

# Implicit Neural Representations for Sparse CT Reconstruction of Pulmonary Nodules: A Radiometric and Geometric Analysis

Alessandro Costanzo Ciano

5 February 2026

## Abstract

Accurate pulmonary nodule assessment relies on volumetric CT, yet clinical thick-slice acquisitions suffer from data sparsity that degrades both geometric shape and internal texture. This study evaluates Sinusoidal Representation Networks (SIREN)—a class of Implicit Neural Representations—for recovering volumetric density from sparse data. Using a 50-nodule test set from the LIDC-IDRI database, SIREN was benchmarked against longitudinal cubic and morphological interpolation under a 1-in-5 slice training protocol. A distinct performance dichotomy was observed. Radiometrically, SIREN excelled at global structural recovery, reducing global RMSE by 15.6% (0.051 vs. 0.061) and improving PSNR by 1.43 dB across the cohort. Furthermore, SIREN achieved superior global structural similarity (SSIM 0.89 vs 0.87). However, this performance inverted within the masked tumor core, where cubic interpolation demonstrated higher fidelity in both pixel-wise (PSNR 18.58 dB vs. 17.20 dB) and structural metrics. Geometrically, internal variance in the implicit representation compromised boundary definition, yielding a lower Dice Similarity Coefficient (0.69) than morphological interpolation (0.84). These results suggest that while implicit representations offer superior global signal fidelity, they introduce high-frequency variance that degrades local segmentation precision, highlighting the need for hybrid reconstruction strategies in clinical AI.

## 1 Introduction

Accurate assessment of pulmonary nodules is crucial to early lung cancer detection and treatment monitoring, as lung cancer remains the leading cause of cancer-related mortality worldwide [1]. In clinical practice, the “gold standard” for measurement relies on high-resolution Computed Tomography (CT). However, to reduce radiation dose and data storage requirements, scans are frequently reconstructed with large slice intervals (e.g., 2.5 mm to 5 mm). This degradation severely compromises volumetric measurements—vital for RECIST criteria [2]—and hampers the extraction of subtle texture features required for radiomic analysis [3].

Traditionally, the recovery of missing volumetric data has relied on linear or longitudinal cubic interpolation. While computationally efficient, these methods act as low-pass filters [4]. Mathematically, this smoothness assumption minimizes mean squared error in homogeneous regions; clinically, however, it results in the loss of high-frequency spectral information essential for characterizing pathology. Malignant nodules are distinguished by irregular internal textures and sharp density variations [5]. Consequently, while interpolation generates statistically ‘safe’ smooth approximations, it lacks the radiometric fidelity required to distinguish heterogeneous malignant nodules from homogeneous benign tissue.

Implicit Neural Representations (INRs) provide an alternative by parameterizing volumetric density as a continuous function of coordinates,  $F(\mathbf{x}) \rightarrow \sigma$ , rather than a discrete voxel grid [6]. This continuous

parameterization is theoretically suited for medical super-resolution, where anatomy is biologically continuous but radiologically discretized. Specifically, Sinusoidal Representation Networks (SIREN) [7] are investigated. By leveraging periodic activation functions to mitigate the spectral bias inherent in standard ReLU-based architectures, SIRENs are positioned to recover high-frequency anatomical gradients lost during sparse acquisition.

In this work, a comparative analysis of SIREN-based reconstruction versus standard interpolation is proposed using lung nodule data from the LIDC-IDRI dataset [8]. To simulate aggressive clinical sparsity, a strict evaluation protocol is employed where only 20% of the data (1 out of 5 slices) is used for training. Performance is assessed via a dual-axis framework: **Geometric Consistency** evaluates the definition of anatomical boundaries against a morphological baseline, while **Radiometric Fidelity** analyzes the preservation of internal Hounsfield Unit (HU) distributions using frequency analysis and PSNR.

The primary contribution of this research is the demonstration of a critical spectral trade-off in volumetric recovery. Results reveal that while linear methods effectively minimize variance within the masked tumor core, INRs are superior for global structural recovery and the preservation of high-frequency boundaries. This suggests that clinical AI reconstruction strategies must be context-dependent, balancing linear smoothing for homogeneity with implicit representations for structural sharpness.

## 2 Methodology

### 2.1 Dataset Selection and Preprocessing

This study uses the Lung Image Database Consortium (LIDC-IDRI) [8], a publicly available thoracic CT dataset. To establish a rigorous ground truth, strict inclusion criteria were enforced to ensure that only high-resolution, clinically significant data was utilized.

Scans were filtered to obtain only those with a slice thickness of  $d_z \leq 1.0$  mm. This constraint ensures that the “ground truth” volumes contain sufficient high-frequency information to validate the reconstruction capability of the model. Furthermore, to focus on the challenging heterogeneous textures characteristic of pathology, nodule selection was restricted to those with an average radiologist-assigned malignancy score greater than 3 (on a scale of 1 to 5).

For each selected candidate, a localized volume was extracted centered on the radiologist consensus bounding box with an additional spatial context padding of 10 mm. The voxel intensities were clipped to a standard lung window range of  $[-1000, 400]$  Hounsfield Units (HU) and normalized to the range  $[0, 1]$  to stabilize neural network training. Spatial coordinates  $(x, y, z)$  were normalized to lie within the unit hypercube  $[-1, 1]^3$ .

To rigorously evaluate the proposed implicit neural representation, a two-tiered data partitioning strategy was employed, addressing both the dataset-level separation for hyperparameter selection and the instance-level splitting required for the SIREN training protocol.

### 2.2 Experimental Design and Data Partitioning

**Dataset Partitioning:** The experimental cohort consisted of 57 high-malignancy lung nodules extracted from the LIDC-IDRI database. This dataset was stratified into two distinct subsets:

- **Development Set ( $N = 7$ ):** A small, representative subset used exclusively for hyperparameter tuning and network architecture search. No performance metrics from this set are reported in the final results to prevent data leakage.

- **Test Set** ( $N = 50$ ): The remaining 50 nodules served as the independent test set. These samples were reserved solely for the final evaluation of the interpolation and segmentation performance.

**Instance-Specific Training Protocol:** Unlike conventional deep learning approaches that train a single model to generalize across a dataset, the SIREN architecture acts as a continuous implicit representation for a specific signal. Consequently, a separate network instance was initialized and optimized from scratch for each nodule in the test set.

For each nodule volume, an intra-instance split was applied to assess the model’s ability to interpolate unseen data:

- **Training Views (Sparse Input):** The volumetric data was subsampled along the axial dimension with a stride of 5. Only these selected slices (representing 20% of the total volume) were used to compute the loss and update the network weights during optimization.
- **Evaluation Views (Dense Reconstruction):** The remaining intermediate slices (80% of the volume), which were unseen by the network during optimization, served as the validation set to quantify the fidelity of the dense volumetric reconstruction.

### 2.3 Implicit Neural Representation (SIREN)

To recover the continuous volumetric density function from sparse discrete sampling, a Sinusoidal Representation Network (SIREN) [7] is employed. Unlike conventional Multi-Layer Perceptrons (MLPs) that utilize piecewise linear activations (e.g., ReLU)—which bias networks toward low-frequency functions and result in over-smoothed textures—SIREN leverages periodic activation functions to enable the modeling of high-frequency spatial details [9].

The neural representation is defined as a continuous function  $F_\theta : \mathbb{R}^3 \rightarrow \mathbb{R}$ , parametrized by weights  $\theta$ . The network takes a 3D spatial coordinate  $\mathbf{x} = (x, y, z) \in [-1, 1]^3$  as input and outputs the estimated normalized radiodensity  $\hat{\sigma} \in [0, 1]$ .

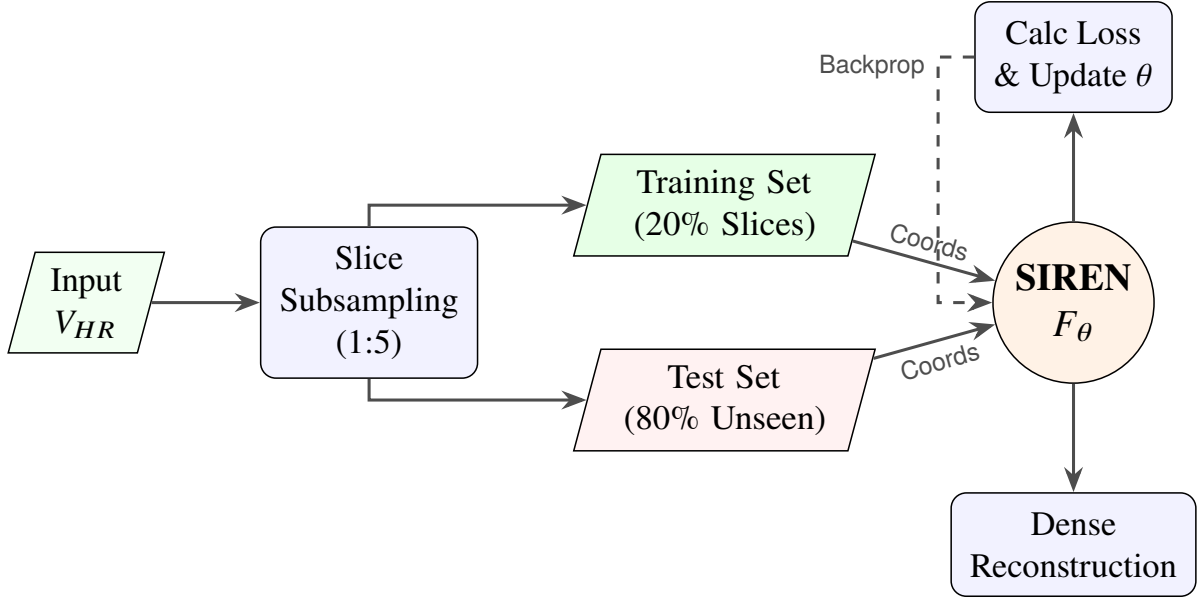
The architecture consists of a sequence of fully connected layers with sinusoidal non-linearities. For a layer  $i$ , the operation is defined as:

$$\mathbf{h}_i = \sin(\omega_0(\mathbf{W}_i \mathbf{h}_{i-1} + \mathbf{b}_i)) \quad (1)$$

where  $\mathbf{W}_i$  and  $\mathbf{b}_i$  are the weight matrix and bias vector, and  $\omega_0$  is a hyperparameter governing the spatial frequency spectrum of the signal.

**Network Architecture:** The proposed *TumorSIREN* model consists of an input layer mapping  $\mathbb{R}^3 \rightarrow \mathbb{R}^{256}$ , followed by 3 hidden layers of 256 units. To prevent overfitting on the sparse training coordinates, a Dropout layer ( $p = 0.5$ ) is introduced after every sinusoidal activation in the hidden blocks. The final layer projects the hidden features to a scalar density value  $\hat{\sigma}$  via a linear transformation without activation, allowing the network to output the full dynamic range of the normalized Hounsfield units.

The complete training and evaluation pipeline is visualized in Figure 1. Let the original high-resolution volume be denoted as  $V_{HR}$ . This framework allows for exact voxel-wise error calculation. Unlike real-world clinical scenarios where the missing intermediate data is permanently lost, this simulation ensures that the reconstructed volume  $V_{rec}$  can be directly compared against the original  $V_{HR}$  to evaluate the recovery of internal density and biological texture.



**Figure 1: Instance-Specific Training Pipeline.** A unique SIREN network is optimized for each patient. The model overfits the sparse Training Set (top path) to learn the implicit volume, then reconstructs the unseen Evaluation Views (bottom path).

**Initialization Scheme:** Proper initialization is critical for the convergence of deep sinusoidal networks. The scheme proposed by Sitzmann et al. [7] is followed to ensure that the distribution of activations remains standard normal throughout the network. Weights are initialized from a uniform distribution  $\mathcal{U}(-\frac{\sqrt{6/n}}{\omega_0}, \frac{\sqrt{6/n}}{\omega_0})$ , where  $n$  is the number of input features. A value of  $\omega_0 = 30$  is used to prioritize high-frequency texture recovery.

## 2.4 Optimization and Training Strategy

The training process is formulated as a coordinate-based regression task. The network is optimized to minimize the reconstruction error on the *Observed Slices* (training set  $\mathcal{D}_{train}$ ), while generalization is monitored on the *Unobserved Slices* (validation set  $\mathcal{D}_{val}$ ). Each model instance of the test set was trained for 2,000 epochs, a value selected to ensure full convergence, as evidenced by the asymptotic behavior of the loss curves during preliminary evaluation on the development cohort.

**Loss Function:** The loss employed is Mean Squared Error (MSE) loss between the predicted density  $F_\theta(\mathbf{x})$  and the ground truth voxel intensity  $\sigma_{GT}(\mathbf{x})$  :

$$\mathcal{L}(\theta) = \frac{1}{|\mathcal{B}|} \sum_{\mathbf{x} \in \mathcal{B}} |F_\theta(\mathbf{x}) - \sigma_{GT}(\mathbf{x})|^2 \quad (2)$$

where  $\mathcal{B}$  represents a mini-batch of coordinate-intensity pairs sampled from  $\mathcal{D}_{train}$ .

**Optimization Protocol:** The network is trained using the Adam optimizer with an initial learning rate of  $5 \times 10^{-5}$  and a batch size of 2,048 coordinates. To enhance training stability and convergence, a comprehensive scheduling and regularization strategy was implemented:

1. **Learning Rate Schedule:** A Linear Warmup is utilized for the first 50 epochs to stabilize early

gradients, followed by Cosine Annealing with Warm Restarts. The primary cycle length is set to  $T_0 = 100$  epochs with a period multiplier of  $T_{mult} = 2$ , allowing the optimizer to escape local minima during restart phases.

2. **Gradient Clipping:** To mitigate exploding gradients inherent to high-frequency periodic activations, gradient norms are clipped to a maximum value of 1.0.
3. **Coordinate Jittering:** To enforce local smoothness and prevent the network from memorizing the discrete training grid, Gaussian noise injection is applied to the input coordinates during training:  $\mathbf{x}' = \mathbf{x} + \mathcal{N}(0, \sigma_{noise}^2)$ , with  $\sigma_{noise} = 10^{-4}$ .

## 2.5 Comparative Baselines

To validate the efficacy of the proposed INR approach, performance is benchmarked against standard clinical reconstruction techniques.

**Density Baseline (Longitudinal Cubic Interpolation):** To reconstruct the missing intermediate slices, 1D cubic spline interpolation is applied along the longitudinal (z) axis. Specifically, for each fixed spatial coordinate  $(x, y)$  in the axial plane, a piecewise cubic polynomial is fitted to the voxel intensities of the sparse training slices. Missing values are recovered by sampling this continuous spline at the target z-coordinates. Unlike spatial interpolation methods (e.g., trilinear or tricubic) which smooth data across neighboring voxels, this approach preserves the original spatial resolution within the  $xy$ -plane, effectively testing the assumption that tissue density evolves smoothly across the slice gap.

**Geometric Baseline (Morphological Linear Interpolation):** To provide a rigorous upper bound for shape recovery, a shape-based interpolation method is employed. Given the binary masks of two training slices  $M_z$  and  $M_{z+k}$ , the intermediate mask at slice  $z + i$  is generated via linear interpolation of the binary morphology:

$$M_{z+i} = ((1 - \alpha) \cdot M_z + \alpha \cdot M_{z+k}) > 0.5 \quad (3)$$

where  $\alpha = i/k$ . This baseline assumes that the tumor boundary creates a convex hull that transforms linearly across the slice gap. This method serves as the Geometric Upper Bound for geometric consistency in this study, representing the best possible shape estimation derivable from sparse data without learned semantic priors.

## 2.6 Evaluation Axis I: Radiometric Fidelity

The primary objective of this study is to assess the recovery of internal tumor density (texture). Evaluation is performed strictly on the *Unobserved Slices* of the Test Set. Let  $V_{GT}$  be the ground truth volume and  $V_{rec}$  be the reconstructed volume.

To quantify absolute density errors and signal fidelity, four complementary metrics are employed:

- **Mean Absolute Error (MAE):** MAE provides a linear representation of the average deviation in Hounsfield Units. It is robust to outliers and directly reflects the average accumulation of density error per voxel:

$$MAE = \frac{1}{N} \sum_{i=1}^N |V_{rec}^{(i)} - V_{GT}^{(i)}| \quad (4)$$

- **Root Mean Squared Error (RMSE):** Unlike MAE, RMSE penalizes large errors heavily. This is critical in medical imaging, where large local deviations can mimic false calcifications or obscure low-density necrosis:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (V_{\text{rec}}^{(i)} - V_{\text{GT}}^{(i)})^2} \quad (5)$$

- **Peak Signal-to-Noise Ratio (PSNR):** PSNR assesses the fidelity of the signal reconstruction relative to noise and interpolation artifacts. As the data is normalized to the range  $[0, 1]$ , the dynamic range is  $\text{MAX} = 1.0$ :

$$\text{PSNR} = 20 \cdot \log_{10} \left( \frac{\text{MAX}}{\sqrt{\text{MSE}}} \right) \quad (6)$$

Higher PSNR indicates superior preservation of structural information and reduced blurring of high-frequency internal textures.

- **Structural Similarity Index (SSIM):** Unlike voxel-wise metrics, SSIM assesses the recovery of local patterns by comparing luminance, contrast, and structure. This is particularly relevant for maintaining the textural heterogeneity of tumor tissue. The index is defined as:

$$\text{SSIM}(V_{\text{rec}}, V_{\text{GT}}) = \frac{(2\mu_{\text{rec}}\mu_{\text{GT}} + C_1)(2\sigma_{\text{rec},\text{GT}} + C_2)}{(\mu_{\text{rec}}^2 + \mu_{\text{GT}}^2 + C_1)(\sigma_{\text{rec}}^2 + \sigma_{\text{GT}}^2 + C_2)} \quad (7)$$

where  $\mu$  and  $\sigma$  denote the local mean and standard deviation, respectively, and  $\sigma_{\text{rec},\text{GT}}$  represents the local covariance. The constants  $C_1 = (k_1 L)^2$  and  $C_2 = (k_2 L)^2$  are included to ensure numerical stability when the denominators approach zero. Following standard practice, the dynamic range  $L = 1.0$  is used with stability coefficients  $k_1 = 0.01$  and  $k_2 = 0.03$ .

**Stratified Evaluation:** Standard global metrics over the full slice can be dominated by high-contrast transitions in the lung parenchyma (e.g., air-tissue boundaries), potentially masking performance differences within the subtler soft-tissue range of the lesion. To address this, the evaluation is stratified into two spatial domains:

1. **Global Field of View:** Metrics are computed over the entire spatial domain of the slice to assess general anatomical consistency and boundary recovery.
2. **Intra-tumoral Region:** Metrics are restricted to the voxel set  $\Omega_{\text{tumor}} = \{\mathbf{x} \mid M_{\text{GT}}(\mathbf{x}) = 1\}$ , defined by the ground truth consensus mask. This isolates the reconstruction fidelity specifically within the malignant tissue, ensuring that background noise or high-frequency parenchymal edges do not bias the assessment of internal tumor texture.

## 2.7 Evaluation Axis II: Geometric Consistency

While radiometric metrics assess internal texture, geometric metrics evaluate the fidelity of the tumor boundary.

For the baseline method, the experimental setup differs due to the nature of the algorithm. As a deterministic, non-learning-based approach, thresholding operations are applied directly to the full consensus mask (utilizing all volumetric slices). Global thresholding is susceptible to false positives from distant anatomical structures. To focus specifically on nodule boundary definition, a Region of

Interest (ROI) strategy is employed. For every nodule, a search volume  $\mathcal{V}_{ROI}$  is generated by expanding the ground truth binary mask  $M_{GT}$  via 3D morphological dilation with a spherical structuring element ( $r = 3.0$  mm). Within this constrained volume, binary prediction masks  $\hat{M}$  are generated by thresholding the reconstructed density  $V_{rec}$  at clinically relevant Hounsfield Unit (HU) levels. To assess robustness, a sensitivity analysis is performed across a range of thresholds  $\tau \in \{-500, -450, -400, -350, -300\}$  HU:

$$\hat{M}_\tau = (V_{rec}(\mathbf{x}) > \tau) \cap (\mathbf{x} \in \mathcal{V}_{ROI}) \quad (8)$$

The resulting binary masks are evaluated using three complementary metrics based on physical spatial dimensions:

- **Dice Similarity Coefficient (DSC):** DSC measures the volumetric overlap between the predicted mask  $M_{pred}$  and ground truth mask  $M_{GT}$ . It is defined as:

$$\text{DSC} = \frac{2|M_{pred} \cap M_{GT}|}{|M_{pred}| + |M_{GT}|} \quad (9)$$

- **Hausdorff Distance (HD):** To measure boundary distortion, the Symmetric Hausdorff Distance is computed in physical units (mm). Let  $S_{pred}$  and  $S_{GT}$  be the sets of surface voxels for the predicted and ground truth masks, respectively. The metric is defined as:

$$H(S_{pred}, S_{GT}) = \max\{h(S_{pred}, S_{GT}), h(S_{GT}, S_{pred})\} \quad (10)$$

where  $h(A, B) = \max_{a \in A} \{\min_{b \in B} d(a, b)\}$  is the directed Hausdorff distance and  $d(\cdot)$  denotes the Euclidean distance. This provides a robust, worst-case estimate of boundary deviation.

- **Absolute Relative Volume Error (ARVE):** ARVE quantifies the percentage error in total tumor volume estimation, normalized by the ground truth volume:

$$\text{ARVE} = \frac{|V_{pred} - V_{GT}|}{V_{GT}} \quad (11)$$

### 3 Results

On the hold-out test set ( $N = 50$ ), a distinct performance gap was observed between shape-based and density-based reconstruction methods.

#### 3.1 Radiometric Fidelity and Signal Reconstruction

SIREN showed significantly higher performances for the global field of view with respect to the masked tumor core (Table 1).

**Global Reconstruction (Full Slice):** Over the full spatial domain, SIREN quantitatively outperformed the interpolation baseline. Averaged across the test cohort, SIREN reduced the global RMSE by 15.6% ( $0.051 \pm 0.016$  vs.  $0.061 \pm 0.020$ ) and achieved a higher Global PSNR of **26.26** dB compared to 24.83 dB for interpolation (+5.8%). Crucially, in terms of perceptual quality, SIREN demonstrated superior structural preservation, achieving a Global SSIM of **0.893** compared to 0.869 for interpolation (+2.8%).

**Table 1: Radiometric Reconstruction Performance (Global vs. Local).** Comparison of voxel-wise error metrics on unobserved test slices ( $N = 50$ ). Values represent mean  $\pm$  std. dev.  $\Delta$  indicates relative performance of SIREN vs. Interpolation.

Metric	Global (Full Slice)			Local (Intra-Tumour Mask)		
	Interp.	SIREN	$\Delta$ (%)	Interp.	SIREN	$\Delta$ (%)
MAE	$0.033 \pm 0.012$	<b><math>0.031 \pm 0.011</math></b>	-8.2%	<b><math>0.084 \pm 0.029</math></b>	$0.109 \pm 0.053$	+29.9%
RMSE	$0.061 \pm 0.020$	<b><math>0.051 \pm 0.016</math></b>	-15.6%	<b><math>0.123 \pm 0.038</math></b>	$0.151 \pm 0.067$	+23.1%
PSNR (dB)	$24.83 \pm 2.96$	<b><math>26.26 \pm 2.84</math></b>	+5.8%	<b><math>18.58 \pm 2.58</math></b>	$17.20 \pm 3.80$	-7.5%
SSIM	$0.869 \pm 0.089$	<b><math>0.893 \pm 0.077</math></b>	+2.8%	<b><math>0.779 \pm 0.135</math></b>	$0.746 \pm 0.134$	-4.2%

Qualitative results of the density reconstruction across the full field of view are displayed in Figure 2 (Columns 1-3).

**Local Reconstruction (Intra-Tumour):** Within the masked tumor core, both methods yielded significantly lower error rates. Cubic Interpolation achieved an Intra-Tumour RMSE of **0.123**, whereas SIREN resulted in 0.151 (+23.1% relative error). Similarly, Interpolation maintained a higher local PSNR (18.58 dB) compared to SIREN (17.20 dB). This trend extended to structural metrics; Interpolation retained a higher Local SSIM (**0.779**) compared to SIREN (0.746), indicating that the implicit representation struggled to cohere the specific structural arrangement of the intratumoral heterogeneity. The spatial distribution of these residuals is visualized in the difference heatmaps (Figure 2, Columns 4-5).

### 3.2 Texture Preservation and Spectral Analysis

To investigate the source of the intra-tumoral variance observed in the radiometric analysis, statistical moments and frequency distributions were evaluated. Results are summarized in Table 2 and Figure 3.

**Intensity Distribution:** The ground truth nodule tissue exhibited a non-Gaussian intensity distribution with a mean Kurtosis of 4.62. Cubic Interpolation produced a flattened distribution with a reduced Kurtosis of 3.13.

**Variance and Frequency:** Variance preservation analysis indicated distinct spectral behaviors. Cubic Interpolation resulted in a standard deviation ratio ( $\sigma_{pred}/\sigma_{GT}$ ) of  $1.07 \pm 0.14$  within the tumor core ( $p = 0.0015$ ). SIREN further expanded the variance, yielding a ratio of  $1.10 \pm 0.16$  ( $p < 0.001$ ). A paired t-test confirmed the difference in variance preservation between SIREN and Interpolation was statistically significant ( $t = -2.217$ ,  $p = 0.031$ ).

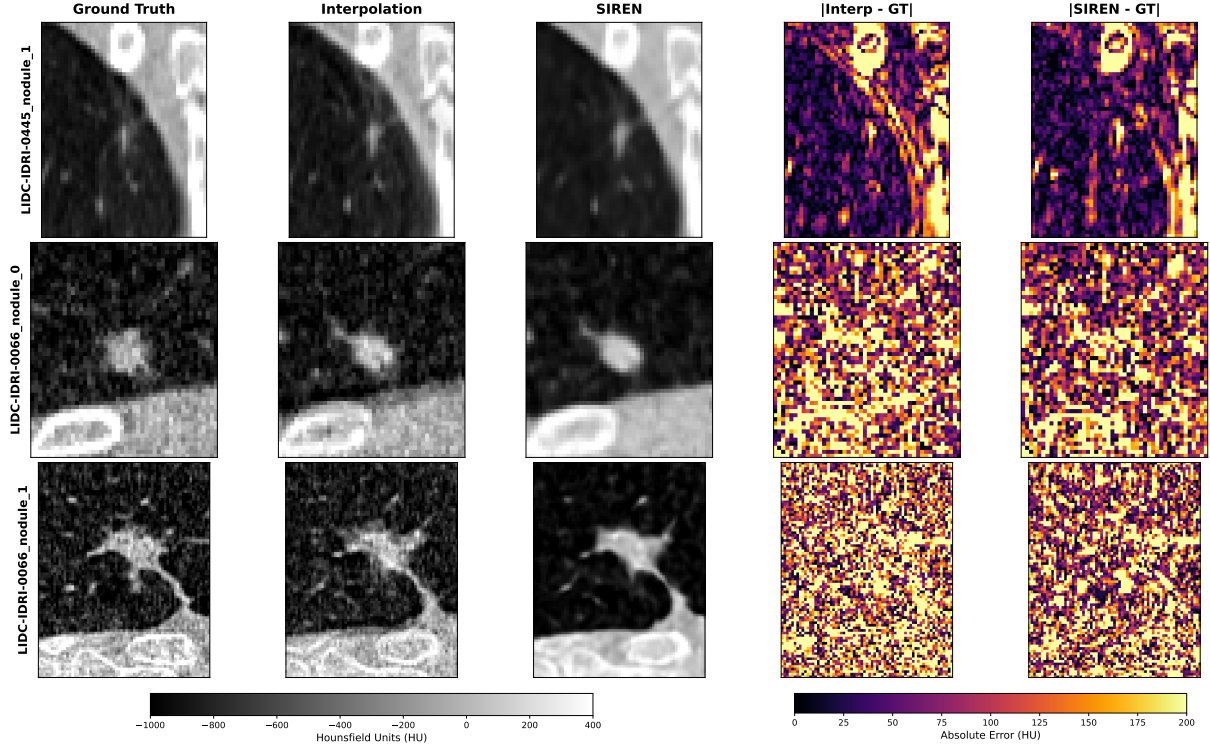
Power Spectral Density (PSD) analysis (Figure 3B) revealed that at low spatial frequencies ( $k < 2$  cycles/pixel), SIREN exhibited elevated power levels compared to both the ground truth and the baseline. At medium and high spatial frequencies ( $k \geq 2$ ), the spectral profiles of all three representations converged.

### 3.3 Geometric Consistency and Segmentation

Segmentation masks derived from density reconstructions were benchmarked against the morphological linear interpolation baseline. Table 3 summarizes the comparative performance.

The Morphological Baseline provided the upper bound for shape fidelity, achieving a Dice Similarity Coefficient (DSC) of  $0.837 \pm 0.061$  and a Hausdorff Distance (HD) of  $1.98 \pm 0.55$  mm. This outperformed the optimal thresholds for both density-based reconstruction methods across all metrics.





**Figure 2: Qualitative comparison of volumetric reconstruction on unseen intermediate slices.**

Rows represent three distinct nodules from the test set; slices correspond to the central planes excluded from training. **Row 1:** A small nodule (LIDC-IDRI-0445) illustrating the difficulty of reconstructing fine-scale features. **Rows 2-3:** Larger nodules (LIDC-IDRI-0066) showing improved fidelity in SIREN. The third row highlights a complex case where the nodule is attached to other non-tumorous anatomical objects. **Columns 1-3:** Visual comparison of ground truth (GT), Longitudinal Cubic Interpolation, and SIREN. SIREN reconstructions are notably smoother in the anatomical background, consistently limiting noise artifacts compared to the interpolation baseline which exhibits blurring. **Columns 4-5:** Absolute difference heatmaps ( $|Reconstruction - GT|$ ). Interpolation errors are structurally concentrated at high-contrast edges (e.g., the nodule boundary in Row 3), whereas SIREN errors appear as distributed high-frequency noise.

Sensitivity to threshold variations is detailed in Figure 4. In volumetric estimation SIREN minimized error at lower thresholds (ARVE 0.582 at -500 HU), whereas Cubic Interpolation proved more robust at higher thresholds (ARVE 0.444 at -300 HU).

In terms of boundary adherence (Figure 4C), SIREN exhibited performance degradation as the threshold increased. While Cubic Interpolation maintained a stable HD ( $\approx 3.05$  mm) between -500 and -350 HU, the SIREN HD increased monotonically from 3.09 mm to 3.39 mm. Similarly, Cubic Interpolation consistently achieved higher DSC values than SIREN, with the margin widening to 0.027 at -300 HU (0.662 vs. 0.635). Qualitative inspection (Figure 5) reveals that density-based methods produced boundary distortions, particularly in juxtapleural contexts, which were not present in the Morphological baseline.

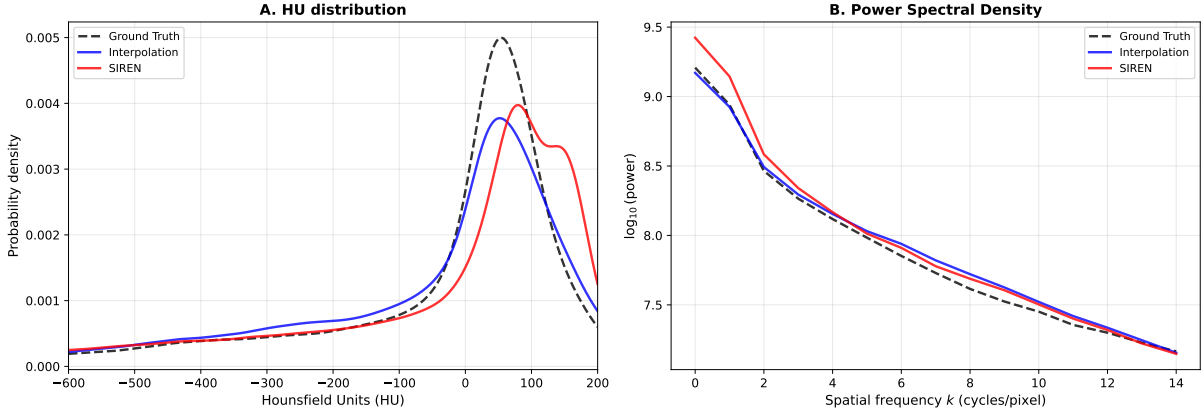
### 3.4 Summary of Findings

Evaluation on the hold-out test set ( $N = 50$ ) demonstrated a distinct performance trade-off between global structural recovery and local density precision.

In terms of **Radiometric Fidelity**, results depended on the spatial domain. Globally, SIREN proved

**Table 2: Frequency and Statistical Texture Analysis.**  $\mu$  and  $\sigma$  represent Mean and Std Dev in Hounsfield Units (HU).

Metric	Full Volume (Global)			Intra-Tumour (Local)		
	GT	Interp.	SIREN	GT	Interp.	SIREN
$\mu$ (HU)	$-638 \pm 149$	$-642 \pm 143$	$-646 \pm 145$	$-133 \pm 162$	$-158 \pm 157$	$-185 \pm 176$
$\sigma$ (HU)	$359 \pm 111$	$345 \pm 100$	$343 \pm 114$	$205 \pm 65$	$213 \pm 46$	$222 \pm 56$
$\sigma$ Ratio	1.00	$0.97 \pm 0.04$	$0.94 \pm 0.07$	1.00	$1.07 \pm 0.14$	$1.10 \pm 0.16$
Kurtosis	$7.49 \pm 8.89$	$6.18 \pm 6.73$	$6.81 \pm 7.90$	$4.62 \pm 5.15$	$3.13 \pm 1.61$	$3.59 \pm 2.89$



**Figure 3: Intra-tumour texture analysis.** Analyses are restricted to voxels within the tumour mask. (A) **Hounsfield unit (HU) distribution:** Kernel density estimate of intensity within the ROI. Ground truth (black dashed) exhibits a sharp, heavy-tailed profile; interpolation (blue) flattens the distribution, reducing peak density; SIREN (red) recovers a sharper peak but displays a secondary density shoulder in the high-HU range (approx. 100-150 HU). (B) **Power spectral density:** Radial average of magnitude-squared 2D FFT per slice (y-axis:  $\log_{10}(\text{power})$ ). SIREN exhibits slightly elevated power at low spatial frequencies ( $k < 2$ ), while the spectral profiles of ground truth, Interpolation, and SIREN are largely congruent at medium and high spatial frequencies.

superior, reducing RMSE by 15.6% and improving PSNR by 5.8% relative to the interpolation baseline. Conversely, within the masked tumor core, Cubic Interpolation minimized variance more effectively, yielding higher intra-tumoral PSNR (18.58 dB vs. 17.20 dB).

Texture analysis indicated that Cubic Interpolation acted as a low-pass filter, flattening the intensity distribution (Kurtosis 3.13). SIREN recovered a distribution shape closer to the ground truth (Kurtosis 3.59) but introduced statistically significant variance expansion ( $\sigma_{ratio} = 1.10$ ) and elevated power in the low-frequency spectrum.

Regarding **Geometric Consistency**, the Morphological baseline established the upper bound for performance (DSC 0.84). Comparing density reconstruction methods, Longitudinal Cubic Interpolation consistently outperformed SIREN. While Interpolation maintained stable boundary adherence across thresholds ( $HD \approx 3.05$  mm), SIREN exhibited degradation at higher Hounsfield levels, with surface distance error increasing to 3.39 mm.

## 4 Discussion

The reconstruction of sparse volumetric medical data represents a fundamental inverse problem where the goal is to recover a continuous biological signal from discrete, undersampled observations. In this work, the efficacy of Implicit Neural Representations—specifically Sinusoidal Representation Networks

**Table 3: Geometric Consistency Metrics.** Comparison of the Morphological Baseline against density-based methods across varying Hounsfield Unit (HU) thresholds.

Method	Thresh.	DSC ( $\uparrow$ )	HD (mm) ( $\downarrow$ )	ARVE ( $\downarrow$ )
<b>Morphological</b>	–	<b><math>0.837 \pm 0.061</math></b>	<b><math>1.98 \pm 0.55</math></b>	<b><math>0.039 \pm 0.048</math></b>
<b>Cubic Interp.</b>	-500 HU	$0.696 \pm 0.152$	$3.04 \pm 0.61$	$0.674 \pm 0.727$
	-450 HU	$0.697 \pm 0.154$	$3.04 \pm 0.62$	$0.581 \pm 0.644$
	-400 HU	$0.691 \pm 0.160$	$3.05 \pm 0.61$	$0.517 \pm 0.568$
	-350 HU	$0.680 \pm 0.168$	$3.08 \pm 0.59$	$0.471 \pm 0.508$
	-300 HU	$0.662 \pm 0.179$	$3.21 \pm 1.23$	$0.444 \pm 0.449$
<b>SIREN</b>	-500 HU	$0.691 \pm 0.160$	$3.09 \pm 0.65$	$0.582 \pm 0.682$
	-450 HU	$0.684 \pm 0.165$	$3.11 \pm 0.65$	$0.516 \pm 0.612$
	-400 HU	$0.672 \pm 0.175$	$3.15 \pm 0.69$	$0.483 \pm 0.552$
	-350 HU	$0.657 \pm 0.184$	$3.31 \pm 1.25$	$0.471 \pm 0.492$
	-300 HU	$0.635 \pm 0.198$	$3.39 \pm 1.33$	$0.471 \pm 0.445$

(SIREN)—is evaluated against standard clinical interpolation for recovering lung nodules from sparse CT acquisitions. This work represents the first attempt to rigorously quantify the disparity between *global* anatomical recovery and *local* intra-tumoral fidelity within the framework of SIREN-based medical reconstruction. The findings reveal a critical dichotomy: while SIRENs offer superior global regularization and denoising, they exhibit a paradoxical loss of fidelity within the heterogeneous tumor core, suggesting that the optimal reconstruction strategy is highly dependent on the downstream clinical task.

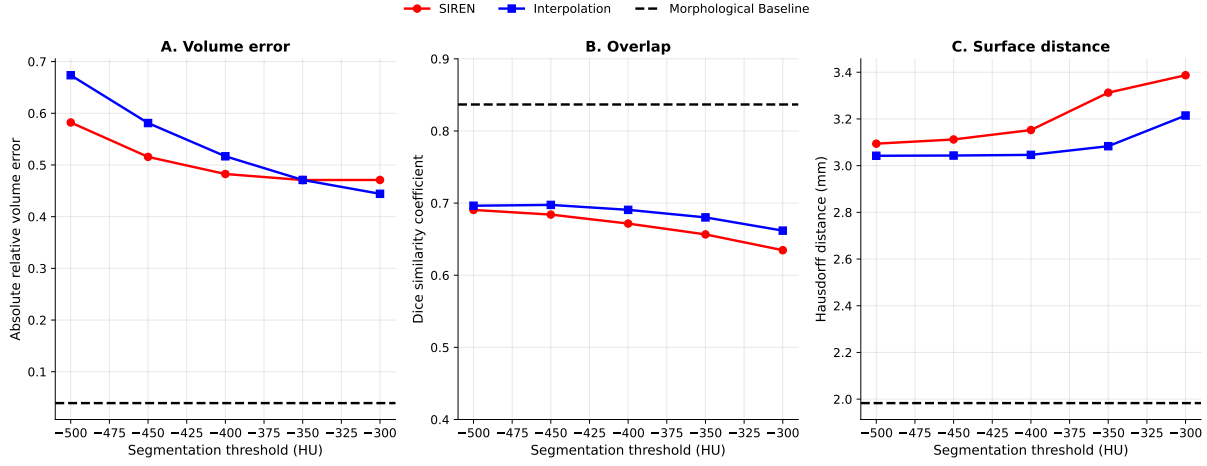
#### 4.1 The Global-Local Performance Gap

The most significant finding of this study is the divergence in performance between the global field of view and the localized tumor volume, a trend consistent across both pixel-wise (RMSE/PSNR) and perceptual (SSIM) metrics.

Globally, the SIREN architecture demonstrated a clear advantage, improving PSNR by 5.8% and reducing RMSE by more than 15% compared to cubic interpolation. More importantly, the improvement in Global SSIM (0.89 vs 0.87) indicates that SIREN acts as an effective global prior. In medical imaging, where Structural Similarity (SSIM) often correlates better with radiological assessment than absolute pixel error, this suggests that SIREN successfully models the low-frequency macro-architecture of the thorax features (e.g., chest wall, rib cage, and parenchymal boundaries).

However, this global superiority masked a critical failure within the region of clinical interest. Within the tumor core, the performance inverted. The interpolation baseline achieved significantly higher spectral fidelity and structural similarity (Local SSIM 0.78 vs 0.75). This distinction is crucial: while RMSE quantifies absolute intensity differences, SSIM measures the degradation of structural information (luminance, contrast, and structure). The drop in Local SSIM confirms that the “variance expansion” observed in the spectral analysis ( $\sigma_{ratio} = 1.10$ ) is not merely a shift in intensity distribution, but a degradation of the spatial arrangement of tissue textures.

This can be attributed to the fundamental difference in signal characteristics. Global anatomy is dominated by lower-frequency structural components which the SIREN parameterizes efficiently. Conversely, malignant nodule texture is characterized by stochastic, high-frequency heterogeneity. The SIREN, driven by the inductive bias of its periodic activations, effectively “hallucinates” high-frequency texture in the gap regions [10]. While this results in a kurtosis closer to the ground truth, the generated



**Figure 4: Geometric metrics versus segmentation threshold (HU).** The horizontal dashed line indicates the **simple morphological propagation** baseline, which consistently outperforms both SIREN and longitudinal cubic interpolation across all three metrics (DSC 0.84, HD 1.98 mm, ARVE 0.04). (A) **Absolute relative volume error (ARVE)**: SIREN achieves lower error at lower thresholds (−500 to −425 HU) but is overtaken by Interpolation at higher thresholds. (B) **Dice similarity coefficient (DSC)**: Cubic interpolation consistently achieves higher overlap than SIREN, with the performance gap widening at higher thresholds. (C) **Hausdorff distance (HD)**: Interpolation attains lower (better) HD than SIREN across thresholds; SIREN shows a marked increase in surface distance error as the threshold increases.

texture lacks the specific structural coherence of the original pathology, leading to penalized SSIM scores. This highlights a limitation in current INR architectures: the inability to decouple high-frequency signal recovery from stochastic noise generation in the absence of dense sampling.

## 4.2 Geometric Fidelity and the Boundary Problem

Accurate volume estimation is a prerequisite for RECIST-based treatment monitoring. The geometric analysis demonstrates that for pure shape recovery, simpler is better. The morphological linear interpolation baseline provided the tightest upper bound for segmentation accuracy (DSC 0.84), outperforming both density-based reconstruction methods.

The degradation of SIREN’s performance at valid tumor boundaries (Figure 4) is likely a manifestation of the Gibbs phenomenon, adapted to neural networks. As the network attempts to fit the sharp step-function of the tumor density boundary using continuous sinusoidal basis functions, it introduces ringing artifacts—oscillations in density values near the edge. This explains the monotonic increase in Hausdorff Distance as the thresholding stringency increased. While cubic interpolation is biologically naive, its inherent smoothness prevents these high-frequency boundary artifacts, making it accidentally more robust for pure volumetric segmentation tasks, provided the threshold is sufficiently permissive.

## 4.3 Clinical Implications

These findings have immediate implications for the integration of AI-based super-resolution in clinical workflows.

1. **For Visualization:** SIREN-based reconstruction is preferable. Its ability to denoise the global volume and sharpen features improves the perceived image quality, potentially aiding radiologists in qualitative assessment and incidental finding detection outside the nodule.



**Figure 5: Qualitative comparison of segmentation masks derived from density reconstruction.**

Visual results are presented for three representative nodules: a small isolated nodule (Top; LIDC-IDRI-0066\_0), a nodule adjacent to high-density structures (Middle; LIDC-IDRI-0066\_1), and a juxtapleural nodule (Bottom; LIDC-IDRI-0445\_1). Columns display the volumetric density, the central ground truth mask, the CT slice of the expanded Region of Interest (ROI), and the predicted masks for Cubic Interpolation and SIREN, respectively. The displayed slices correspond to the central intermediate planes (unseen test data) situated between the sparse training slices (1-in-5 protocol). Note the smoothing of the SIREN implicit representation compared to the discretization steps in interpolation.

2. **For Radiomics and Quantification:** Caution is warranted. The variance expansion and spectral alteration introduced by SIREN within the tumor core can significantly skew radiomic feature extraction (e.g., texture entropy, gray-level run length), which are known to be highly sensitive to variations in reconstruction kernels and intensity distributions [11]. For quantitative follow-up, current results suggest that linear interpolation—despite its blurring effect—remains the “safer” estimator because its errors are systematic (smoothing) rather than stochastic (hallucination).

#### 4.4 Limitations and Future Work

This study focused on a 1-in-5 sparsity protocol to simulate aggressive dose reduction; results may vary under less distinct downsampling regimes. Additionally, the vanilla SIREN architecture is utilized to isolate the impact of periodic activations. Recent advancements in positional encoding, such as Hash-Grid encodings (Instant-NGP) [12] or hybrid grid-implicit representations [13], may mitigate the local-global trade-off by permitting spatially adaptive frequency responses. Finally, this study relied on the LIDC-IDRI dataset; external validation on multi-center data with varying scanner protocols is necessary to confirm the generalizability of these spectral biases.

## **5 Conclusion**

This work provides the first quantitative evidence that implicit neural representations in medical imaging are subject to a distinct local-global performance trade-off. While SIRENs represent a powerful tool for continuous anatomical modeling and global image enhancement, they currently lack the constraints necessary to accurately recover the stochastic texture of malignant lesions from sparse data. Future development must move beyond global spectral priors toward spatially-aware architectures able to distinguish between smooth functional continuity of anatomy and high-frequency irregularity of pathology.

## References

- [1] Rebecca L Siegel, Kimberly D Miller, Nikita Sandeep Wagle, and Ahmedin Jemal. Cancer statistics, 2023. *CA: a cancer journal for clinicians*, 73(1), 2023.
- [2] Elizabeth A Eisenhauer, Patrick Therasse, Jan Bogaerts, Lawrence H Schwartz, Danielle Sargent, Robert Ford, Janet Dancey, Stephen Arbuck, Steve Gwyther, Margaret Mooney, et al. New response evaluation criteria in solid tumours: revised recist guideline (version 1.1). *European journal of cancer*, 45(2):228–247, 2009.
- [3] Philippe Lambin, Emmanuel Rios-Velazquez, Ralph Leijenaar, Sara Carvalho, Ruud GPM Van Stiphout, Patrick Granton, Catharina ML Zegers, Robert Gillies, Ronald Boellard, André Dekker, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *European journal of cancer*, 48(4):441–446, 2012.
- [4] Erik HW Meijering, Wiro J Niessen, and Max A Viergever. Quantitative evaluation of convolution-based methods for medical image interpolation. *Medical image analysis*, 5(2):111–126, 2001.
- [5] Hugo JWL Aerts, Emmanuel Rios Velazquez, Ralph TH Leijenaar, Chintan Parmar, Patrick Grossmann, Sara Carvalho, Johan Bussink, René Monshouwer, Benjamin Haibe-Kains, Derek Rietveld, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature communications*, 5(1):4006, 2014.
- [6] Hayit Greenspan, Bram Van Ginneken, and Ronald M Summers. Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE transactions on medical imaging*, 35(5):1153–1159, 2016.
- [7] Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. *Advances in neural information processing systems*, 33:7462–7473, 2020.
- [8] Samuel G Armato III, Geoffrey McLennan, Luc Bidaut, Michael F McNitt-Gray, Charles R Meyer, Anthony P Reeves, Binsheng Zhao, Denise R Aberle, Claudia I Henschke, Eric A Hoffman, et al. The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans. *Medical physics*, 38(2):915–931, 2011.
- [9] Matthew Tancik, Pratul Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan Barron, and Ren Ng. Fourier features let networks learn high frequency functions in low dimensional domains. *Advances in neural information processing systems*, 33:7537–7547, 2020.
- [10] Vegard Antun, Francesco Renna, Clarice Poon, Ben Adcock, and Anders C Hansen. On instabilities of deep learning in image reconstruction and the potential costs of ai. *Proceedings of the National Academy of Sciences*, 117(48):30088–30095, 2020.
- [11] Dennis Mackin, Xenia Fave, Lifei Zhang, David Fried, Jinzhong Yang, Brian Taylor, Edgardo Rodriguez-Rivera, Cristina Dodge, Aaron Kyle Jones, et al. Measuring computed tomography scanner variability of radiomics features. *Investigative radiology*, 50(11):757–765, 2015.

- [12] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM transactions on graphics (TOG)*, 41(4):1–15, 2022.
- [13] Julien NP Martel, David B Lindell, Connor Z Lin, Eric R Chan, Marco Monteiro, and Gordon Wetzstein. Acorn: Adaptive coordinate networks for neural scene representation. *arXiv preprint arXiv:2105.02788*, 2021.