Paper Summary BACON

Abstract

BACON: Band-limited Coordinate Networks for Multiscale Scene Representation. Coordinate-based network are trained to map continuous input coordinate to the value of a signal at each point. **Training at a single scale** will result in artifacts when naive downsampling and upsampling. BACON can be designed based on the **spectral characteristic of the represented signal** at unsupervised signal. Demonstrate BACON for

1. Multiscale neural representation of images.
2. radiance fields.
3. 3D scenes using **signed distance functions**

1. Introduction

Neural representations approximate signals using a continuous function that is embedded in the learned weights of a fully-connected NN. Since it is designed to represent signals at a single scale, the behavior of the NN at unsupervised coordinate is difficult to predict.

The key properties of BACON architecture are:

1. the maximum frequency at each layer can be manipulated analytically.
2. The behavior of a trained network is entirely characterized by its **Fourier spectrum**.

Contribution:

1. Introducing BACON for representing and optimizing
2. Developing methods for spectral analysis of the architecture and proposing initialization scheme

2. Related Work

Architectures for Scene Representation

NN architectures for scene representation networks can be classified as **feature-based**, **coordinate-based**, and **hybrid**.

Feature-based: quickly evaluated, but have a large memory footprint.

Coordinate-based: Map from an input coordinate to a signal value, and the proposed one in this paper is coordinate-based. Didn't use MLP but **multiplicative filter networks**(MFNs, recently proposed).

3. Method

3.1 Band-limited Coordinate Networks

MLP employ a Hadamard product b/t linear layers and sine activation functions. Extending the theoretical understanding and practicality of MFN by:

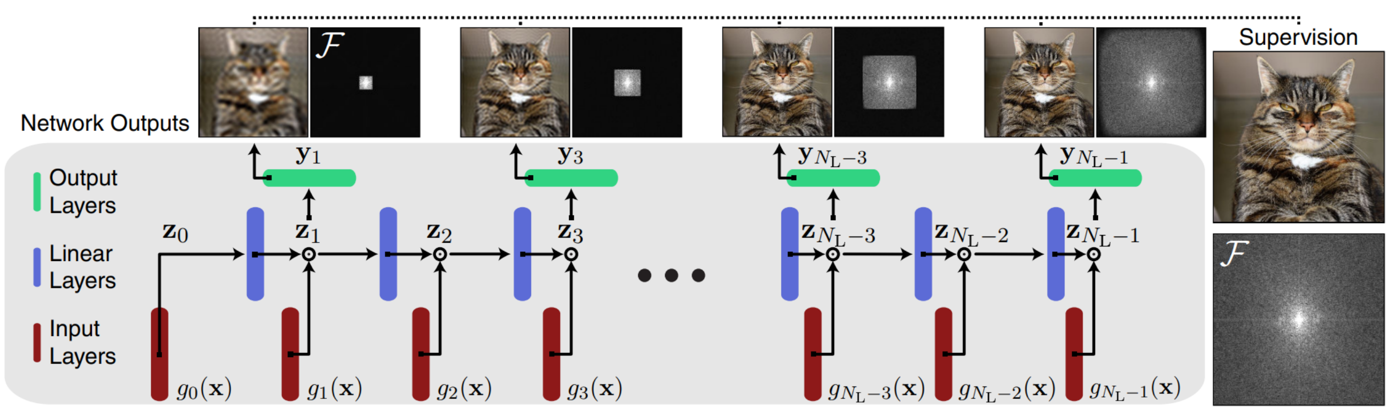
1. Architecture change to achieve multi-scale, band-limited outputs
2. Deriving formulas to quantify the expected frequencies in the representation
3. Deriving initialization scheme preventing vanishing activations in deep networks

$x \in R^{d_{in}}$ -> $g_i(x) = sin(\omega_i x + \phi_i), i = 0, \ldots, N_L - 1$, and N_L is the number of layers in the network. We refer to the intermediate activations as $z_i \in R^{d_h}$ and allow intermediate outputs of the network $y^i \in R^{d_out}$ at the i^{th} layer. The given expression:

$z_0 = g_0(x)$

$z_i = g_i(x) \circ (W_i z_{i-1} + b_i) \quad 0 \leq i \leq N_L$

where $\circ$ indicates the Hadamard product $\quad y_i = W_i^{out} z_i + b_i^{out}$



Some expressions:

$$y_i = \sum_{j=0}^{N_{sine}^{(i)}-1} \bar{\alpha}_j sin(\bar{\omega}_j x + \bar{\phi}_j)$$

$$N_{sine}^{(N_L)} = \sum_{i=0}^{N_L-1} 2^i \, d_h^{i+1}$$

3.2 Frequency Spectrum

MFN can be expressed as a sum of sines to create band-limited networks.