A Structured Dictionary Perspective on Implicit Neural Representations.

Gizem Yüce*, Guillermo Ortiz-Jimenez*, Beril Besbinar and Pascal Frossard

Can we unify all INRs under one formulation?

Sure, we can.

Can we characterize their expressive power?

Yes! They are signal dictionaries of input harmonics.

...and their inductive bias?

Of course! Just compute the eigenvectors of their NTK.

An overview of INRs.

INRs represent signals $\mathbb{g} : \mathbb{R}^D \rightarrow \mathbb{R}^C$ using neural networks $\mathcal{f}_\theta : \mathbb{R}^D \rightarrow \mathbb{R}^C$ that map coordinates $r \in \mathbb{R}^D$ (e.g., pixels) to signal values $\mathbb{g}(r) \in \mathbb{R}^C$ (e.g., color).

INRs circumvent spectral bias, but how?

Their architecture can be unified as:

$$\mathcal{f}_\theta(r) = \sum_{k=0}^{N} \sum_{l=1}^{D} \mathcal{W}^{(k)} \mathcal{U}^{(l)} r_l^{(l-1)} + b^{(k)}.$$ 

Dictionary perspective.

Fourier Feature Networks (FFNs)

Sinusoidal Representations (SIREN)

Expressive power of INRs.

Theorem (Informal): An INR $\mathcal{f}_\theta : \mathbb{R}^D \rightarrow \mathbb{R}^C$ with polynomial activations $\rho(s) = \sum_{k=0}^{N} a_k s^k$ can only represent functions formed by linear combinations of integer harmonics of its input mapping $\gamma(r) = \sin(2\pi r + \phi)$.

$$\mathcal{f}_\theta(r) = \sum_{k=0}^{N} \sum_{l=0}^{D} \mathcal{W}^{(k)} \mathcal{U}^{(l)} r_l^{(l-1)} + b^{(k)}.$$ 

Only linear combination

Proof intuition: Activations split input in harmonics $\gamma(r) = 2\pi r \rightarrow \rho(\gamma(r)) = \rho(2\pi r) = \sum_{k=0}^{N} a_k s^k$.

As a result, depth $\ell$ allows to represent high spectral details, i.e., $O(2^\ell \mathcal{K}^2)$ frequencies with only $O(\ell \mathcal{K}^2)$ parameters. Efficiency comes from the low rank frequency mixing structure.

Failure modes of INRs.

Imperfect recovery

If the set of harmonics $\mathcal{H}(\ell)$ does not properly cover the spectrum of the signal, INRs lead to imperfect recovery.

Only the harmonics of $z \times \ell_4$ are reconstructed.

Aliasing

If the set of harmonics $\mathcal{H}(\ell)$ contains very high frequencies, the network cannot resolve the ambiguity, leading to aliasing.

Increasing the high frequency emphasis $w_4$ creates very high frequency atoms in $\mathcal{H}(\ell)$ that lead to aliasing at higher rates.

Inductive bias of INRs.

Overview of NTK

Linearize neural network using Taylor

$\Theta(x, z) = \Theta(0, z) + \Theta(x, z) \cdot \nabla \Theta(0, z)$

Linearization is equivalent to kernel predictor given by:

In kernel methods, the inductive bias is given by alignment with the first eigenfunctions of kernel.

$$\Theta(x_1, z_1) = \sum_{i=1}^{\text{Eigenvale}} \lambda_i \phi_i(x_1) \phi_i(x_2)$$

$\mathcal{L}(x) = \sum_{k=0}^{N} \mathcal{W}^{(k)} \mathcal{U}^{(l)} r_l^{(l-1)} + b^{(k)}$

Energy concentration

Atom projection

Meta-learning is dictionary learning

Meta-learning has been proposed to improve efficiency of INRs.

Using the signal dictionary perspective we can see that meta-learning reshapes the eigenfunctions of the NTK using combinations of the meta-training examples.

This boosts the energy concentration akin to classical dictionary learning.