2017.

Simon Baker and Takeo Kanade. Limits on super-resolution and how to break them. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(9):1167–1183, 2002. doi: 10.1109/TPAMI.2002.1033210.

Jonathan T. Barron, Ben Mildenhall, Matthew Tancik, Peter Hedman, Ricardo Martin-Brualla, and Pratul P. Srinivasan. Mip-NeRF: A multiscale representation for anti-aliasing neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 5835–5844, 2021.

Jonathan T. Barron, Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, and Peter Hedman. Zip-NeRF: Anti-aliased grid-based neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 19640–19648, 2023.

Alexander Becker, Rodrigo Caye Daudt, Dominik Narnhofer, Torben Peters, Nando Metzger, Jan Dirk Wegner, and Konrad Schindler. Thera: Aliasing-free arbitrary-scale super-resolution with neural heat fields. *Transactions on Machine Learning Research*, 2025. URL https://openreview.net/forum?id=GU8YOfmqyg.

Yinbo Chen, Sifei Liu, and Xiaolong Wang. Learning continuous image representation with local implicit image function. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8624–8634, 2021. doi: 10.1109/CVPR46437.2021.00852.

Zixuan Chen, Lingxiao Yang, Jianhuang Lai, and Xiaohua Xie. APRF: Anti-aliasing projection representation field for inverse problem in imaging. *arXiv preprint arXiv:2307.05270*, 2023.

Netalee Efrat, Daniel Glasner, Alexander Apartsin, Boaz Nadler, and Anat Levin. Accurate blur models vs. image priors in single image super-resolution. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 2832–2839, 2013. doi: 10.1109/ICCV.2013.352.

Rizal Fathony, Anit Kumar Sahu, Devin Willmott, and J. Zico Kolter. Multiplicative filter networks. In *International Conference on Learning Representations (ICLR)*, 2021. URL https://openreview.net/forum?id=OmtmcPkkhT.

Haoan Feng, Diana Aldana, Tiago Novello, and Leila De Floriani. SASNet: Spatially-adaptive sinusoidal networks for INRs. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 41964–41973, June 2026.

Rich Franzen. Kodak lossless true color image suite. http://r0k.us/graphics/kodak/, 1999. PhotoCD PCD0992, 24 images.

Sander Riisøen Jyhne, Christian Igel, Morten Goodwin, Per-Arne Andersen, Serge Belongie, and Nico Lang. SuperF: Neural implicit fields for multi-image super-resolution. In *International Conference on Learning Representations (ICLR)*, 2026. URL https://openreview.net/forum?id=FiiItlSqqL.

Namhoon Kim and Sara Fridovich-Keil. Grids often outperform implicit neural representation at compressing dense signals. In D. Belgrave, C. Zhang, H. Lin, R. Pascanu, P. Koniusz, M. Ghassemi, and N. Chen (eds.), *Advances in Neural Information Processing Systems*, volume 38, Main Conference, pp. 51926–51980. Curran Associates, Inc., 2025. doi: 10.52202/085713-1731. URL https://proceedings.neurips.cc/paper_files/paper/2025/file/4a9195eba1bb530fb4effb30cea616ca-Paper-Conference.pdf.

Jaewon Lee and Kyong Hwan Jin. Local texture estimator for implicit representation function. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1919–1928, 2022. doi: 10.1109/CVPR52688.2022.00197.

David B. Lindell, Dave Van Veen, Jeong Joon Park, and Gordon Wetzstein. BACON: Band-limited coordinate networks for multiscale scene representation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 16231–16241, 2022.

Zhen Liu, Hao Zhu, Qi Zhang, Jingde Fu, Weibing Deng, Zhan Ma, Yanwen Guo, and Xun Cao. 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 (CVPR)*, pp. 2713–2722, June 2024.

Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In *ECCV*, 2020.

Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics*, 41(4):102:1–102:15, July 2022. doi: 10.1145/3528223.3530127. URL https://doi.org/10.1145/3528223.3530127.

Yuan Ni, Zhantao Chen, Shizhou Xu, Cheng Peng, Rajan Plumley, Chun Hong Yoon, Jana B Thayer, and Joshua J Turner. Multi-resolution enhancement for full-spectrum neural representations. *Nature Machine Intelligence*, pp. 1–14, 2026.

Nasim Rahaman, Aristide Baratin, Devansh Arpit, Felix Draxler, Min Lin, Fred Hamprecht, Yoshua Bengio, and Aaron Courville. On the spectral bias of neural networks. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 5301–5310. PMLR, 2019. URL https://proceedings.mlr.press/v97/rahaman19a.html.

Vishwanath Saragadam, Daniel LeJeune, Jasper Tan, Guha Balakrishnan, Ashok Veeraraghavan, and Richard G. Baraniuk. WIRE: Wavelet implicit neural representations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 18507–18516, June 2023.

Liyue Shen, John Pauly, and Lei Xing. NeRP: Implicit neural representation learning with prior embedding for sparsely sampled image reconstruction. *IEEE Transactions on Neural Networks and Learning Systems*, 35(1):770–782, 2024. doi: 10.1109/TNNLS.2022.3177134.

Vincent Sitzmann, Julien N. P. Martel, Alexander W. Bergman, David B. Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. In *Advances in Neural Information Processing Systems*, volume 33, pp. 7462–7473. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper/2020/hash/53c04118df112c13a8c34b38343b9c10-Abstract.html.

Matthew Tancik, Pratul P. Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan T. Barron, and Ren Ng. Fourier features let networks learn high frequency functions in low dimensional domains. In *Advances in Neural Information Processing Systems*, volume 33, pp. 7537–7547. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper/2020/hash/55053683268957697aa39fba6f231c68-Abstract.html.

Pandula Thennakoon, Avishka Ranasinghe, Mario De Silva, Buwaneka Epakanda, Roshan Godaliyadda, Parakrama Ekanayake, and Vijitha Herath. COSMO-INR: Complex sinusoidal modulation for implicit neural representations. In *International Conference on Learning Representations (ICLR)*, 2026. URL https://openreview.net/forum?id=yGJrvSU6wK.

Michael Unser. Splines: A perfect fit for signal and image processing. *IEEE Signal Processing Magazine*, 16(6):22–38, 1999.

Michael Unser. Sampling—50 years after Shannon. *Proceedings of the IEEE*, 88(4):569–587, April 2000.

Michael Unser and Akram Aldroubi. A general sampling theory for nonideal acquisition devices. *IEEE Transactions on Signal Processing*, 42(11):2915–2925, November 1994. doi: 10.1109/78.330352.

Michael Unser, Akram Aldroubi, and Murray Eden. B-spline signal processing: Part I—theory. *IEEE Transactions on Signal Processing*, 41(2):821–833, 1993.

Shu Wang, Yanbo Gao, Shuai Li, Chong Lv, Xun Cai, Chuankun Li, Hui Yuan, and Jinglin Zhang. MetricGrids: Arbitrary nonlinear approximation with elementary metric grids based implicit neural representation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 21381–21391, June 2025.

Mustafa B. Yaldiz, Ishit Mehta, Nithin Raghavan, Andreas Meuleman, Tzu-Mao Li, and Ravi Ramamoorthi. Spectral prefiltering of neural fields. In *SIGGRAPH Asia 2025 Conference Papers*, pp. 1–12, 2025. doi: 10.1145/3757377.3763901.

Siqi Ye, Liyue Shen, Md Tauhidul Islam, and Lei Xing. Super-resolution biomedical imaging via reference-free statistical implicit neural representation. *Physics in Medicine & Biology*, 68(20), 2023. doi: 10.1088/1361-6560/acfdf1.

A Sensitivity to Quadrature Resolution

The Spatial supervision operator is evaluated numerically using K^2 quadrature queries per measurement. We use $K = 16$ throughout the main experiments. To verify that this choice provides a sufficiently accurate numerical approximation, we repeat the reconstruction with lower quadrature orders using MetricGrids on DIV2K image `0801`, with all other settings unchanged. Table 2 shows that reconstruction quality is already close to saturation at $K = 8$.

Increasing K from 8 to 16 changes field PSNR by at most 0.033 dB across the three sampling scales. This indicates that the $K = 16$ setting used in the main experiments is already near numerical saturation for this ablation.

B Sensitivity to Gaussian-Width Mismatch

We evaluate how reconstruction changes when the Gaussian width $\tilde{\sigma}$ assumed by the Spatial supervision operator differs from the measurement-generating width σ. Table 3 reports FINER on Kodak-24 at $\times 4$, averaged over three seeds per image. The correctly specified width, $\tilde{\sigma}/\sigma = 1$, reproduces the corresponding FINER result from the main reconstruction experiment and gives the largest improvement. The mean gain over Point supervision remains positive throughout the tested $\pm 30\%$ mismatch range.

Thus, the observed benefit is not restricted to exact knowledge of the Gaussian width. Approximately specified Spatial supervision remains better than Point supervision on average across the tested range. Even at $\tilde{\sigma}/\sigma = 1.30$, the mean gain remains $+0.147$ dB. The effect weakens as the assumed width moves away from the correctly specified value, with an observed asymmetry between under- and over-estimation.

C Experimental Details

Reference field and measurement lattices. Each image defines a 512×512 reference field, which is extended continuously by bilinear interpolation. Measurements are synthesized from this continuous field through the Gaussian–pixel acquisition operator on regular lattices of 512^2, 256^2, and 128^2 samples for the $\times 1$, $\times 2$, and $\times 4$ settings, respectively. The operator uses $\sigma/h = 0.5$, where h denotes the measurement-grid spacing (equivalently, the low-resolution pixel width). Reconstruction is evaluated against the unfiltered 512×512 reference field in all settings. No intermediate resized low-resolution image is constructed.

Quadrature. The Gaussian–pixel operator is evaluated numerically using K^2 quadrature samples per measurement, with K samples along each spatial axis. Measurement synthesis uses $K = 24$, whereas the spatial supervision operator used during reconstruction uses $K = 16$. Along each axis, the quadrature nodes are equal-mass quantiles of the one-dimensional Gaussian–pixel kernel, so all two-dimensional weights equal $1/K^2$ and the kernel shape is encoded by the node locations. We intentionally use different quadrature orders for synthesis and reconstruction so that the spatial condition does not reuse the exact discrete numerical operator used to generate the measurements.

Optimization. All models are optimized for $3{,}000$ Adam updates. Within each paired comparison, the representation, initialization, optimizer, learning-rate schedule, training measurements, minibatch sequence, and number of optimization steps are held fixed; only the supervision operator changes. Point and Spatial conditions therefore start from the same initialization and are optimized using identical measurements and training schedules. We use three seeds per image, average results over seeds within each image, and treat the image as the statistical unit.

D Training and Inference Cost

Point and Spatial supervision use the same representation, trainable parameter count, optimizer, and number of optimization steps. They do not, however, have equal training cost: at $K = 16$, Spatial supervision evaluates 256 field locations per measurement, whereas Point supervision evaluates one.

Table 2: **Sensitivity to quadrature resolution.** Field PSNR (dB) of Spatial supervision for different quadrature orders using MetricGrids on DIV2K image `0801`. The main experiments use $K = 16$.

K	Queries / measurement	$\times 1$	$\times 2$	$\times 4$
4	16	30.068	27.732	23.883
8	64	30.155	27.876	23.988
16	256	**30.188**	**27.905**	**24.015**

Table 3: **Sensitivity to Gaussian-width mismatch at $\times 4$.** Mean field-PSNR gain over Point supervision for FINER on Kodak-24 as the Gaussian width $\tilde{\sigma}$ assumed by the Spatial supervision operator deviates from the measurement-generating width σ. Results are averaged over three seeds per image.

$\tilde{\sigma}/\sigma$	ΔPSNR (dB)
0.70	$+0.965$
0.85	$+1.247$
1.00	$\mathbf{+1.451}$
1.15	$+1.187$
1.30	$+0.147$

Table 4 reports controlled timing and memory measurements for SIREN, FINER, and MetricGrids on an RTX A6000 over 24 images using CUDA-synchronized wall-clock timing. Spatial supervision requires substantially more training time and peak memory, while inference remains essentially unchanged because the supervision operator is used only during training.

The additional computation is therefore incurred during training rather than through increased model size or deployment-time complexity. Across the three measured representations, Spatial supervision increases training time by approximately 9–$24\times$ and peak memory by approximately 26–$35\times$, while measured inference time differs from Point supervision by only 0.6–1.6%.

E Acquisition-Consistency Control

The main experiments keep the measurements fixed while varying the reconstruction acquisition model. We additionally perform a complementary control in which each reconstruction condition uses measurements generated by its corresponding acquisition model. Specifically, the Point condition uses point-sampled measurements with point reconstruction,

$$\psi = \tilde{\psi} = \delta,$$

whereas the Spatial condition uses measurements generated and reconstructed with the Gaussian–pixel acquisition model. This comparison tests whether the acquisition-dependent differences observed in the main experiments persist when the measurement-generation and reconstruction models correspond within each condition.

For compactness, we denote the original main-paper Point condition by ψ/δ, the corresponding Point condition by δ/δ, and the corresponding Spatial condition by ψ/ψ. We define

$$\Delta_{\mathrm{acq}} = \mathrm{PSNR}(\psi/\psi) - \mathrm{PSNR}(\delta/\delta), \tag{8}$$

where positive values indicate higher field PSNR under Spatial acquisition.

As shown in Table 5, at $\times 4$ the corresponding-operator comparison retains a 1.07–1.29 dB advantage for Spatial acquisition across SIREN, FINER, and MetricGrids, with the same direction on all 24 Kodak images for each representation. The acquisition-dependent separation observed at this scale therefore persists when each condition uses its corresponding measurement model.

At $\times 2$, the difference is smaller: $+0.27$, $+0.04$, and $+0.50$ dB for SIREN, FINER, and MetricGrids, respectively. For FINER, the difference is not statistically detectable ($12/24$ images, $p = 0.83$). This indicates that the magnitude of the acquisition effect depends on the sampling regime, with a substantially clearer separation at $\times 4$.

Table 4: **Training and inference cost.** CUDA-synchronized measurements on an RTX A6000 over 24 images. Training-time ratios denote Spatial/Point and are computed per image before aggregation; brackets report 95% bootstrap confidence intervals over images. Peak-memory ratios denote Spatial/Point. Inference ratios are computed from the median render times over the measured inference runs.

Method	Training time			Peak memory			Inference
	Point (s)	Spatial (s)	Ratio	Point (GB)	Spatial (GB)	Ratio	Ratio
SIREN	12.16	262.28	22.80 [20.41, 25.72]	0.0439	1.3111	29.8	1.016
FINER	14.08	335.33	24.40 [23.05, 26.75]	0.0515	1.8111	35.1	1.007
MetricGrids	28.96	243.51	8.95 [7.90, 10.27]	0.0376	0.9656	25.7	1.006

Table 5: **Acquisition-consistency control on Kodak-24.** Field PSNR (dB) for the original Point condition (ψ/δ), corresponding Point acquisition and reconstruction (δ/δ), and corresponding Spatial acquisition and reconstruction (ψ/ψ). Results use seed 0. Δ_{acq} compares the latter two conditions.

Scale	Backbone	ψ/δ	δ/δ	ψ/ψ	Δ_{acq}
$\times 2$	SIREN	26.00	26.40	26.67	+0.27†
	FINER	27.05	28.89	28.93	+0.04
	MetricGrids	27.05	28.70	29.20	+0.50†
$\times 4$	SIREN	24.27	24.44	25.51	**+1.07**†
	FINER	24.30	24.46	25.76	**+1.29**†
	MetricGrids	24.15	24.26	25.52	**+1.26**†

† Two-sided Wilcoxon signed-rank $p < 0.05$ over the 24 images. At $\times 4$, Spatial acquisition is higher on all $24/24$ images for every backbone ($p = 1.19 \times 10^{-7}$).

F Capacity and Optimization-Budget Controls

We test whether the reconstruction difference can be explained by insufficient capacity or optimization of the Point baseline. Using FINER on Kodak-24 at $\times 4$, we enlarge only the Point condition while keeping the measurement data and remaining training protocol unchanged. The reference model has width 256, depth 3, and 198,915 parameters. We increase the parameter count to approximately $2\times$ and $4\times$ by using widths 363 and 513, yielding 398,577 ($2.004\times$) and 794,127 ($3.992\times$) parameters, respectively. We separately extend optimization beyond the reference 3,000 updates. For these longer-training controls, the original learning-rate schedule is kept on its absolute timeline rather than rescaled to the longer run.

As shown in Table 6, increasing the Point model's capacity or optimization budget does not reduce the reconstruction gap in this setting. Approximately $2\times$ and $4\times$ parameter counts leave Spatial ahead by 1.55 and 1.83 dB, respectively, while extending Point optimization from 3,000 to 54,142 steps leaves the gap at 1.53 dB. Thus, for this FINER/Kodak-24/$\times 4$ setting, the observed difference is not explained by insufficient Point-model capacity or training duration.

G Spectral Error Computation

For an RGB image I, we compute the two-dimensional Fourier transform independently for each colour channel,

$$\widehat{I}(u, v, c) = \mathcal{F}\{I(\cdot, \cdot, c)\}(u, v), \tag{9}$$

where (u, v) denotes spatial frequency and c indexes the RGB channels. We then define the channel-averaged spectral power as

$$P_I(u, v) = \frac{1}{3} \sum_{c=1}^{3} \left| \widehat{I}(u, v, c) \right|^2 . \tag{10}$$

Table 6: **Capacity and optimization-budget controls.** Field PSNR (dB) for FINER on Kodak-24 at $\times 4$. Only the Point condition is modified. The fixed Spatial reference uses the original architecture and 3,000-step training budget and achieves 25.756 dB. The final column reports Spatial-minus-Point PSNR.

Control	Point setting	Point PSNR	Spatial − Point
Reference	198,915 params, 3k steps	24.305	+1.451
Capacity	398,577 params (2.004×)	24.206	+1.550
	794,127 params (3.992×)	23.923	+1.833
Optimization	6k steps	24.265	+1.491
	12k steps	24.256	+1.500
	54,142 steps	24.226	+1.530

Let $\bar{P}_I(\rho)$ denote the radially averaged power spectrum, where ρ is the normalized spatial frequency. We compute the corresponding log-power spectrum as

$$L_I(\rho) = \log_{10}\left(\max\left(\bar{P}_I(\rho), \epsilon\right)\right), \qquad \epsilon = 10^{-12}. \tag{11}$$

For image i and acquisition condition $a \in \{\mathrm{Point}, \mathrm{Spatial}\}$, we define the spectral error as

$$d_{i,a}(\rho) = \left| L_{i,a}^{\mathrm{recon}}(\rho) - L_i^{\mathrm{ref}}(\rho) \right|. \tag{12}$$

Smaller values therefore indicate closer agreement between the reconstructed and reference power spectra at spatial frequency ρ.

H Additional Qualitative Examples

Reconstruction examples. Figure 5 extends Figure 2 to two additional Kodak reference fields. Within each Point–Spatial comparison, the representation, measurements, initialization, optimizer, learning-rate schedule, and number of optimization steps are held fixed; only the supervision operator changes.

Across both reference fields, Spatial supervision improves reconstruction quality over Point supervision for every representation and sampling scale. The improvement holds across all 24 model–scale comparisons, with gains of $+0.46$–$+2.80$ dB on `kodim20` and $+0.07$–$+3.13$ dB on `kodim23`. The corresponding normalized absolute-error maps show similar reductions across architectures and sampling regimes, extending the qualitative behavior observed in Figure 2.

Spatial error maps. Figures 6 and 7 extend the signed spatial-error analysis of Figure 3 to additional Kodak-24 and DIV2K reference fields. Positive values of $\Delta e(\mathbf{x})$ indicate lower local reconstruction error under Spatial acquisition, whereas negative values indicate lower error under Point acquisition. The additional examples show similar spatially nonuniform differences across image content, representations, and sampling scales.

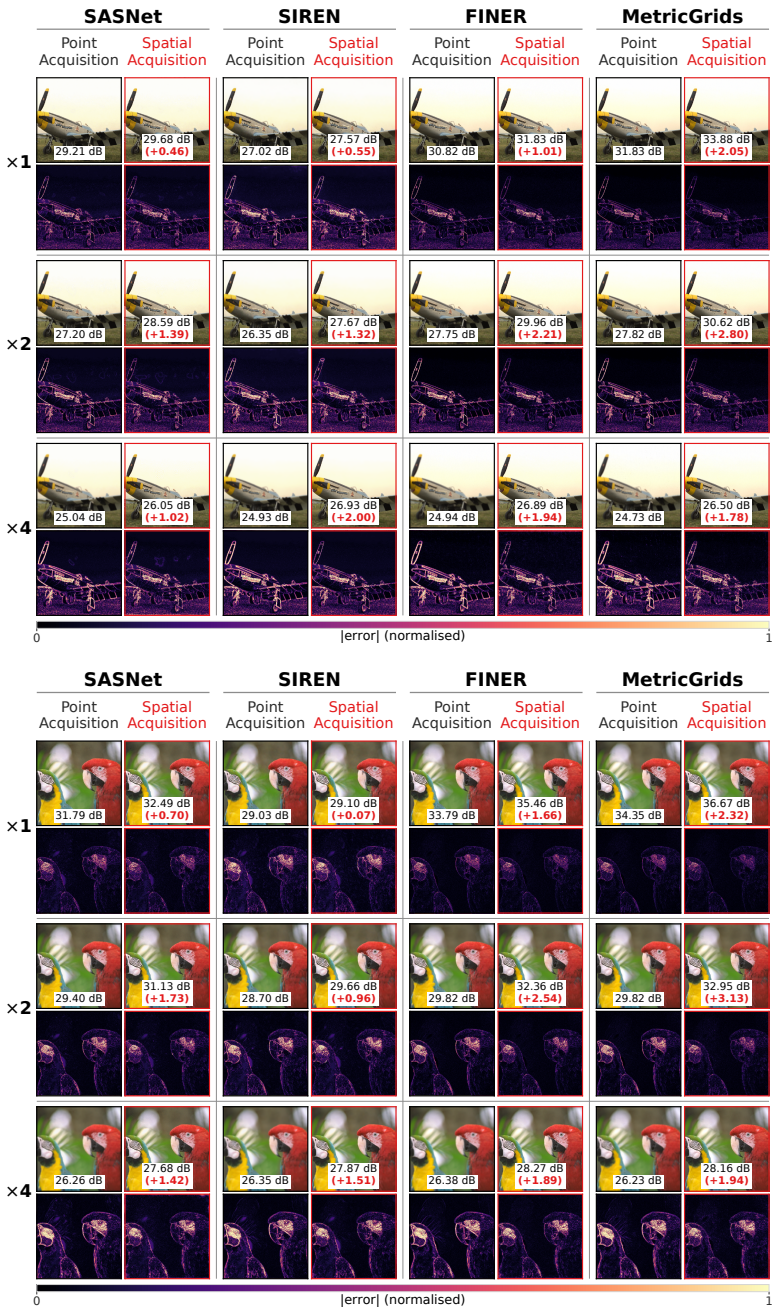

Figure 5: **Additional qualitative comparisons across sampling scales.** Top: `kodim20`. Bottom: `kodim23`. Each example shows reconstructions and normalized absolute-error maps at $\times 1$, $\times 2$, and $\times 4$ for SASNet, SIREN, FINER, and MetricGrids. Measurements are synthesized through the Gaussian–pixel acquisition operator ψ; Point supervision uses $\tilde{\psi} = \delta$, whereas Spatial supervision uses $\tilde{\psi} = \psi$. Insets report field PSNR against the unfiltered 512×512 reference field.

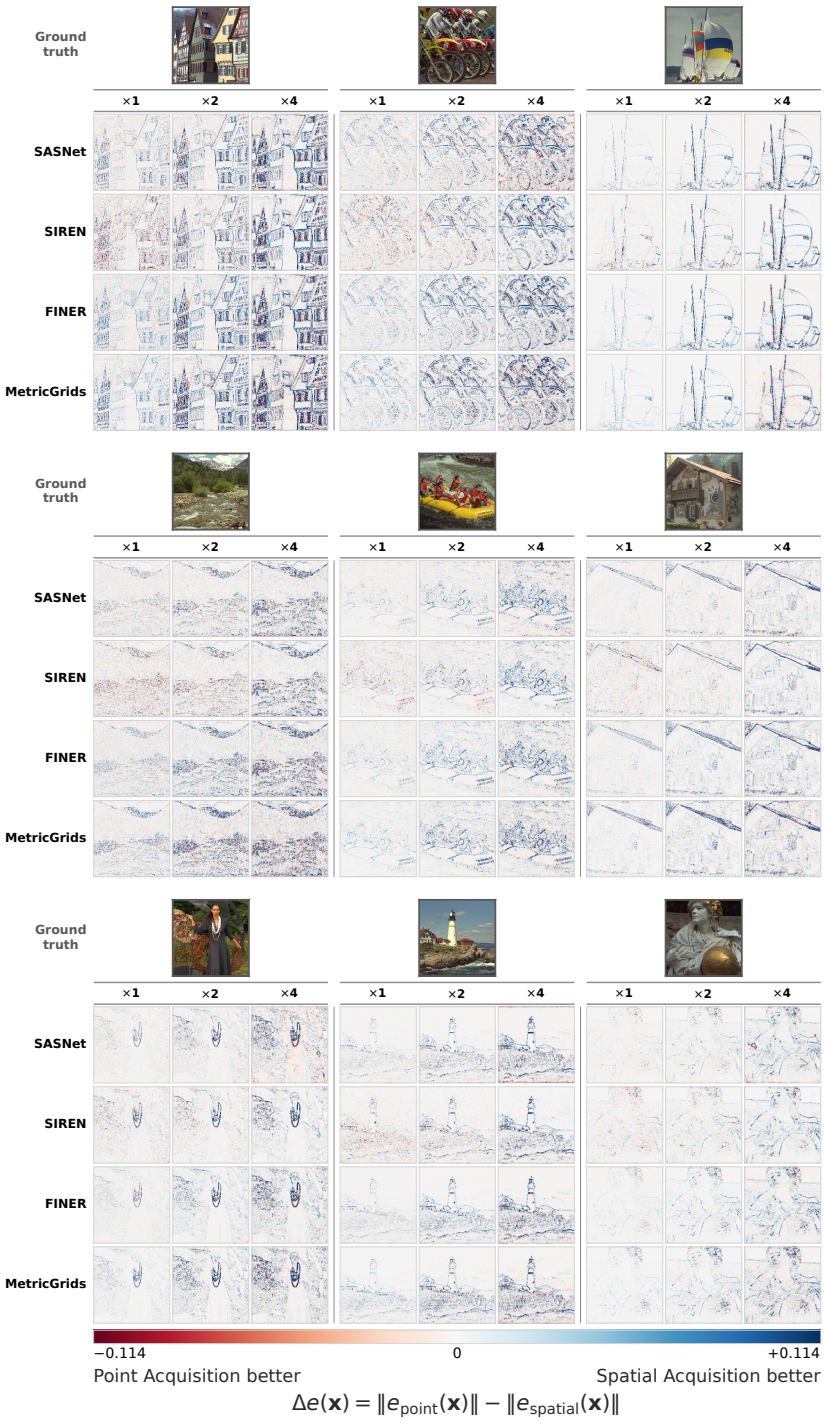

Figure 6: **Additional spatial error maps on Kodak-24.** Signed Point-minus-Spatial reconstruction-error differences for nine additional reference fields across SASNet, SIREN, FINER, and MetricGrids at $\times 1$, $\times 2$, and $\times 4$ sampling. Positive values indicate lower local reconstruction error under Spatial acquisition, whereas negative values indicate lower error under Point acquisition. The symmetric color scale is shared across all examples in the figure.

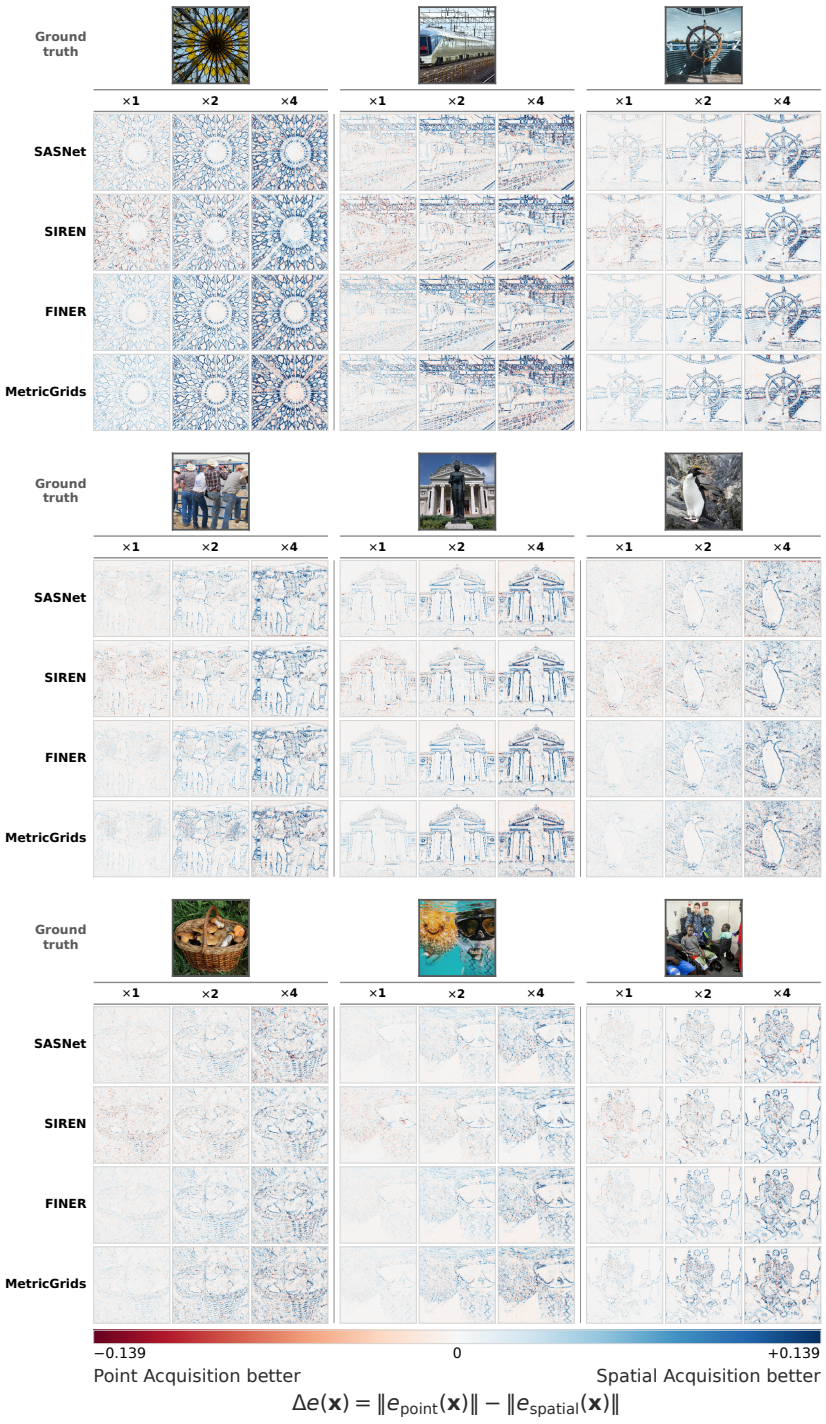

Figure 7: **Additional spatial error maps on DIV2K.** Signed Point-minus-Spatial reconstruction-error differences for nine additional reference fields across SASNet, SIREN, FINER, and MetricGrids at $\times 1$, $\times 2$, and $\times 4$ sampling. Positive values indicate lower local reconstruction error under Spatial acquisition, whereas negative values indicate lower error under Point acquisition. The symmetric color scale is shared across all examples in the figure.