






# Superpixel-informed Implicit Neural Representation for Multi-Dimensional Data

Jiayi Li<sup>1</sup>, Xile Zhao<sup>1</sup>, Jianli Wang<sup>2</sup>, Chao Wang<sup>3</sup>, and Min Wang<sup>4</sup>

<sup>1</sup> University of Electronic Science and Technology of China, Chengdu, China  
lijiayi03531@gmail.com, xlzhao122003@163.com

<sup>2</sup> Southwest Jiaotong University, Chengdu, China  
wangjianli\_123@163.com

<sup>3</sup> Southern University of Science and Technology, Shenzhen, China  
chaowang.hk@gmail.com

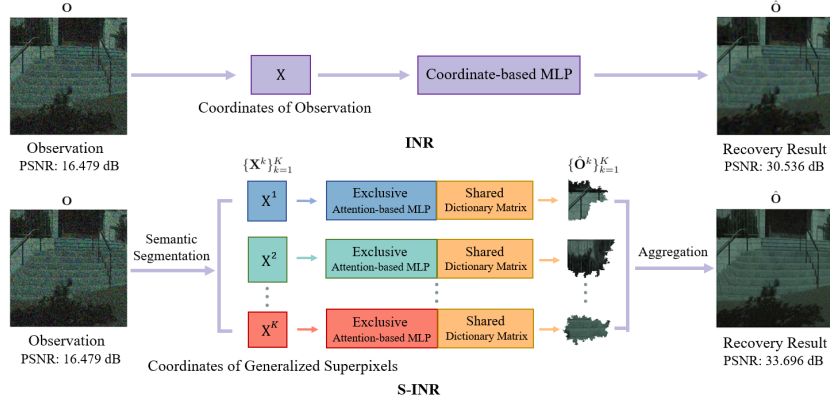
<sup>4</sup> Jiangxi University of Finance and Economics, Nanchang, China  
minwang1989@126.com

**Abstract.** Recently, implicit neural representations (INRs) have attracted increasing attention for multi-dimensional data recovery. However, INRs simply map coordinates via a multi-layer perceptron (MLP) to corresponding values, ignoring the inherent semantic information of the data. To leverage semantic priors from the data, we propose a novel Superpixel-informed INR (S-INR). Specifically, we suggest utilizing generalized superpixel instead of pixel as an alternative basic unit of INR for multi-dimensional data (e.g., images and weather data). The coordinates of generalized superpixels are first fed into exclusive attention-based MLPs, and then the intermediate results interact with a shared dictionary matrix. The elaborately designed modules in S-INR allow us to ingeniously exploit the semantic information within and across generalized superpixels. Extensive experiments on various applications validate the effectiveness and efficacy of our S-INR compared to state-of-the-art INR methods.

**Keywords:** Multi-Dimensional Data · Implicit Neural Representation · Superpixel

## 1 Introduction

Implicit neural representations (INRs) [39] have attracted significant research interest in recent years as a novel method for representing multi-dimensional data, such as images [21, 44, 48], point clouds [3, 9], videos [5, 24, 53], and audio signals [20, 40]. INR utilizes a coordinate-based multilayer perceptron (MLP), which takes spatial coordinates as input and outputs corresponding values. Essentially, INR implicitly represents data via deep neural networks to achieve a continuous representation. The strong representation capacity of INR opens up new possibilities for multi-dimensional data modeling and analysis, overcoming the limitations of traditional discrete representations that rely on fixed mesh-grids or matrices.



**Fig. 1:** Illustration of the overall processing of the traditional INR and our proposed Superpixel-informed INR (S-INR) on image denoising task. In the proposed S-INR, generalized superpixels are used as basic units. The coordinates of generalized superpixels are fed into the exclusive attention-based MLPs, and then the intermediate results interact with a shared dictionary matrix to obtain recovery results.

INR has achieved success in a variety of applications on multi-dimensional data recovery tasks with its strong representation capacity, including image super-resolution [6, 25, 52], completion [47, 49], generation [33, 41], and enhancement [51], 3D shape representation [10, 29, 46], novel view synthesis [18, 19, 32], and physics-based simulation and inference [12, 50]. Recently, INR, i.e., coordinate-based MLP, has been shown to perform poorly in encoding signals with high-frequency components, due to the corresponding neural tangent kernel (NTK) of MLPs being prone to high-frequency fall-offs [17, 30]. To address the above issue, researchers have focused on improving the expressiveness of the INR via two primary aspects: positional encoding and nonlinear activation functions. For positional encoding, the prevalent methods used Fourier features to project low-dimensional inputs into higher-dimensional space by applying a positional encoding layer before the MLP [28, 42]. Later, Long et al. [26] proposed a phase-varying positional encoding module that exploits the relationship between phase information in sinusoidal functions and their displacements. For nonlinear activation functions, Sitzmann et al. used a periodic activation function, Ramasinghe et al. [31] proposed a unified framework for activation, Saragadam et al. [35] used a continuous Gabor wavelet activation function, Shen et al. proposed [38] a novel function TRIDENT for INR characterized by a trilogy of nonlinearities, and Danzel et al. [37] introduced the hyperbolic oscillation function, which is designed to capture high-frequency information of the underlying data. Other improvement methods such as band-limited coordinate networks [23] and multiplicative filter networks [8] have also been proven effective.

Although INR and its variants have achieved success in various applications [4, 13, 27, 34, 36], previous INR methods ignore the rich inherent semantic information of the data, which could be beneficial for data representations and

subsequent applications. A natural question followed: can we develop a novel data representation that efficiently exploits the inherent semantic information of the data under the INR framework?

To address this challenge, we first suggest utilizing generalized superpixels instead of pixels as alternative basic units of INR for multi-dimensional data, which encode rich semantic information. The generalized superpixels, going beyond traditional superpixels, are not limited to image data, but can also be applied to point data arising from real-world applications in general. Then, to fully exploit the semantic information within and across generalized superpixels, we propose a novel Superpixel-informed INR (S-INR). Specifically, we elaborately design two key modules in S-INR, i.e., exclusive attention-based MLPs and a shared dictionary matrix. The exclusive attention-based MLP enhances the expressiveness of S-INR in feature dimensions within each generalized superpixel. The shared dictionary matrix allows us to capture correlations between generalized superpixels. Fig. 1 clearly illustrates the superior representation capacity of our proposed S-INR when compared to traditional INR. Our proposed S-INR effectively exploits the semantic information within and across generalized superpixels, resulting in outstanding performance. The contributions of this paper are as follows:

(i) We suggest utilizing generalized superpixels instead of pixels as alternative basic units of INR, which encode rich semantic information. The generalized superpixels are not limited to image data, but also suitable for more general point data arising from real-world applications than traditional superpixels.

(ii) To exploit the semantic information within and across generalized superpixels, we propose a novel superpixel-informed implicit neural representation (termed as S-INR). The key modules in S-INR, i.e., exclusive attention-based MLPs and the shared dictionary matrix, are elaborately designed to respect the individuality of each generalized superpixel and capture the commonalities between them.

(iii) Extensive experiments including image reconstruction, image completion, image denoising, weather data completion, and 3D surface completion, validate the broad applicability and superiority of S-INR compared to state-of-the-art INR methods.

## 2 The Proposed Model

**Notation.**  $x$ ,  $\mathbf{x}$ , and  $\mathbf{X}$  denote the scalar, vector, and matrix, respectively. Here,  $\|\cdot\|$  denotes the vector norm.

### 2.1 Definition of Generalized Superpixel

Pixels<sup>5</sup> are considered as the basic unit of INR for multi-dimensional data. However, existing INRs, utilizing pixels as the basic unit, do not effectively exploit

---

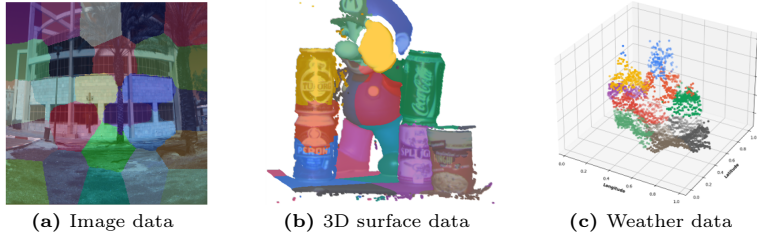
<sup>5</sup> For point data, pixels correspond to a feature value vector.

the inherent semantic information of the data. To address this problem, we propose a generalized superpixel and employ it as an alternative basic unit to enable the exploitation of the inherent semantic information of multi-dimensional data as priors under the INR framework.

**Definition 1 (Generalized Superpixel).** Given a point data  $\mathbf{O} \in \mathbb{R}^{s \times n}$  consisting of  $n$  points, i.e.,  $\{\mathbf{o}_i \in \mathbb{R}^s\}_{i=1}^n$ , where  $s$  is the number of feature dimensions, and the number of generalized superpixels  $K$ .  $\{\mathbf{O}^k = \{\mathbf{o}_i^k\}_{i=1}^{n_k}\}_{k=1}^K$ ,  $n_1 + \dots + n_K = n$ , representing assigned groups of data points are called generalized superpixels if and only if they satisfy the following two conditions:

1. **Disjointness.**  $\{\mathbf{O}^k\}_{k=1}^K$  are pairwise disjoint, i.e.,  $\mathbf{O}^k \cap \mathbf{O}^{k'} = \emptyset$  for all  $k \neq k'$ .
2. **Spatial Connectivity.** Each  $\{\mathbf{O}^k\}_{k=1}^K$  forms a spatially connected blob, meaning that all data points within a generalized superpixel are contiguous or proximate in the spatial domain.

*Remark 1.* The traditional superpixel is limited to image data. In contrast, our proposed generalized superpixel is beyond image data. It is also suitable for more general point data arising from real-world applications, such as 3D surface data and weather data, as illustrated in Figs. 2b and 2c.



**Fig. 2:** Illustration of generalized superpixel segmentation on the (a) image data, (b) 3D surface data, and (c) weather data.

Next, we will introduce our method for finding generalized superpixels from point data. General clustering algorithms (e.g.,  $k$ -means++ [2]) can be deployed to find groups. However, they do not satisfy the above condition of generalized superpixels (i.e., spatial connectivity). To ensure the resulting partitioning of generalized superpixel satisfies both the two above conditions, we develop a generalized superpixel segmentation algorithm (GSSA) in Algorithm 1. Specifically, we consider the importance of corresponding coordinates  $\mathbf{X} \in \mathbb{R}^{c \times n}$  (i.e.,  $\{\mathbf{x}_i \in \mathbb{R}^c\}_{i=1}^n$ ) of the point data  $\mathbf{O} \in \mathbb{R}^{s \times n}$  (i.e.,  $\{\mathbf{o}_i \in \mathbb{R}^s\}_{i=1}^n$ ) to satisfy the spatial connectivity. Our example results are shown in Fig. 2.

With GSSA, the generalized superpixels that encode rich semantic information are ready as basic units instead of pixels in INRs.

---

**Algorithm 1** The Generalized Superpixel Segmentation Algorithm (GSSA)
 

---

**Input:** The point data  $\{\mathbf{o}_i \in \mathbb{R}^s\}_{i=1}^n$  and its coordinate  $\{\mathbf{x}_i \in \mathbb{R}^c\}_{i=1}^n$ , weight  $\alpha$ , and the number of generalized superpixels  $K$ .

**Initialization:** Set initial clustering centers  $\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_K$  using the  $k$ -means++ initialization method and  $\{\mathbf{x}_{\boldsymbol{\mu}_k}\}_{k=1}^K$  are their corresponding coordinates.

- 1: **repeat**
- 2:   **for**  $i = 1$  to  $n$  **do**
- 3:      $m_{ik} = \begin{cases} 1 & \text{if } k = \arg \min_k \|\mathbf{o}_i - \boldsymbol{\mu}_k\|^2 + \alpha \|\mathbf{x}_i - \mathbf{x}_{\boldsymbol{\mu}_k}\|^2 \\ 0 & \text{otherwise} \end{cases}$
- 4:   **end for**
- 5:   **for**  $k = 1$  to  $K$  **do**
- 6:      $\boldsymbol{\mu}_k = \frac{\sum_{i=1}^n m_{ik} \mathbf{o}_i}{\sum_{i=1}^n m_{ik}}$
- 7:      $\mathbf{x}_{\boldsymbol{\mu}_k} = \frac{\sum_{i=1}^n m_{ik} \mathbf{x}_i}{\sum_{i=1}^n m_{ik}}$
- 8:   **end for**
- 9: **until** Convergence

**Output:** Generalized superpixels  $\mathbf{O}^k = \{\mathbf{o}_i | m_{ik} = 1, i = 1, \dots, n\}$  and coordinates  $\mathbf{X}^k = \{\mathbf{x}_i | m_{ik} = 1, i = 1, \dots, n\}, k = 1, \dots, K$ .

---

## 2.2 The Proposed S-INR

To fully exploit the semantic information within and across the generalized superpixels, we propose a novel superpixel-informed implicit neural representation with elaborately designed modules, i.e., exclusive attention-based MLPs and a shared dictionary matrix. Next, we first introduce the basic structure of INR and then elaborate on our proposed S-INR.

INR maps input coordinates to corresponding output values parameterized by an MLP, which is typically continuous and differentiable. Specifically, for a given data  $\mathbf{O} \in \mathbb{R}^{s \times n}$  (i.e.,  $\{\mathbf{o}_i \in \mathbb{R}^s\}_{i=1}^n$ ), INR uses the MLP with the sinusoidal activation function to map its coordinate  $\mathbf{X} \in \mathbb{R}^{c \times n}$  (i.e.,  $\{\mathbf{x}_i \in \mathbb{R}^c\}_{i=1}^n$ ) to corresponding values  $\hat{\mathbf{O}} \in \mathbb{R}^{s \times n}$  (i.e.,  $\{\hat{\mathbf{o}}_i \in \mathbb{R}^s\}_{i=1}^n$ ), serving as the recovery results. The INR, mapping  $\mathbb{R}^c$  to  $\mathbb{R}^s$ , can be formulated as follows:

$$\Phi_\theta(\mathbf{x}) = \mathbf{W}_L(\sin(\mathbf{W}_{L-1}(\cdots(\sin(\mathbf{W}_0\mathbf{x} + \mathbf{b}_0))\cdots) + \mathbf{b}_{L-1})) + \mathbf{b}_L, \quad (1)$$

where  $\theta \triangleq \{\mathbf{W}_l, \mathbf{b}_l\}_{l=0}^L$  denotes the learnable parameters,  $\{\mathbf{W}_l \in \mathbb{R}^{c_{l+1} \times c_l}\}_{l=0}^L$  are weight matrices, and  $\{\mathbf{b}_l \in \mathbb{R}^{c_{l+1}}\}_{l=0}^L$  denote bias vectors. Here,  $c_{L+1} = s$  is the output dimension and  $L$  denotes the number of layers in INR.

Different from using pixels as basic units in INR, we suggest utilizing generalized superpixels as the basic units of S-INR, which encode rich inherent semantic information of the data. The details of S-INR are presented below.

First, we propose a standard S-INR to exploit the semantic information within and across generalized superpixels. Specifically, we suggest utilizing the shared dictionary matrix  $\mathbf{D} \in \mathbb{R}^{s \times r}$ , which can interact with each generalized superpixel. With  $\mathbf{D}$ , the mathematical expression of the standard S-INR is as follows:

$$\hat{\mathbf{o}}^k = \mathbf{D}(\Phi_{\theta_k}(\mathbf{x}^k)), \quad k = 1, \dots, K, \quad (2)$$

where our proposed S-INR maps the corresponding coordinate of the  $k$ -th generalized superpixels  $\mathbf{X}^k \in \mathbb{R}^{c \times n_k}$  (i.e.,  $\{\mathbf{x}_i^k \in \mathbb{R}^c\}_{i=1}^{n_k}$ ) to corresponding values  $\hat{\mathbf{O}}^k \in \mathbb{R}^{s \times n_k}$  (i.e.,  $\{\hat{\mathbf{o}}_i^k \in \mathbb{R}^s\}_{i=1}^{n_k}$ ), which is then aggregated to obtain recovery results  $\hat{\mathbf{O}} \in \mathbb{R}^{s \times n}$  (i.e.,  $\{\hat{\mathbf{o}}_i \in \mathbb{R}^s\}_{i=1}^n$ ),  $n_1 + \dots + n_K = n$ . Here,  $\Phi_{\theta_k} : \mathbb{R}^c \rightarrow \mathbb{R}^r$  denotes the  $k$ -th INR representing the  $k$ -th generalized superpixel with learnable parameters  $\theta_k \triangleq \{\mathbf{W}_l^k, \mathbf{b}_l^k\}_{l=0}^L$  and  $K$  denotes the number of generalized superpixels. Note that in our work we set the dictionary matrix  $\mathbf{D}$  to be a learnable parameter matrix, which allows for more flexible representation learning compared to using a predefined coding matrix in the dictionary learning field.

Second, we develop an advanced S-INR that significantly enhances the expressiveness of S-INR in representing each generalized superpixel. Concretely, we propose an exclusive attention-based MLP  $\Psi_{\theta_k}$ , which is tailored to capture correlations across feature dimensions within each generalized superpixel. With  $\Psi_{\theta_k}$ , the formal formulation of S-INR is formulated as follows:

$$\hat{\mathbf{o}}^k = \mathbf{D}(\Psi_{\theta_k}(\mathbf{x}^k)), \quad k = 1, \dots, K. \quad (3)$$

Specifically, for the proposed exclusive attention-based MLP  $\Psi_{\theta_k}$ , we add a self-attention function  $\psi_l^k : \mathbb{R}^{c_{l+1}} \rightarrow \mathbb{R}^{c_{l+1}}$  [14] after the  $l$ -th layer in the  $k$ -th INR  $\Phi_{\theta_k}$ , so that the output of the original  $l$ -th layer is further processed by  $\psi_l^k$ . Thus, the parameters of the advanced S-INR can be denoted as  $\{\mathbf{D}, \{\theta_k\}_{k=1}^K\}$ , where  $\theta_k \triangleq \{\mathbf{U}_l^k, \mathbf{V}_l^k, \mathbf{W}_l^k, \mathbf{b}_l^k\}_{l=0}^L$  denotes the learnable parameters. The mathematical expression of  $\Psi_{\theta_k}$  is formulated as follows:

$$\begin{aligned} \Psi_{\theta_k}(\mathbf{x}^k) = & \psi_L^k(\mathbf{W}_L^k(\psi_{L-1}^k(\sin(\mathbf{W}_{L-1}^k(\dots \psi_0^k(\sin(\mathbf{W}_0^k \mathbf{x}^k + \mathbf{b}_0^k)) \dots) \\ & + \mathbf{b}_{L-1}^k))) + \mathbf{b}_L^k), \end{aligned} \quad (4)$$

where  $\psi_l^k(\mathbf{z}_{l+1}^k) = \eta(\mathbf{U}_l^k(\delta(\mathbf{V}_l^k(\tau(\mathbf{z}_{l+1}^k)))) \otimes \mathbf{z}_{l+1}^k)$ ,  $l = 0, \dots, L$ . Here, if  $l = 0$ ,  $\mathbf{z}_1^k = \sin(\mathbf{W}_0^k \mathbf{x}^k + \mathbf{b}_0^k)$ , if  $l = 1, \dots, L-1$ ,  $\mathbf{z}_{l+1}^k = \sin(\mathbf{W}_l^k \mathbf{z}_l^k + \mathbf{b}_l^k)$ , and if  $l = L$ ,  $\mathbf{z}_{l+1}^k = \mathbf{W}_l^k \mathbf{z}_l^k + \mathbf{b}_l^k$ . The symbol  $\tau(\cdot)$  denotes channel-wise average pooling,  $\delta(\cdot)$  denotes the ReLU activation function,  $\eta(\cdot)$  denotes the sigmoid function,  $\otimes$  denotes the channel-wise product, and  $\{\mathbf{U}_l^k \in \mathbb{R}^{c_{l+1} \times c_{l+1}}\}_{l=0}^L$  and  $\{\mathbf{V}_l^k \in \mathbb{R}^{c_{l+1} \times c_{l+1}}\}_{l=0}^L$  are learnable parameter matrices.

In summary, our proposed S-INR fully exploits the semantic information within and across generalized superpixels. The elaborately designed S-INR successfully as an effective tool exploits the inherent semantic information of the data under the INR framework.

### 2.3 The Proposed Recovery Model

To examine the representation capacity of our proposed S-INR, we propose an S-INR-based data recovery model tailored for multi-dimensional data recovery tasks. This model can be formulated as follows:

$$\min_{\mathbf{D}, \{\theta_k\}_{k=1}^K} \mathcal{L}(\mathbf{O}, \hat{\mathbf{O}}) = \sum_{k=1}^K \sum_{i=1}^{n_k} \mathcal{L}(\mathbf{o}_i^k, \mathbf{D}(\Psi_{\theta_k}(\mathbf{x}_i^k))), \quad (5)$$

where  $\Psi_{\theta_k}(\cdot)$  denotes the exclusive self-attention MLP to fully exploit semantic information within corresponding generalized superpixel,  $\mathbf{D}$  denotes the shared dictionary matrix that interacts with each generalized superpixel to capture commonalities between them, the generalized superpixel  $\mathbf{O}^k \in \mathbb{R}^{s \times n_k}$  (i.e.,  $\{\mathbf{o}_i^k \in \mathbb{R}^s\}_{i=1}^{n_k}$ ) is obtained from the observation  $\mathbf{O} \in \mathbb{R}^{s \times n}$  (i.e.,  $\{\mathbf{o}_i \in \mathbb{R}^s\}_{i=1}^n$ ),  $\mathbf{x}_i^k$  is corresponding coordinate of the  $\mathbf{o}_i^k$ ,  $K$  is the number of generalized superpixels, and  $\mathcal{L}$  denotes the loss function. The aggregation of  $\{\hat{\mathbf{O}}^k = \{\mathbf{D}(\Psi_{\theta_k}(\mathbf{x}_i^k))\}_{i=1}^{n_k}\}_{k=1}^K$  is the final recovery result  $\hat{\mathbf{O}}$ . Next, with our proposed S-INR-based data recovery model, we consider three specific data recovery tasks:

1. **Data reconstruction** aims at approximating the ground truth. The loss function is  $\sum_{k=1}^K \sum_{i=1}^{n_k} \|\mathbf{o}_i^k - \mathbf{D}(\Psi_{\theta_k}(\mathbf{x}_i^k))\|^2$ , where  $\mathbf{o}_i^k$  is the  $i$ -th point from  $k$ -th clean generalized superpixel.
2. **Data completion** aims at recovering the underlying patterns and structures from incomplete observations. The loss function is  $\sum_{k=1}^K \sum_{i=1}^{n_k} \|(\mathbf{o}_i^k - \mathbf{D}(\Psi_{\theta_k}(\mathbf{x}_i^k)))_{\Omega}\|^2$ , where  $\mathbf{o}_i^k$  is the  $i$ -th point from  $k$ -th incomplete generalized superpixel and  $\Omega$  is the support of the observed data.
3. **Data denoising** aims at recovering the underlying patterns and structures from noisy corrupted observation. We consider Gaussian noise with different standard deviations. The loss function is  $\sum_{k=1}^K \sum_{i=1}^{n_k} \|\mathbf{o}_i^k - \mathbf{D}(\Psi_{\theta_k}(\mathbf{x}_i^k))\|^2$ , where  $\mathbf{o}_i^k$  is the  $i$ -th point from  $k$ -th noisy corrupted generalized superpixel.

The proposed S-INR-based data recovery model is unsupervised which solely requires the observed data without training dataset. To solve the highly nonconvex and nonlinear problem Eq. (5), we use prevalent gradient descent methods, i.e., the efficient adaptive moment estimation algorithm (Adam) [22], to update the parameters  $\mathbf{D}$  and  $\{\theta_k\}_{k=1}^K$ . Note that in the data completion, we utilize our proposed GSSA to find generalized superpixels from the incomplete observation after interpolation.

### 3 Experiments

In this section, we evaluate the representation capacity of our proposed S-INR on a range of multi-dimensional data recovery tasks. We first evaluate image data for tasks such as image reconstruction, image completion, and image denoising. Then, we extend the evaluation to more general point data arising from real-world applications, including weather data completion and 3D surface completion tasks. All experiments are conducted on a computer equipped with an Intel(R) UHD Graphics 630 CPU and an RTX 2080 Ti GPU.

#### 3.1 Experimental Settings

To evaluate the performance of S-INR in different tasks, we use several evaluation metrics in experiments. For tasks involving image data, we evaluate the results using the peak signal-to-noise ratio (PSNR) [16] and the structural similarity index (SSIM) [45]. For tasks involving point data, we report the normalized root mean square error (NRMSE) and the R-Square for evaluations.

**Image Data.** To comprehensively evaluate the performance of our proposed S-INR on image data, we conduct experiments on three important image processing tasks: image reconstruction, image completion, and image denoising. These tasks allow us to assess different capabilities of S-INR, including reconstructing complex data, completing missing data, and recovering noisy corrupted data. (1) Image Representation: We utilize a RGB image *Kodim* ( $512 \times 768 \times 3$ ) from Kodak dataset<sup>6</sup> and a hyperspectral image (HSI) *Pavia* ( $256 \times 256 \times 32$ ) from dataset<sup>7</sup> as testing data. (2) Image Completion: We use a multispectral image (MSI) *Mor* ( $256 \times 256 \times 31$ ) in the ICVL dataset<sup>8</sup> as testing data. We adopt the random sampling rates of 0.025 and 0.05, labeled as Case1 and Case2, respectively. (3) Image Denoising: We adopt a MSI *Lehavam* ( $256 \times 256 \times 31$ ) also in the ICVL dataset as testing data. We consider Gaussian noise with standard deviations of 0.15 and 0.2, labeled as Case1 and Case2, respectively. We compare our method with state-of-the-art INR methods including Fourier [42] (i.e., ReLU+Pos.Enc), Gauss [31], WIRE [35], and SIREN [39]. The image reconstruction task is also compared with DIP [43].

**Point Data.** We further evaluate S-INR on processing point data tasks: weather data completion and 3D surface completion. These allow us to assess the ability of S-INR to model unstructured point data under real-world conditions. (1) 3D Surface Completion: We consider the 3D surface completion task with a six-dimensional point set  $(x, y, z) - (R, G, B)$  formed by  $n$  points, to estimate the color information of a given point cloud  $\mathbf{O} \in \mathbb{R}^{3 \times n}$  from its corresponding coordinate  $\mathbf{X} \in \mathbb{R}^{3 \times n}$ . We use three scenes, i.e., *Scene1*, *Scene2*, and *Scene3*, acquired with spacetime stereo [7], from the SHOT website<sup>9</sup>. We adopt the random sampling rates of 0.025, 0.05, 0.075, and 0.1 as four cases, respectively. (2) Weather Data Completion: The weather data we utilize includes five values, namely precipitation (Prep), soil water evaporation (SWE), vapor pressure (VP), and extreme temperatures (Tmax and Tmin). We utilize three datasets<sup>10</sup> from the North America located at ( $63^\circ\text{N}$ ,  $157^\circ\text{W}$ ), ( $61^\circ\text{N}$ ,  $141^\circ\text{W}$ ), and ( $62^\circ\text{N}$ ,  $149^\circ\text{W}$ ), respectively. We adopt the random sampling rates of 0.1, 0.15, 0.2, and 0.25 as four cases, respectively. For the above two tasks, we report the average quantitative results for their corresponding four cases, respectively. We compare our method with standard regression methods including  $K$ -neighbors regressor (KNN) [1], decision tree (DT) [15], and random forest (RF) [11]. Additionally, we compare our method with the classical INR method, i.e., SIREN [39].

**Hyperparameters Settings.** The configuration of hyperparameters is crucial for both our proposed S-INR method and other comparison methods. For WIRE,

<sup>6</sup> <https://r0k.us/graphics/kodak/>

<sup>7</sup> [https://www.ehu.eus/ccwintco/index.php/Hyperspectral\\_Remote\\_Sensing\\_Scenes](https://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes)

<sup>8</sup> <https://icvl.cs.bgu.ac.il/hyperspectral/>

<sup>9</sup> <http://www.vision.deis.unibo.it/research/80-shot>

<sup>10</sup> <https://daac.ornl.gov/cgi-bin/dsviewer.pl?dsid=2130>



as in the original paper, we set the size of the hidden layer to 300. For other INR methods, we set the size of the hidden layer to 256. Regarding the number of hidden layers, we set it to 5 for all INR methods. To ensure a fair comparison, we apply other hyperparameters from the original papers when available. Otherwise, we select optimal values within the ranges stated in the papers. For DIP, we set its hyperparameters as in the original paper. To be fair, we reduce the size of the hidden layer to keep its parameter size similar to other INR methods.

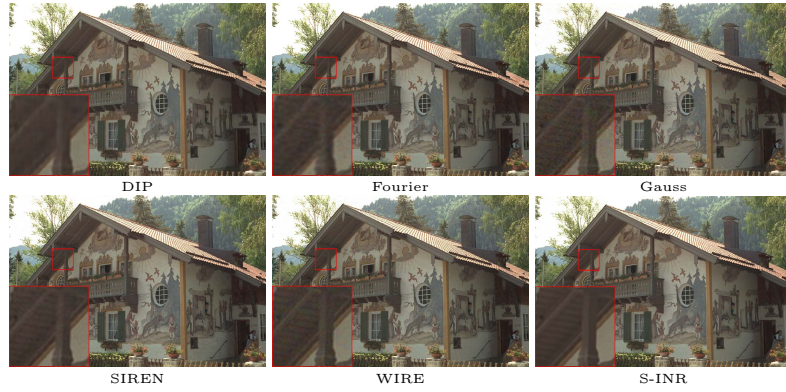
For our S-INR, we set the size of the hidden layer to 35 and the number of hidden layers to 5. Other important hyperparameters include the number of superpixels  $K$ , the weight  $\alpha$  of GSSA, the initialization parameter  $\omega_0$ , the learning rate, and the size of the dictionary matrix  $r$ . The number of superpixels  $K$ , depending on the scale of the input data, is chosen from  $\{15, 25, 50\}$ . The weight  $\alpha$  is selected from the set  $\{1, 5, 20\}$ , with the specific choice dependent on the data type. The initialization parameter  $\omega_0$ , uses the same weight initialization scheme for  $\omega_0$  as SIREN [39], but the tunable value of  $\omega_0$  is selected from  $\{30, 150, 300\}$ , which is influenced by  $K$ . The learning rate is searched within  $\{1 \times 10^{-5}, 5 \times 10^{-5}, 1 \times 10^{-4}, 5 \times 10^{-4}, 1 \times 10^{-3}\}$ . The size of the dictionary matrix  $r$  is set to a multiple of  $s$ , where  $s$  is the output dimension of the data, and the scale factor is set to 5 for RGB images, 4 for point data, and  $\{2, 3, 4\}$  for HSIs.

### 3.2 Experimental Results

**Image Reconstruction Results.** The quantitative and qualitative results of image reconstruction are illustrated in Table 1 and Fig. 3. We can observe that S-INR achieves superior PSNR and SSIM values compared to other comparison methods in Table 1. Meanwhile, we can observe that DIP, SIREN, and Fourier fail to represent image details, and the results of Gauss and WIRE are difficult to remove noise. In contrast, our method preserves rich textures and details without introducing noise, as shown in Fig. 3. The promising reconstruction performance of our method can be largely attributed to the inherent semantic information captured by using the proposed generalized superpixels as basic units.

**Table 1:** The quantitative results for image reconstruction task. The **best** and second best values are highlighted. (PSNR  $\uparrow$ , SSIM  $\uparrow$ )

Data	Metrics	DIP	Fourier	Gauss	SIREN	WIRE	S-INR
<i>Kodim</i>	PSNR	30.154	32.101	30.188	33.052	<u>33.199</u>	<b>36.077</b>
	SSIM	0.882	0.899	0.862	<u>0.932</u>	0.918	<b>0.965</b>
<i>Pavia</i>	PSNR	36.283	37.982	36.413	37.727	<u>38.455</u>	<b>39.102</b>
	SSIM	0.919	0.935	0.923	0.937	<u>0.941</u>	<b>0.949</b>



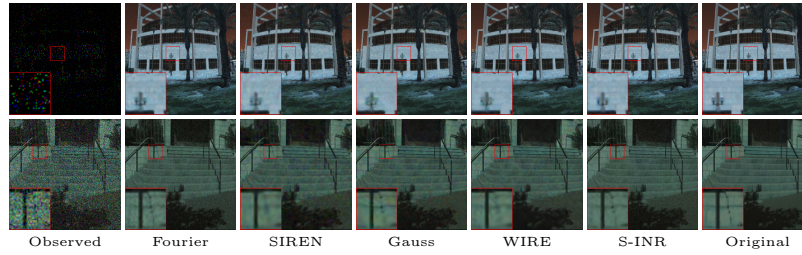
**Fig. 3:** The results of image reconstruction by different methods on RGB image *Kodim*.

**Image Completion Results.** The quantitative and qualitative results of image completion are shown in Table 2 and the first row of Fig. 4. We can observe that our method obtains better quantitative results as compared with other INR methods in terms of PSNR and SSIM. From the visual results in Fig. 4, we can see that Fourier and SIREN fail to represent image details and produce blurry results, Gauss and WIRE show artifacts in the form of green discoloration and extraneous patterns. Meanwhile, S-INR can capture fine details without artifacts, validating the effectiveness of our method in image completion tasks. The superior performance of S-INR is attributed to its elaborately designed components to jointly exploit the inherent semantic information of the data.

**Table 2:** The quantitative results for image completion task. The **best** and second best values are highlighted. (PSNR  $\uparrow$ , SSIM  $\uparrow$ )

Data	Cases	Metrics	Observed	Fourier	Gauss	SIREN	WIRE	S-INR
<i>Mor</i>	Case1	PSNR	9.056	27.835	28.064	27.918	<u>28.087</u>	<b>29.068</b>
		SSIM	0.030	0.801	<u>0.808</u>	0.763	0.804	<b>0.823</b>
	Case2	PSNR	9.169	31.017	31.026	30.929	<u>31.609</u>	<b>32.281</b>
		SSIM	0.047	0.886	0.895	0.860	<u>0.897</u>	<b>0.900</b>

**Image Denoising Results.** The qualitative and quantitative results of image denoising are shown in the second row of Fig. 4 and Table 3. The visual quality of our method is apparent, particularly in its ability to represent intricate details and remove heavy noise, such as the split on the wall reflected in the zoomed views, as shown in Fig. 4. On the contrary, Gauss, SIREN, and WIRE



**Fig. 4:** From top to bottom list the results of image completion and image denoising by different methods on *Mor* in Case2 and *Lehavam* in Case1, respectively.

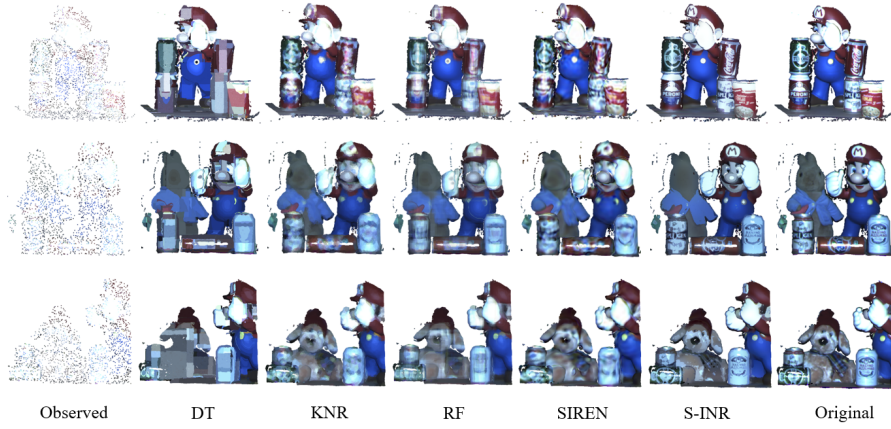
**Table 3:** The quantitative results for image denoising task. The **best** and second best values are highlighted. (PSNR  $\uparrow$ , SSIM  $\uparrow$ )

Data	Cases	Metrics	Observed	Fourier	Gauss	SIREN	WIRE	S-INR
<i>Lehavam</i>	Case1	PSNR	16.479	32.514	32.853	30.536	<u>33.196</u>	<b>33.696</b>
		SSIM	0.224	0.832	0.856	0.784	<u>0.863</u>	<b>0.867</b>
	Case2	PSNR	13.977	<u>31.333</u>	31.303	29.348	30.123	<b>32.394</b>
		SSIM	0.144	0.796	<u>0.819</u>	0.745	0.792	<b>0.831</b>

fail to remove all the noise, while Fourier has chromatic aberration and introduces background noise. In addition, as compared with other INR methods, our method achieves competitive PSNR and SSIM values in Table 3. This indicates that S-INR is not only superior in the accuracy of visual quality but also in the quantitative metrics compared to other comparison methods. We attribute the ability to recover noisy images even without any regularization to the expressiveness and robustness of S-INR. It is achieved by respecting the individuality of each generalized superpixel and capturing the commonalities between them.

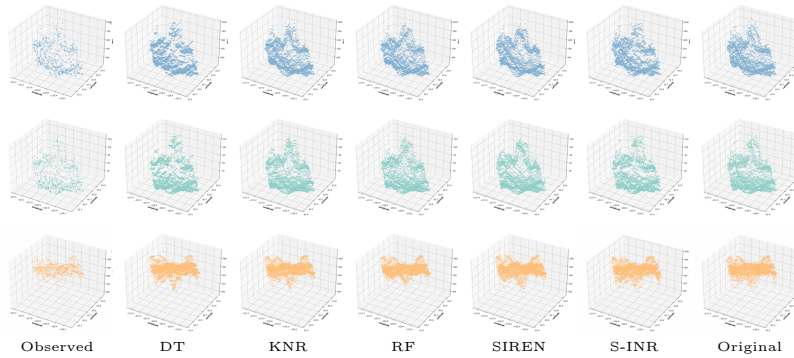
**Table 4:** The average quantitative results by different methods for point data recovery tasks. The **best** and second best values are highlighted. (NRMSE  $\downarrow$ , R-Square  $\uparrow$ )

Data	Weather Data Completion						3D Surface Completion					
	(63°N, 157°W)		(61°N, 141°W)		(62°N, 149°W)		<i>Scene1</i>		<i>Scene2</i>		<i>Scene3</i>	
Method	NRMSE	R-Square	NRMSE	R-Square	NRMSE	R-Square	NRMSE	R-Square	NRMSE	R-Square	NRMSE	R-Square
DT	0.101	0.698	0.149	0.484	0.131	0.486	0.171	0.703	0.126	0.761	0.141	0.721
KNR	<u>0.072</u>	<u>0.849</u>	0.099	0.766	<u>0.093</u>	0.739	0.112	0.868	0.093	0.867	0.094	0.874
RF	0.076	0.829	0.111	0.712	0.099	0.705	<u>0.107</u>	<u>0.880</u>	<u>0.090</u>	0.875	0.092	0.879
SIREN	0.078	0.818	<u>0.096</u>	<u>0.776</u>	<u>0.093</u>	<u>0.741</u>	0.109	0.878	<u>0.090</u>	<u>0.886</u>	<u>0.087</u>	<u>0.885</u>
S-INR	<b>0.058</b>	<b>0.900</b>	<b>0.070</b>	<b>0.877</b>	<b>0.062</b>	<b>0.889</b>	<b>0.074</b>	<b>0.944</b>	<b>0.063</b>	<b>0.945</b>	<b>0.060</b>	<b>0.945</b>



**Fig. 5:** From top to down respectively list the results of 3D surface completion recovery by different methods for the 3D surface data *Scene1*, *Scene2*, and *Scene3* with the random sampling rate of 0.025.

**3D Surface Completion Results.** The quantitative and qualitative results of the 3D surface completion are shown in Table 4 and Fig. 5. We can observe that our S-INR achieves higher performance than other comparison methods in terms of NRMSE and R-Square. From the visual results, we can see that SIREN and other regression methods have difficulty in accurately recovering intricate details. In contrast, the proposed method can better represent the details due to its ability to exploit the inherent semantic information of the data, as evidenced by the patterns of the bottle in Fig. 5. These results confirm the effectiveness and superiority of our S-INR for complex point data recovery tasks beyond images.

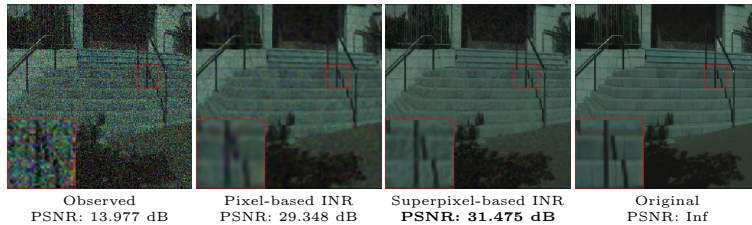


**Fig. 6:** From top to down respectively list the results of weather data completion by different methods on Prcp, SWE, and VP for the weather data located at ( $63^{\circ}\text{N}$ ,  $157^{\circ}\text{W}$ ) with the random sampling rate of 0.25.

**Weather Date Completion Results.** The results on weather data completion are shown in Table 4 and Fig. 6. Compared with other INR methods, our proposed S-INR achieves superior performance in quantitative results in terms of NRMSE and R-square in Table 4. Visually, we can observe that the results of the proposed method are generally closest to the true values in Fig. 6. In contrast, other regression methods struggle to recover the complex structure of weather data, resulting in their results being quite different from the original realistic results. The superior performance of our method can be attributed to the elaborately designed components in S-INR to exploit the semantic information within and across generalized superpixels.

### 3.3 Discussions

**The Role of Generalized Superpixels.** To better understand the roles of basic units, we compare two categories of basic units (i.e., pixels and generalized superpixels) based INR, namely the traditional pixel-based INR [39] and the superpixel-based INR. For fairness, we adopt the same INR architecture<sup>11</sup> that uses MLP with the sinusoidal activation function to map the coordinates to corresponding values. Both model parameter scales are kept on the same scale by adjusting certain model hyperparameters.



**Fig. 7:** The results of image denoising by pixel-based INR [39] and superpixel-based INR on MSI *Lehavim* in Case1. When utilizing our proposed generalized superpixels as basic units, INR can restore finer details and remove more noise by exploiting the inherent semantic information of the data. The **best** values are highlighted.

We conduct quantitative and qualitative comparisons between the pixel-based INR and the superpixel-based INR, with results presented in Fig. 7. We can observe that superpixel-based INR performs better than pixel-based INR as evidenced by higher PSNR. Visually, we can observe that superpixel-based INR preserves finer details with fewer artifacts than pixel-based INR, e.g., we can find in the zoomed region that superpixel-based INR can represent more

<sup>11</sup> Specifically, the superpixel-based INR is based on using  $K$  INRs to represent  $K$  superpixels, where  $K$  is the number of generalized superpixels.

**Table 5:** Quantitative results of image denoising on *Lehavam* in Case2. The **best** values are highlighted.

Components			<i>Lehavam</i>		Parameters (Mb)
Generalized Superpixels	Exclusive MLPs	Shared Dictionary Matrix	PSNR	SSIM	
<del>X</del>	<del>X</del>	<del>X</del>	29.348	0.745	0.786
✓	<del>X</del>	<del>X</del>	31.475	0.817	<b>0.741</b>
✓	<del>X</del>	✓	31.997	0.818	0.797
✓	✓	<del>X</del>	31.012	0.787	0.902
✓	✓	✓	<b>32.394</b>	<b>0.831</b>	0.995

intricate details, while its background contains less noise. The superior performance of superpixel-based INR is attributed to utilizing our proposed basic unit, namely generalized superpixel, which encode rich inherent semantic information of the data.

**Ablation Experiment.** To understand the roles of S-INR’s key components, we perform an ablation experiment on the image denoising task. The important components include the use of generalized superpixels as basic units (generalized superpixels), exclusive attention-based MLPs (exclusive MLPs), and a shared dictionary matrix.

Table 5 shows the recovery results of quantitative comparisons with different components in S-INR. We can observe that the three components in S-INR all contribute to the superior performance of S-INR. The results of S-INR without the generalized superpixels as basic units, exclusive attention-based MLPs, and the shared dictionary matrix show the performance degradation of approximately 3.0dB, 0.4dB, and 1.4dB in terms of PSNR, respectively, compared to the performance of our proposed S-INR. This demonstrates that all three components are important to S-INR, and by working together, they can effectively exploit the inherent semantic information of the data to achieve state-of-the-art performance.

## 4 Conclusion

In this work, we first suggested utilizing generalized superpixels instead of pixels as alternative basic units of INR for multi-dimensional data that encode rich semantic information. Then, to fully exploit the semantic information within and across generalized superpixels, we proposed a novel superpixel-informed implicit neural representation (termed as S-INR), which includes two elaborately designed modules, namely the exclusive attention-based MLPs and a shared dictionary matrix. Extensive experiments have validated consistent improvements in our proposed S-INR over state-of-the-art INR methods in terms of quantitative metrics and visual quality. We believe that S-INR is a potential tool for more versatile multi-dimensional applications.

**Acknowledgments:** This research is supported by NSFC (No. 12371456, 12171-072, 62131005, 62306248), Sichuan Science and Technology Program (No. 2024NS-FJQ0038, 2023ZYD0007), and National Key Research and Development Program of China (No. 2020YFA0714001).

## References

1. Altman, N.S.: An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician* **46**(3), 175–185 (1992)
2. Arthur, D., Vassilvitskii, S.: k-means++: the advantages of careful seeding. In: *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. pp. 1027–1035 (2007)
3. Ben-Shabat, Y., Koneputugodage, C.H., Gould, S.: Digs: Divergence guided shape implicit neural representation for unoriented point clouds. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 19323–19332 (June 2022)
4. Cervantes, P., Sekikawa, Y., Sato, I., Shinoda, K.: Implicit neural representations for variable length human motion generation. In: *Proceedings of the European Conference Computer Vision (ECCV)*. pp. 356–372 (2022)
5. Chen, H., He, B., Wang, H., Ren, Y., Lim, S.N., Shrivastava, A.: Nerv: Neural representations for videos. In: *Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)*. vol. 34, pp. 21557–21568 (2021)
6. Chen, Y., Liu, S., Wang, X.: Learning continuous image representation with local implicit image function. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 8628–8638 (June 2021)
7. Davis, J., Nehab, D., Ramamoorthi, R., Rusinkiewicz, S.: Spacetime stereo: a unifying framework for depth from triangulation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27**(2), 296–302 (2005)
8. Fathony, R., Sahu, A.K., Willmott, D., Kolter, J.Z.: Multiplicative filter networks. In: *Proceedings of the International Conference on Learning Representations (ICLR)* (2021)
9. Fujiwara, K., Hashimoto, T.: Neural implicit embedding for point cloud analysis. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 11734–11743 (June 2020)
10. Henzler, P., Mitra, N.J., Ritschel, T.: Learning a neural 3D texture space from 2D exemplars. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 8353–8361 (June 2020)
11. Ho, T.K.: Random decision forests. In: *Proceedings of 3rd International Conference on Document Analysis and Recognition (ICDAR)*. vol. 1, pp. 278–282 (1995)
12. Hofherr, F., Koestler, L., Bernard, F., Cremers, D.: Neural implicit representations for physical parameter inference from a single video. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. pp. 2093–2103 (January 2023)
13. Hong, S., Nam, J., Cho, S., Hong, S., Jeon, S., Min, D., Kim, S.: Neural matching fields: Implicit representation of matching fields for visual correspondence. In: *Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)*. vol. 35, pp. 13512–13526 (2022)
14. Hu, J., Shen, L., Albanie, S., Sun, G., Wu, E.: Squeeze-and-Excitation networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **42**(8), 2011–2023 (August 2020)

15. Hunt, E.B., Marin, J., Stone, P.J.: Experiments in induction. (1966)
16. Huynh-Thu, Q., Ghanbari, M.: Scope of validity of psnr in image/video quality assessment. *Electronics Letters* **44**, 800–801 (02 2008)
17. Jacot, A., Gabriel, F., Hongler, C.: Neural tangent kernel: Convergence and generalization in neural networks. In: *Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)*. vol. 31, p. 6 (2018)
18. Jayasundara, V., Agrawal, A., Heron, N., Shrivastava, A., Davis, L.S.: FlexNeRF: Photorealistic free-viewpoint rendering of moving humans from sparse views. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 21118–21127 (June 2023)
19. Kaneko, T.: MIMO-NeRF: Fast neural rendering with multi-input multi-output neural radiance fields. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. pp. 3273–3283 (October 2023)
20. Kazerooni, A., Azad, R., Hosseini, A., Merhof, D., Bagci, U.: Incode: Implicit neural conditioning with prior knowledge embeddings. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. pp. 1298–1307 (January 2024)
21. Khan, M.O., Fang, Y.: Implicit neural representations for medical imaging segmentation. In: *Proceedings of the Medical Image Computing and Computer Assisted Intervention (MICCAI)*. pp. 433–443 (2022)
22. Kingma, D., Ba, J.: Adam: A method for stochastic optimization. *Proceedings of the International Conference on Learning Representations (ICLR)* p. 13 (May 2015)
23. Lindell, D.B., Van Veen, D., Park, J.J., Wetzstein, G.: BACON: Band-limited coordinate networks for multiscale scene representation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 16252–16262 (June 2022)
24. Lu, Y., Wang, Z., Liu, M., Wang, H., Wang, L.: Learning spatial-temporal implicit neural representations for event-guided video super-resolution. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 1557–1567 (June 2023)
25. Ma, C., Yu, P., Lu, J., Zhou, J.: Recovering realistic details for magnification-arbitrary image super-resolution. *IEEE Transactions on Image Processing* **31**, 3669–3683 (2022)
26. Mai, L., Liu, F.: Motion-adjustable neural implicit video representation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 10738–10747 (June 2022)
27. Martel, J.N.P., Lindell, D.B., Lin, C.Z., Chan, E.R., Monteiro, M., Wetzstein, G.: Acorn: adaptive coordinate networks for neural scene representation. *ACM Transactions on Graphics* **40**(4) (July 2021)
28. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: NeRF: Representing scenes as neural radiance fields for view synthesis. In: *Proceedings of the European Conference Computer Vision (ECCV)*. pp. 405–421 (2020)
29. Oechsle, M., Mescheder, L., Niemeyer, M., Strauss, T., Geiger, A.: Texture fields: Learning texture representations in function space. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. pp. 4531–4540 (October 2019)
30. Rahaman, N., Baratin, A., Arpit, D., Draxler, F., Lin, M., Hamprecht, F., Bengio, Y., Courville, A.: On the spectral bias of neural networks. In: *Proceedings of the*



- 36th International Conference on Machine Learning (ICML). vol. 97, pp. 5301–5310 (June 2019)
31. Ramasinghe, S., Lucey, S.: Beyond Periodicity: Towards a unifying framework for activations in coordinate-mlps. In: *Proceedings of the European Conference Computer Vision (ECCV)*. pp. 142–158 (2022)
32. Rebain, D., Jiang, W., Yazdani, S., Li, K., Yi, K.M., Tagliasacchi, A.: DeRF: Decomposed radiance fields. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 14153–14161 (June 2021)
33. Reed, A.W., Kim, H., Anirudh, R., Mohan, K.A., Champley, K., Kang, J., Jayasuriya, S.: Dynamic CT reconstruction from limited views with implicit neural representations and parametric motion fields. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. pp. 2258–2268 (October 2021)
34. Reiser, C., Peng, S., Liao, Y., Geiger, A.: KiloNeRF: Speeding up neural radiance fields with thousands of tiny mlps. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. pp. 14335–14345 (October 2021)
35. Saragadam, V., LeJeune, D., Tan, J., Balakrishnan, G., Veeraraghavan, A., Baraniuk, R.G.: WIRE: Wavelet implicit neural representations. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 18507–18516 (June 2023)
36. Saragadam, V., Tan, J., Balakrishnan, G., Baraniuk, R.G., Veeraraghavan, A.: MINER: Multiscale implicit neural representations. In: *Proceedings of the European Conference Computer Vision (ECCV)*. pp. 318–333 (2022)
37. Serrano, D., Szymkowiak, J., Musialski, P.: Hosc: A periodic activation function for preserving sharp features in implicit neural representations (2024)
38. Shen, Z., Cheng, Y., Chan, R.H., Liò, P., Schönlieb, C.B., Aviles-Rivero, A.I.: Trident: The nonlinear trilogy for implicit neural representations (2023)
39. Sitzmann, V., Martel, J., Bergman, A., Lindell, D., Wetzstein, G.: Implicit neural representations with periodic activation functions. *Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)* **33**, 7462–7473 (2020)
40. Su, K., Chen, M., Shlizerman, E.: Inras: Implicit neural representation for audio scenes. In: *Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)*. vol. 35, pp. 8144–8158 (2022)
41. Sun, Y., Liu, J., Xie, M., Wohlberg, B., Kamilov, U.S.: CoIL: Coordinate-based internal learning for tomographic imaging. *IEEE Transactions on Computational Imaging* **7**, 1400–1412 (2021)
42. Tancik, M., Srinivasan, P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J., Ng, R.: Fourier features let networks learn high frequency functions in low dimensional domains. In: *Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS)*. vol. 33, pp. 7537–7547 (2020)
43. Ulyanov, D., Vedaldi, A., Lempitsky, V.: Deep image prior. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 9446–9454 (June 2018)
44. Umpller, Y., Postels, J., Yang, R., Gool, L.V., Tombari, F.: Implicit neural representations for image compression. In: *Proceedings of the European Conference Computer Vision (ECCV)*. pp. 74–91 (2022)
45. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* **13**(4), 600–612 (2004)

46. Xiang, F., Xu, Z., Hašan, M., Hold-Geoffroy, Y., Sunkavalli, K., Su, H.: NeuTex: Neural texture mapping for volumetric neural rendering. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 7119–7128 (June 2021)
47. Xu, D., Wang, P., Jiang, Y., Fan, Z., Wang, Z.: Signal processing for implicit neural representations. In: Proceedings of the International Conference on Neural Information Processing Systems (NeurIPS). vol. 35, pp. 13404–13418 (2022)
48. Xu, R., Yao, M., Chen, C., Wang, L., Xiong, Z.: Continuous spectral reconstruction from RGB images via implicit neural representation. In: Proceedings of the European Conference Computer Vision (ECCV). pp. 78–94 (2023)
49. Xu, W., Jiao, J.: Revisiting implicit neural representations in low-level vision. In: International Conference on Learning Representations Workshop (2023)
50. Yang, L., Kim, B., Zoss, G., Gözcü, B., Gross, M., Solenthaler, B.: Implicit neural representation for physics-driven actuated soft bodies. *ACM Transactions on Graphics* **41**(4) (July 2022)
51. Yang, S., Ding, M., Wu, Y., Li, Z., Zhang, J.: Implicit neural representation for cooperative low-light image enhancement. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 12918–12927 (October 2023)
52. Zhang, K., Zhu, D., Min, X., Zhai, G.: Implicit neural representation learning for hyperspectral image super-resolution. *IEEE Transactions on Geoscience and Remote Sensing* **61**, 1–12 (2023)
53. Zhao, Q., Asif, M.S., Ma, Z.: Dnerv: Modeling inherent dynamics via difference neural representation for videos. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2031–2040 (June 2023)