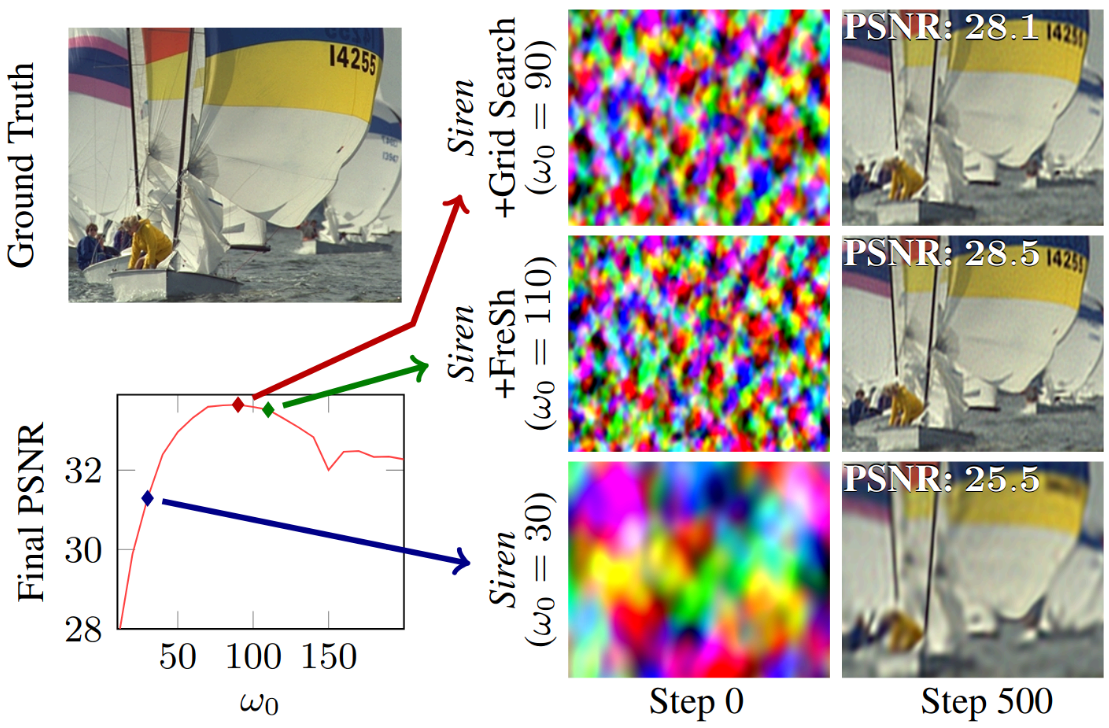
# FreSh: Frequency Shifting For Accelerated Neural Representation Learning

Adam Kania[1], Marko Mihajlovic[2], Sergey Prokudin[2 3], Jacek Tabor[1], Przemysław Spurek[1 4]
Jagiellonian University[1]; ETH Zurich[2]; Balgrist University Hospital[3]; IDEAS NCBR[4]

## TL;DR

- **Implicit Neural Representations** (INRs) are neural models used for continuously representing signals such as images and 3D shapes. They typically consist of a MLP preceded by a frequency embedding.
- We show the importance of choosing a good **embedding frequency** for achieving good performance.
- We propose FreSh - an algorithm for choosing **optimal embedding frequency without expensive grid searches**.




## Motivation

**Implicit Neural Representations (INRs)**
INRs are MLP-based models that have recently gained attention as a powerful approach for continuously representing signals such as images, videos, and 3D shapes. Their example use-case is reconstuction of a 3D object just from images of the object.

**Embedding layer**
The first layer of an INR maps coordinates into feature vectors through a high-frequency function, the frequency of which is controlled by a hyperparameter.
For example, in Siren the embedding is given as:

$$\mathrm{embedding}(\mathbf{x}) = \sin(\omega_0 \mathbf{W}\mathbf{x} + \mathbf{b})$$

**Embedding frequency vs. performance**
Frequency of the embedding layer greatly affects both training time and final quality. However, typically no tuning of this parameter is performed as it requires training multiple models for each tested signal.

## Efficient embedding configuration

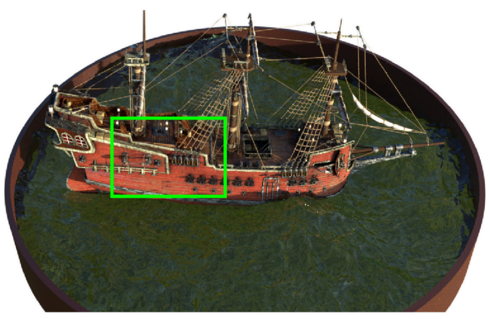
A typical (and costly) approach is to apply a grid search and train models with every possible embedding configuration. We aim to similarly **find the best embedding configurations**, but we want to be efficient.

Instead of training, we analyze frequencies present in the model and the target dataset - **frequency works as a good predictor of performance**, which we use to find a good configuration of the model.

## FreSh works with

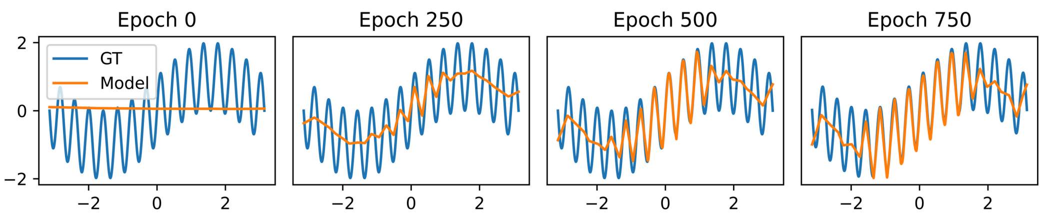
| **Models**: | **Tasks**: |
|---|---|
| Siren | image |
| Finer | video |
| Fourier Features | 3D object (NeRF) |
| Hash Encoding | |

Ground Truth



## Background

**Spectral bias**
Spectral bias (frequency bias) makes neural models slower at fitting higher frequency signals. Frequency embedding are required to fight this bias, as INRs have to fit the high frequency details.
For example, here a MLP tries to fit y=sin(x)+sin(15x):



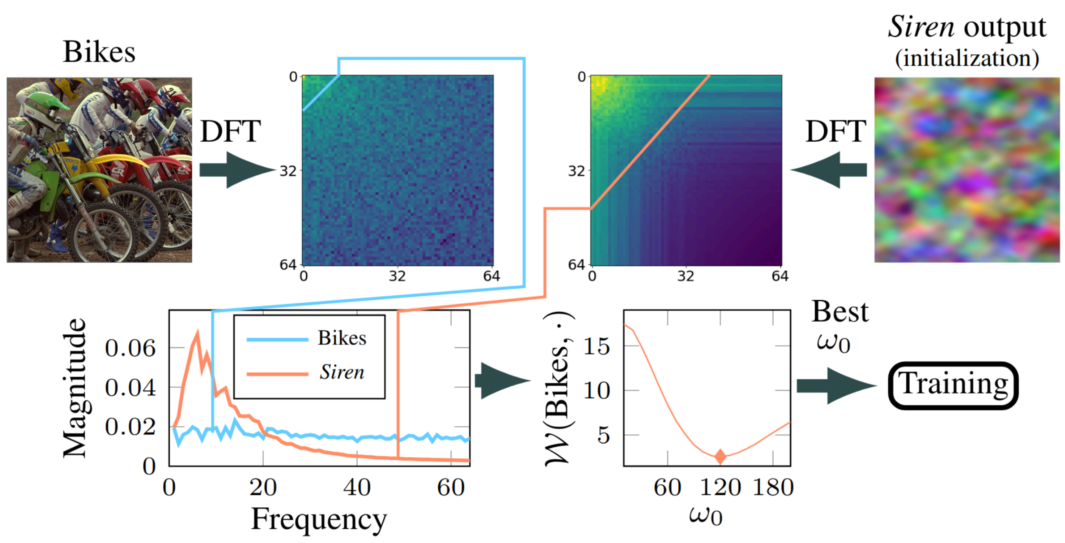

**Discrete Fourier Transform (DFT)**
2-dimensional DFT maps an image $A \in \mathbb{R}^{C\times N\times N}$ into the frequency domain. For the *c*-th channel of the image, it is given as:

$$\mathcal{F}_{j,k}(A_c) = \sum_{m=0}^{N-1} e^{-i2\pi\frac{jm}{N}} \sum_{n=0}^{N-1} e^{-i2\pi\frac{kn}{N}} A_{c,m,n}$$

**Spectrum vector**
DFT describes not only magnitude, but also direction of a given frequency. We remove the unnecessary directional information by summing the same frequencies, resulting in a spectrum vector:

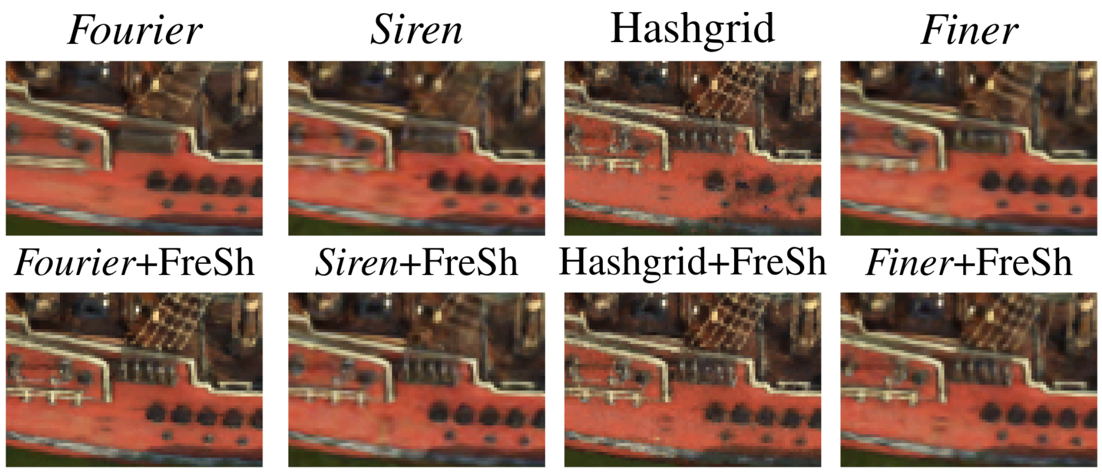
$$\mathcal{S}(A,d) = \sum_{c\in\{0,\dots,C-1\}} \sum_{\substack{(i,j)\in\{0,\dots,N-1\}\\ i+j=d}} |\mathcal{F}_{i,j}(A_c)|$$

**Wasserstein distance**
Wasserstein distance is a distance function between probability distributions. For discrete, 1-dimensional distributions, Wasserstein distance is the L1 norm of difference of CDFs.




## Fresh

FreSh improves INR performance by aligning the frequency of the neural model with its target dataset. FreSh algorithm:
1. Compute output of an untrained model.
2. Calculate spectrum vectors of the dataset and model's output.
3. Compare the output spectrum with the spectrum of the target signal using the Wasserstein distance.
4. Repeat steps 1-3 for each configuration that you wish to test.
5. Train the model using the Wasserstein-best configuration.




1) Rahaman, Nasim, et al. "On the spectral bias of neural networks." International conference on machine learning. PMLR, 2019.
2) Ronen, Basri, et al. "The convergence rate of neural networks for learned functions of different frequencies." Advances in Neural Information Processing Systems 32, 2019.
3)Sitzmann, Vincent, et al. "Implicit neural representations with periodic activation functions." Advances in neural information processing systems 33 (2020): 7462-7473.

Check out the project!


adkania@alumni.uj.edu.pl