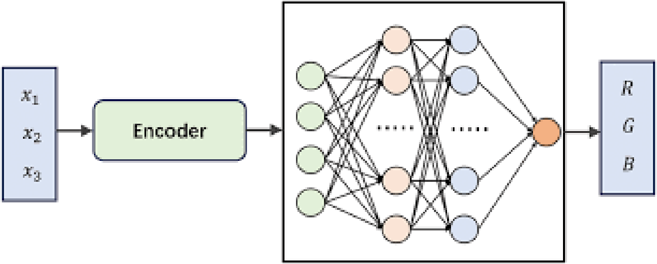
Deep Learning and Neural Networks

Demetrio Labate

April 12, 2026

Implicit Neural Representations (INRs)

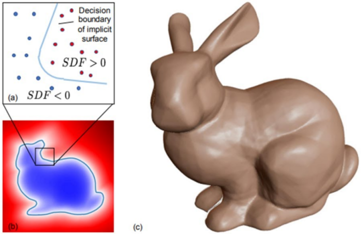
Implicit Neural Representations (INRs) (sometimes also referred to as coordinate-based representations) are coordinate-based, continuous functions parameterized by neural networks that map input coordinates directly to signal values like RGB, signed distance, or density.



INRs

Conventional signal representations are usually discrete – for instance, images are discrete grids of pixels, and 3D shapes are parameterized as grids of voxels or meshes. Such representations are computationally and memory intensive, especially in 3D.

INRs parameterize a signal as a continuous function that maps the domain of the signal (i.e., a coordinate) to some value at that coordinate (e.g., an R,G,B color). These functions are usually not analytically tractable. INRs approximate that function via a neural network.



INRs

Benefits:

- **Low memory.** INRs are not coupled to spatial resolution, in the the way an image is coupled to the number of pixels, for instance. This is because they are continuous functions. Thus, the memory required to parameterize the signal is independent of spatial resolution, and only scales with the complexity of the underlying signal.
- **Infinite resolution.** For the same reason, INRs can be sampled at arbitrary spatial resolutions.
- **Meta-learning.** Generalizing across neural implicit representations amounts to learning a prior over a space of functions, implemented via learning a prior over the weights of neural networks.

INRs - Core Formulation

An INR models any continuous field $y \in \mathbb{R}^m$ as a neural function

$$f(x;\Theta) : \mathbb{R}^D \mapsto \mathbb{R}^m$$

where $x \in \mathbb{R}^D$ is a coordinate (pixel location, 3D position, time, hybrid spatio-angular features, etc.) and Θ are the network weights.

This is achieved via ℓ^2 minimization over a discrete sample $\{(x_i, y_i)\}_{i=1}^N$

$$\theta^* = \arg\min_\theta \sum_{i=1}^N \|f(x_i;\theta) - y_i\|^2$$

INRs - Core Formulation

Key mathematical properties:

- **Continuity:** INRs interpolate naturally between sampling points and are not tied to a discretization scale.
- **Differentiability:** Smooth activations allow for arbitrary-order spatial derivatives, critical for applications involving PDE constraints or signal processing on the latent field [Essakine et al., 2024, Xu et al., 2022].
- **Adaptive capacity:** Width, depth, and underlying activation/encoding parameterization directly control expressive power, with theoretical frameworks such as the neural tangent kernel (NTK) relating initialization and convergence characteristics to harmonic structure [Yüce et al., 2021, Ko et al., 19 Aug 2025].

INRs - Core Formulation

The **main challenge** for INRs is **spectral bias towards lower frequencies**, that is, the tendency to lose high frequency details present signals, e.g., textures in images.

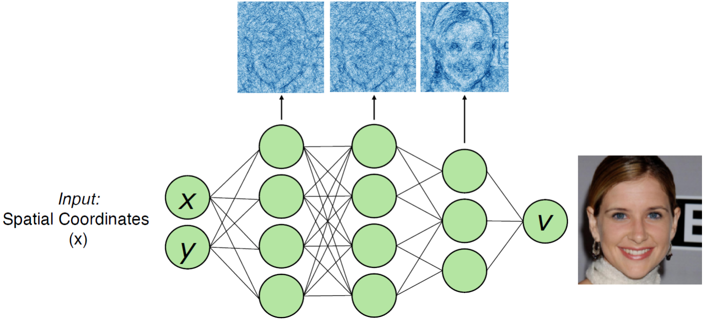


Figure: INRs fit images in a coarse-to-fine manner

[Rahaman, N., Baratin, A., Arpit, D., Dräxler, F., Lin, M., Hamprecht, F.A., Bengio, Y., Courville, A.C. *On the spectral bias of neural networks.* ICML (2018)]

INRs - Core Formulation

Most INR architectures can be decomposed into

1. a mapping function (or, positional encoding) $\gamma\colon \mathbb{R}^D \mapsto \mathbb{R}^T$
2. a MLP with weights W^ℓ, biases b^ℓ and activation functions ρ^ℓ applied componentwise, at each layer $\ell = 1, \ldots, L-1$.

Thus, if we denote by $z^{(\ell)}$ each layer's post activation, the INR architecture computes

$$\begin{aligned} z^{(0)} &= \gamma(r) \\ z^{(\ell)} &= \rho^\ell(W^\ell z^{(\ell-1)} + b^\ell), \quad \ell = 1, \ldots, L-1 \\ f(r; \Theta) &= W^L z^{(L-1)} + b^L \end{aligned}$$

To ameliorate the spectral bias problem and allow faster convergence and greater fidelity of INR, different solutions have been proposed consisting of special selections of the mapping and activation functions.

Fourier feature networks (FFNs)

- 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 domain*. 2020

They proposed a **Fourier mapping**

$$\gamma(r) = \sin(\Omega r + \phi)$$

with parameters $\Omega \in \mathbb{R}^{T \times D}$ and an MLP with $\rho^\ell = ReLU$.

Sinusoidal representation networks (SIRENs)

- Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein. *Implicit neural representations with periodic activation functions*. 2020

They proposed a **Fourier mapping**

$$\gamma(r) = \sin(\Omega r + \phi)$$

with parameters $\Omega \in \mathbb{R}^{T \times D}$ and an MLP with $\rho^\ell = \sin()$.

NERFs

- *NeRF. Representing Scenes as Neural Radiance Fields for View Synthesis.* ECCV 2020, Authors: Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, Ren Ng

Method for synthesizing novel views of complex scenes by optimizing an underlying continuous volumetric scene function

$$f(x, y, z, \theta, \phi; \Theta) \colon \mathbb{R}^5 \mapsto (R, G, B, \sigma) \in \mathbb{R}^4$$

Here the input is a single continuous 5D coordinate (spatial location (x, y, z) and viewing direction (θ, ϕ) and the output is the volume density and view-dependent emitted radiance at that spatial location.

They use a **Fourier mapping**

$$\gamma(r) = \sin(\Omega r + \phi)$$

with parameters $\Omega \in \mathbb{R}^{T \times D}$ and an MLP with $\rho^{\ell} = ReLU$.

Neural Radiance Fields

INRs - Expresive power

The Fourier mapping γ acts as a collection of single frequency basis, whose spectral support is expanded after each non-linear activation into a collection of higher order harmonics.

Fact: *The expressive power of FFNs and SIRENs is restricted to functions that can be expressed as a linear combination of certain harmonics of the mapping function* γ.

That is, these architectures have the same expressive power as a structured signal dictionary whose atoms are sinusoids with frequencies equal to sums and differences of the integer harmonics of the mapping frequencies. The bandwidth increases exponentially with network depth and polynomial activation order.

[Yüce, Gizem, Guillermo Ortiz-Jiménez, Beril Besbinar, and Pascal Frossard. "A structured dictionary perspective on implicit neural representations." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 19228-19238. 2022]

Architectural Variants

Recent surveys systematize INR methods into four major classes.

- **Activation-centric INRs** use periodic or spatially localized kernels to break the spectral bias of vanilla MLPs, which otherwise favor low-frequency content (spectral bias). The SIREN architecture employs scaled sine nonlinearities, enabling direct representation of high-frequency signals. Other architectures employ Gabor and wavelet-inspired activations which combine spatial and frequency localization (Ko et al., 19 Aug 2025, Roddenberry et al., 2023).
- **Encoding-centric INRs** map inputs through fixed or random orthogonal/frequency bases before feeding them to conventional activations. Examples of encoding strategies include Fourier features, random Fourier maps, multi-scale projections. This approach augments the representational bandwidth without increasing MLP size, but choices of scale and distribution must be matched to the task.

Architectural Variants

- **Hybrid and combined approaches** (e.g., FLAIR, Trident) explicitly unite frequency selectivity, space-frequency localization, and region-adaptive input modulation, often combining wavelet transforms or frequency guides with specialized activations (RC-GAUSS in FLAIR) (Ko et al., 19 Aug 2025).
- **Structural optimizations** include dynamically conditioned activations, modularization, cross-layer harmonization, compressive and sparse architectures (Meta-SparseINR), and stacked hybrid blocks for both deep MLPs and deep conv–MLP hybrids (INRN) (Song et al., 2022, Lee et al., 2021).

Kernel-Space Transformations and Coordinate Mappings

Beyond changing the INR architecture, the transformation of input and output domains (kernel-space design) is effective for improving performance.

- Linear input scaling and output shifting (SS-INR, (Zheng et al., 7 Apr 2025)): Scaling input coordinates expands effective receptive bandwidth; output shifting centers signal values, acting as a form of normalization. These transformations effectively acting as zero-parameter "pseudo-layers".
- Coordinate reparameterization via hash embeddings (DINER, (Xie et al., 2022, Zhu et al., 2023)): Learnable hash tables reorder sample coordinates, restructuring the empirical spectrum so the subsequent MLP operates over a "smoother" representation. This nullifies the negative effects of coordinate disorder and spectral bias, dramatically accelerating convergence and enhancing high-frequency modeling.

INRs

Implicit Neural Representations Repository

Implicit Neural Representations for Medical Imaging - Tutorial

Learning 3D Reconstruction in Function Space