Learnable Fourier-based Activations for Implicit Signal Representations

Parsa Mojarad Adi
Shahid Beheshti University
Tehran, Iran
p.mojarad@mail.sbu.ac.ir

Ali Mehrabian
The University of British Columbia
Vancouver, Canada
alimehrabian619@ece.ubc.ca

Abstract

Implicit neural representations (INRs) use neural networks to provide continuous and resolution-independent representations of complex signals with a small number of parameters. However, existing INR models often fail to capture important frequency components specific to each task. To address this issue, in this paper, we propose a Fourier Kolmogorov–Arnold network (FKAN) for INRs. The proposed FKAN utilizes learnable activation functions modeled as Fourier series in the first layer to effectively control and learn the task-specific frequency components. The activation functions with learnable Fourier coefficients improve the ability of the network to capture complex patterns and details, which is beneficial for high-resolution and high-dimensional data. Experimental results show that our proposed FKAN model outperforms four state-of-the-art baseline schemes across various tasks, including image representation, 3D occupancy volume representation, and image inpainting.

1 Introduction

Implicit neural representations (INRs), which model continuous functions from discrete data, have gained attention for their effectiveness in representing 2D images, 3D shapes, neural radiance fields, and other complex structures (Mildenhall et al., 2020; Sitzmann et al., 2020; Park et al., 2019; Shi et al., 2024). Unlike traditional convolutional neural networks (CNNs) which are limited to 3D inputs, coordinate networks use 1D vectors, providing a flexible framework for solving inverse problems in any dimension. INR models build on the multi-layer perceptron (MLP) structure and alternate between linear layers and non-linear activation functions, benefiting from its continuity nature and expressive power. MLP-based INR models avoid the locality bias problem that often restricts the effectiveness of CNNs. However, ReLU-based MLPs in coordinate networks exhibit spectral bias, prioritizing low-frequency signals. As a result, these networks learn high-frequency components more slowly (Rahaman et al., 2019; Xu, 2018; Shi et al., 2024; Radl et al., 2024). This suggests that MLPs generally capture basic patterns in real-world data, focusing on the low-frequency aspects of the target function (Xu, 2018; Arpit et al., 2019).

To overcome the challenge of capturing high-frequency components, several approaches have been explored. Spatial encoding techniques like frequency decomposition, high-pass filtering, and Fourier features (Tancik et al., 2020) help emphasize high-frequency components, while architectural modifications such as multi-scale representations (Saragadam et al., 2022) capture both low-frequency and high-frequency details. Additionally, methods like SIREN (Sitzmann et al., 2020) and WIRE (Saragadam et al., 2023) use periodic activation functions, such as sine functions, for automatic frequency tuning (Ramasinghe & Lucey, 2022; Lindell et al., 2022). However, the aforementioned approaches introduce new challenges. The effectiveness of the SIREN model relies heavily on the

Machine Learning and Compression Workshop at the 38th Conference on Neural Information Processing Systems (NeurIPS 2024).

proper selection of hyperparameters, like frequency. It is sensitive to initialization and requires careful design to prevent random variations. Moreover, due to the unknown frequency distribution of the signal, spatial encoding techniques face a mismatch between the predefined frequency bases and the signal's inherent properties, causing an incomplete or inaccurate representation (Liu et al., 2024a; Xie et al., 2023).

To address the aforementioned issues, in this paper, we propose a novel approach that enhances the hierarchical representation of INRs for improved signal reconstruction in tasks like image representations and 3D structure modeling. We develop an adaptive mapping function that can manage non-linearity and intricate frequency distributions. We hypothesize that a polynomial approximation of activation functions in the initial layer can capture fine-grained high-frequency details. Inspired by Kolmogorov–Arnold networks (Liu et al., 2024b; Xu et al., 2024), we introduce **Fourier Kolmogorov–Arnold network (FKAN)** to learn task-specific frequency components for INRs. Our key contributions are summarized as follows:

- ***FKAN Architecture***: The proposed FKAN adjusts spectral bias using adaptive Fourier coefficients. Specifically, learnable activation functions modeled with the Fourier series enable the network to capture a broad range of frequency information flexibly. By utilizing the spectral characteristics of the Fourier series, they efficiently represent both the low-frequency and high-frequency elements of the input signal.
- ***Performance Evaluation***: We evaluate the performance of the proposed FKAN on signal representation tasks, including image representation and 3D occupancy volume representation, as well as on inverse problems, such as image inpainting. We compare it with the following baselines: SIREN (Sitzmann et al., 2020), WIRE (Saragadam et al., 2023), INCODE (Kazerouni et al., 2024), and FFN (Tancik et al., 2020). Experimental results show that the proposed FKAN can improve the peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) for image-based tasks. The proposed FKAN improves intersection over union (IoU) for the 3D occupancy volume representation task. The proposed FKAN achieves faster convergence than baseline models in both tasks.

2 Problem Formulation

INRs can be interpreted as approximating a function that maps input features to the output signal. As an example, in the context of 2D images, the input features could be spatial coordinates, and the output signal could be pixel values. This mapping function can be parameterized using a neural network. Let $\boldsymbol{x} \in \mathbb{R}^{d_i}$ denote the input features and $\boldsymbol{y} \in \mathbb{C}^{d_o}$ denote the output signal. The neural network that maps the input features to the output signal is denoted as $f(\cdot; \boldsymbol{\Phi}) : \mathbb{R}^{d_i} \to \mathbb{C}^{d_o}$, where $\boldsymbol{\Phi}$ represents the set of neural network parameters.

The parameters $\boldsymbol{\Phi}$ are determined by minimizing the error between the predicted values of the neural network and the ground truth signal. This can be expressed as:

$$\underset{\boldsymbol{\Phi}}{\operatorname{argmin}} \; \frac{1}{N} \sum_{n=1}^{N} \mathcal{L}\left(f\left(\boldsymbol{x}_n; \boldsymbol{\Phi}\right), \boldsymbol{y}_n\right), \tag{1}$$

where $\mathcal{L}$ denotes a pre-defined loss function and N represents the number of training samples. In this paper, we consider the L2 loss function, i.e., $\mathcal{L} = ||f\left(\boldsymbol{x}_n; \boldsymbol{\Phi}\right) - \boldsymbol{y}_n||^2$. Also, $\boldsymbol{x}_n$ and $\boldsymbol{y}_n$ denote the input and output signal for the $n \in \{1, \ldots, N\}$ training sample, respectively.

3 Proposed Fourier Kolmogorov-Arnold Network

To capture task-specific frequency components in a fine-grained manner, we propose FKAN. Motivated by the Kolmogorov-Arnold representation theorem (Kolmogorov, 1957) and KANs (Liu et al., 2024b), which employ learnable activation functions on edges instead of nodes as in vanilla MLPs, our proposed FKAN utilizes learnable activation functions modeled as Fourier series. This approach allows for learning a higher spectral resolution for signals. The first layer of the proposed spectral FKAN can be expressed as follows:

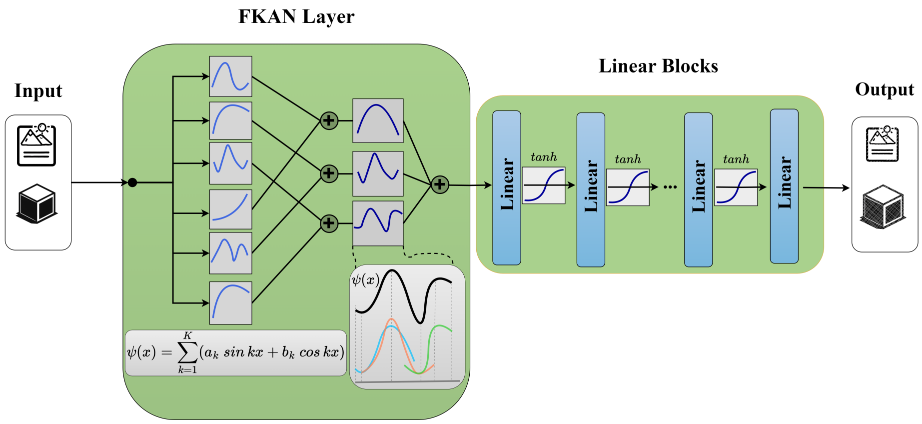

Figure 1: Illustration of the proposed FKAN model. The proposed architecture includes an FKAN block for capturing task-specific frequency components with learnable activation functions and includes L hidden layers to learn non-linear patterns in the signals.

$$\mathbf{\Psi}(\boldsymbol{x}) = \underbrace{\begin{pmatrix} \psi_{1,1}(\cdot) & \psi_{1,2}(\cdot) & \cdots & \psi_{1,d_i}(\cdot) \\ \psi_{2,1}(\cdot) & \psi_{2,2}(\cdot) & \cdots & \psi_{2,d_i}(\cdot) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{H_1,1}(\cdot) & \psi_{H_1,2}(\cdot) & \cdots & \psi_{H_1,d_i}(\cdot) \end{pmatrix}}_{\mathbf{\Psi}(\cdot)} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{d_i} \end{pmatrix}, \tag{2}$$

where $\psi_{i,j}(\cdot) : \mathbb{R} \to \mathbb{R}$ denotes a learnable function. The function matrix $\mathbf{\Psi}(\cdot) : \mathbb{R}^{d_i} \to \mathbb{R}^{H_1}$ transforms the input features into a latent hidden space with dimension H_1. The fundamental idea of KAN is to create an arbitrary function at each hidden neuron through the superposition of multiple non-linear functions applied to the input features.

In (Liu et al., 2024b), spline functions are used to parameterize the learnable functions. However, splines are piecewise polynomial functions, which can be advantageous for localized approximation but require more parameters to achieve similar accuracy globally, resulting in higher training complexity. In addition, splines do not provide a direct frequency-domain representation. To address this issue, as shown in Figure 1, we leverage Fourier series representation (Nussbaumer & Nussbaumer, 1982) to parameterize each learnable function as follows:

$$\psi(x) = \sum_{k=1}^{K} \left(a_k \sin kx + b_k \cos kx\right), \tag{3}$$

where a_k and b_k denote the learnable Fourier coefficients, and K is the number of frequency components (or grid size), which can be fine-tuned as a hyper-parameter. The proposed architecture can control and capture a wide range of frequency components, leveraging the spectral properties of the Fourier series to efficiently represent both low-frequency and high-frequency components of the input signal. Moreover, Fourier series representation has a lower training complexity compared to spline functions. The FKAN with a single layer of learnable activation functions is sufficient to achieve a high-quality spectral representation of the input signal.

To learn the intrinsic non-linear patterns in data, as shown in Figure 1, we utilize L hidden layers, each performing a linear transformation followed by a fixed non-linear activation function. The final layer then applies a linear transformation to generate the output signal. The non-linear activation in hidden layers plays an important role in improving the representation capacity of INRs (Saragadam et al., 2023). To this end, we use the $\tanh(\cdot)$ activation function for the hidden layers. The architecture of the hidden layers is as follows:

$$\begin{aligned} \boldsymbol{h}_i &= \boldsymbol{W}_i \boldsymbol{z}_i + \boldsymbol{b}_i, \\ \gamma_i &= \tanh\left(\omega_0 \boldsymbol{h}_i\right), \quad i = 1, \ldots, L, \end{aligned} \tag{4}$$

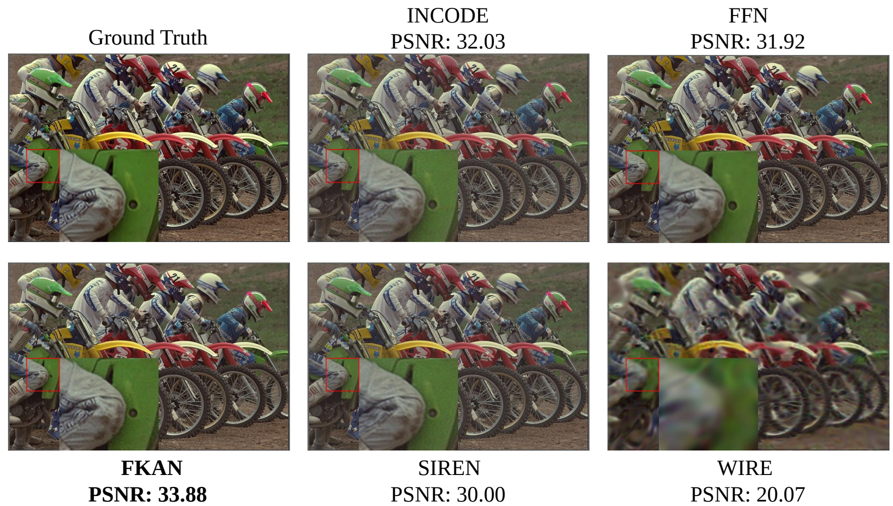

Figure 2: Comparison of the image representation between proposed FKAN and baselines.

where $\boldsymbol{z}_i$ is the input to the i-th hidden layer, with $\boldsymbol{z}_1 = \boldsymbol{\Psi}(\boldsymbol{x})$, $\boldsymbol{W}_i \in \mathbb{R}^{H_i \times H_{i+1}}$ and $\boldsymbol{b}_i \in \mathbb{R}^{H_{i+1}}$ are the learnable weights for the linear transformation in the i-th hidden layer. $\omega_0 \in \mathbb{R}^+$ is a pre-defined positive scalar to control the frequency and convergence of the model, with $\omega_0 = 30$ in our implementations. In addition, we initialize the weights in the hidden layers using uniform distribution $\boldsymbol{W}_i \sim \mathcal{U}\left(-\sqrt{6/d_i}, \sqrt{6/d_i}\right)$.

For the final layer, we apply a linear transformation to generate the output signal as follows:

$$\boldsymbol{y} = \boldsymbol{W}_f \gamma_L + \boldsymbol{b}_f, \tag{5}$$

where $\boldsymbol{W}_f \in \mathbb{C}^{H_L \times d_o}$ and $\boldsymbol{b}_f \in \mathbb{C}^{d_o}$ are the learnable weights for the linear transformation in the final layer.[1]

4 Performance Evaluation

Implementation Details: We evaluate the effectiveness of our proposed FKAN on signal representation tasks, including image representation and 3D occupancy volume representation, as well as on inverse problems, such as image inpainting. Our experiments are conducted on an Nvidia RTX 4070 GPU with 12GB of memory. To implement the neural networks, we use PyTorch library (Paszke et al., 2019) and Adam optimizer (Kingma & Ba, 2015). We choose $H_1 = 128$ for the latent dimension of the FKAN block with grid size $K = 270$. We choose $L = 4$ for the number of hidden layers with 256, 256, 256, and 512 hidden neurons in each layer, respectively. We consider 500 training epochs for image-based tasks and 200 epochs for occupancy volume representation tasks, respectively. We compare the performance of our proposed FKAN with the following baselines: 1 SIREN (Sitzmann et al., 2020), 2 WIRE (Saragadam et al., 2023), 3 INCODE (Kazerouni et al., 2024), and 4 FFN (Tancik et al., 2020).

Table 1: Comparison of the number of parameters and performance for image representation task between methods.

Methods	#Parameters	PSNR (dB)	SSIM
SIREN	$528,387$	33.13 ± 3.78	0.864 ± 0.041
WIRE	$528,643$	30.99 ± 3.44	0.823 ± 0.055
INCODE	$436,775$	$\underline{34.81 \pm 3.78}$	$\underline{0.889 \pm 0.038}$
FFN	$466,179$	33.14 ± 3.28	0.881 ± 0.033
FKAN	$436,367$	$\mathbf{37.91 \pm 3.46}$	$\mathbf{0.939 \pm 0.24}$

[1]As mentioned in Section 2, the output signal can contain complex values. Therefore, we initialize the weights in the final layer as complex numbers to generate a complex-valued output signal. Complex-valued operations are managed based on Wirtinger calculus (Mehrabian & Wong, 2024a; 2024b). For cases where the output signal is real-valued, the weights in the final layer are initialized as real numbers.

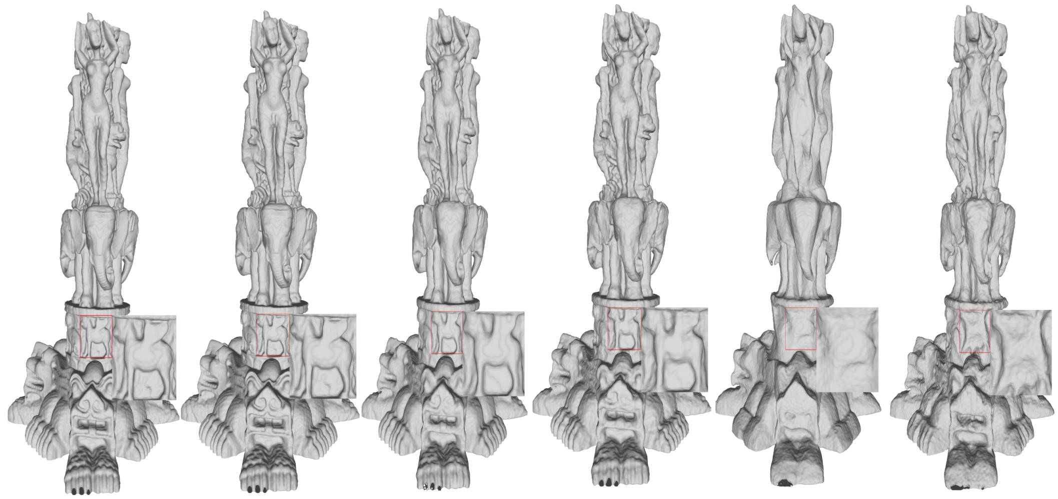

Figure 3: Comparison of the occupancy volume representation between proposed FKAN and baselines.

4.1 Signal Representations

4.1.1 Image Representation

We conducted image representation experiments on the Kodak dataset (Eastman Kodak Company, 1999), which consists of images with resolutions of either 512×768 or 768×512 pixels, all in RGB format. The learning rate is set to 0.0001. To evaluate the performance of the models for the image representation task, we consider PSNR and SSIM metrics. Table I presents the experimental results for the image representation task and the number of parameters for the models. As shown in Table I, the proposed FKAN outperforms all the baselines in both metrics. In particular, the proposed FKAN achieves improvements in PSNR and SSIM metrics, with gains of 8.91% for PSNR and 5.62% for SSIM compared to INCODE as the second-best model, respectively. As depicted in Figure 2, the reconstructed image by FKAN illustrates FKAN's ability to capture intricate details of the ground truth image compared to baselines.

4.1.2 Occupancy Volume Representation

We conduct experiments on the Thai statue dataset from the Stanford 3D Scanning Repository with WIRE system setting (Saragadam et al., 2023), which maps 3D coordinates (i.e., $d_i = 3$) to signed distance function (SDF) values (i.e., $d_o = 1$). We create an occupancy volume through point sampling on a $512 \times 512 \times 512$ grid. The learning rate is set to 0.0001. To evaluate the performance of our proposed FKAN for the occupancy volume representation task, we consider the IoU metric. We plot the reconstructed 3D shapes in Figure 3. We observe that our proposed FKAN model outperforms all the baselines. In particular, the proposed FKAN provides 0.96% improvements on the IoU metric compared to the INCODE. FKAN utilizes learnable activation functions that can capture both low-frequency smooth regions and high-frequency details, resulting in the highest IoU scores.

4.2 Inverse Problems

4.2.1 Image Inpainting

In image inpainting, models are trained on a fraction of the pixel data and are tasked with predicting the entire image. To evaluate the performance of the models, we use an image from the Kodak dataset with a resolution of $768 \times 512 \times 3$. The sampling mask is generated randomly, with an average of 20% of the pixels being sampled. The learning rate is set to 0.001. We plot the reconstructed

Ground Truth	**FKAN**	INCODE	SIREN	FFN	WIRE
	PSNR: 31.09	PSNR: 27.98	PSNR: 27.55	PSNR: 25.38	PSNR: 25.51
	SSIM: 0.8483	SSIM: 0.7617	SSIM: 0.7497	SSIM: 0.6843	SSIM: 0.6712

Figure 4: Comparison of the image inpainting between proposed FKAN and baselines.

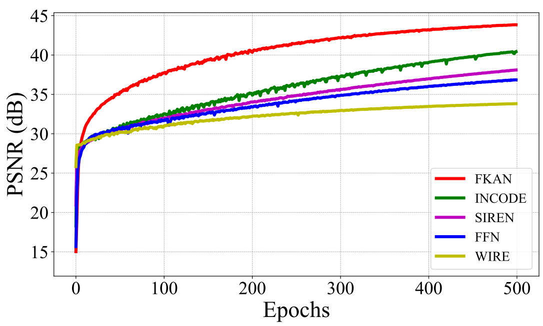

Figure 5: Illustration of the convergence rates of the models for the image representation task.

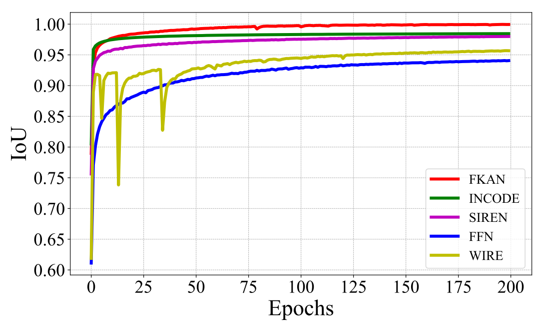

Figure 6: Illustration of the convergence rates of the models for the occupancy volume representation task.

images in Figure 4. As can be seen, the proposed FKAN outperforms all the baselines in both metrics and demonstrates a superior ability to capture intricate features, particularly edges. Unlike other methods that often result in blurred outputs and loss of fine details, FKAN preserves clarity and sharpness, distinguishing itself in terms of image quality. In particular, the proposed FKAN achieves improvements in PSNR and SSIM metrics, with gains of 11.12% for PSNR and 11.37% for SSIM compared to INCODE as the second-best model, respectively.

4.3 Comparison of Convergence Rate

In Figure 5, we plot the convergence rate of the models for the image representation task. We observe that the proposed FKAN has a faster convergence compared to baselines and there is a significant gap between the proposed FKAN and INCODE as the second-best model. In Figure 6, we plot the convergence rate of the models for the occupancy volume representation task. We observe that the proposed FKAN has a faster convergence compared to all baselines. This indicates that FKAN effectively captures and represents complex data structures with high accuracy, achieving superior results in fewer training iterations.

5 Conclusion

In this paper, we proposed FKAN for implicit neural signal representations. The proposed FKAN utilizes learnable activation functions modeled as Fourier series to capture task-specific frequency components and learn complex patterns of high-dimensional signals in a fine-grained manner. We investigated the performance of our proposed FKAN on various tasks, namely image representation, image inpainting, and 3D occupancy volume representation. Experimental results demonstrate that our proposed FKAN outperforms four state-of-the-art baselines with faster convergence. It improves the PSNR and SSIM for the image representation and image inpainting tasks and IoU for the 3D occupancy volume representation task, respectively. For future work, we will consider the neural radiance field task.

References

Devansh Arpit, Stanisław Jastrzebski, Nicolas Ballas, David Krueger, Emmanuel Bengio, Maxinder S Kanwal, Tegan Maharaj, Asja Fischer, Aaron Courville, Yoshua Bengio, et al. A closer look at memorization in deep networks. In *Proc. Int'l Conf. Machine Learning (ICML)*, Sydney, Australia, Aug. 2019.

Amirhossein Kazerouni, Reza Azad, Alireza Hosseini, Dorit Merhof, and Ulas Bagci. INCODE: Implicit neural conditioning with prior knowledge embeddings. In *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis. (WACV)*, Santa Rosa, CA, Jan. 2024.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *Proc. Int'l Conf. Learn. Representations (ICLR)*, San Diego, CA, May 2015.

Andrei Nikolaevich Kolmogorov. On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. In *Doklady Akademii Nauk*, volume 114, pp. 953–956. Russian Academy of Sciences, 1957.

David B. Lindell, Dave Van Veen, Jeong Joon Park, and Gordon Wetzstein. BACON: Band-limited coordinate networks for multiscale scene representation. In *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, New Orleans, LA, June 2022.

Zhen Liu, Hao Zhu, Qi Zhang, Jingde Fu, Weibing Deng, Zhan Ma, Yanwen Guo, and Xun Cao. FINER: Flexible spectral-bias tuning in implicit neural representation by variable-periodic activation functions. In *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, Jun. 2024a.

Ziming Liu, Yixuan Wang, Sachin Vaidya, Fabian Ruehle, James Halverson, Marin Soljačić, Thomas Y Hou, and Max Tegmark. KAN: Kolmogorov-Arnold networks. *arXiv preprint arXiv:2404.19756*, 2024b.

Ali Mehrabian and Vincent W. S. Wong. Joint spectrum, precoding, and phase shifts design for RIS-aided multiuser MIMO THz systems. *IEEE Trans. Commun.*, 72(8):5087–5101, Aug. 2024a.

Ali Mehrabian and Vincent W. S. Wong. Adaptive bandwidth allocation in multiuser MIMO THz systems with graph-transformer networks. In *Proc. of IEEE Int. Conf. Commun. (ICC)*, Denver, CO, Jun. 2024b.

Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing scenes as neural radiance fields for view synthesis. In *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Virtual, Aug. 2020.

Henri J Nussbaumer and Henri J Nussbaumer. *The fast Fourier transform*. Springer, 1982.

Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, and Steven Lovegrove. DeepSDF: Learning continuous signed distance functions for shape representation. In *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Long Beach, CA, Jun. 2019.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. PyTorch: An imperative style, high-performance deep learning library. In *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, Vancouver, Canada, Dec. 2019.

Lukas Radl, Andreas Kurz, Michael Steiner, and Markus Steinberger. Analyzing the internals of neural radiance fields. In *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, Jun. 2024.

Nasim Rahaman, Aristide Baratin, Devansh Arpit, Felix Draxler, Min Lin, Fred Hamprecht, Yoshua Bengio, and Aaron Courville. On the spectral bias of neural networks. In *Proc. Int'l Conf. Machine Learning (ICML)*, Long Beach, CA, Jun. 2019.

Sameera Ramasinghe and Simon Lucey. Beyond periodicity: Towards a unifying framework for activations in coordinate-MLPs. In *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Tel Aviv, Israel, Oct. 2022.

Vishwanath Saragadam, Jasper Tan, Guha Balakrishnan, Richard G Baraniuk, and Ashok Veeraraghavan. Miner: Multiscale implicit neural representation. In *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Tel Aviv, Israel, Oct. 2022.

Vishwanath Saragadam, Daniel LeJeune, Jasper Tan, Guha Balakrishnan, Ashok Veeraraghavan, and Richard G Baraniuk. WIRE: Wavelet implicit neural representations. In *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Vancouver, Canada, Jun. 2023.

Kexuan Shi, Xingyu Zhou, and Shuhang Gu. Improved implicit neural representation with Fourier reparameterized training. In *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, Jun. 2024.

Vincent Sitzmann, Julien N.P. Martel, Alexander W. Bergman, David B. Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. In *Proc. Conf. Neural Inf. Process. Syst. (NeurIPS)*, Virtual, Dec. 2020.

Matthew Tancik, Pratul P. Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan T. Barron, and Ren Ng. Fourier features let networks learn high frequency functions in low dimensional domains. In *Proc. Conf. Neural Inf. Process. Syst. (NeurIPS)*, Virtual, Dec. 2020.

Shaowen Xie, Hao Zhu, Zhen Liu, Qi Zhang, You Zhou, Xun Cao, and Zhan Ma. DINER: Disorder-invariant implicit neural representation. In *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Vancouver, Canada, Jun. 2023.

Jinfeng Xu, Zheyu Chen, Jinze Li, Shuo Yang, Wei Wang, Xiping Hu, and Edith C-H Ngai. FourierKAN-GCF: Fourier Kolmogorov-Arnold network–An effective and efficient feature transformation for graph collaborative filtering. *arXiv preprint arXiv:2406.01034*, 2024.

Zhiqin John Xu. Understanding training and generalization in deep learning by Fourier analysis. *arXiv preprint arXiv:1808.04295*, 2018.

Contents

A Baseline Methods

We adopt four state-of-the-art baseline methods for performance comparison. We introduce the baselines as follows:

1 **SIREN (Sinusoidal Representation Networks)**: SIREN utilizes periodic activations (sine functions) into MLPs for representing continuous signals like images and 3D shapes. The sinusoidal activations allow it to capture high-frequency details and represent structures that standard neural networks struggle with, particularly in neural rendering and solving PDEs. SIREN also overcomes initialization issues associated with periodic activations by using a specialized weight initialization method, ensuring effective gradient flow throughout the network. We download the source code from github.com/vsitzmann/siren and follow the recommended system settings.

2 **WIRE (Wavelet Implicit Neural Representations)**: WIRE utilizes a Gabor wavelet activation function to introduce spatial and frequency compactness, making it more robust for signal processing tasks like image reconstruction. The wavelet transform offers fast signal approximation rates and better fits to image signals, avoiding overfitting to noise. We download the source code from github.com/vishwa91/wire and follow the recommended system settings with 2D Gabor activation function.

3 **INCODE (Implicit Neural Conditioning with Prior Knowledge Embeddings)**: INCODE is designed to handle multi-modal data by embedding prior knowledge into the INR. It extends traditional INRs by conditioning the model on learned embeddings, improving its capacity to reconstruct and generalize across different signal types. We download the source code from github.com/xmindflow/INCODE and follow the recommended system settings.

4 **FFN (Fourier Feature Networks)**: FFN employs Fourier feature mappings to project input coordinates into a high-dimensional space, allowing the neural network to better capture fine details and complex patterns, especially in periodic or high-frequency signals. We download the source code from github.com/tancik/fourier-feature-networks and follow the recommended system setting.