Generalizing Neural Radiance Fields

BAIR Commons Completion Report

Introduction

Coordinate-based neural representations for low dimensional signals are becoming increasingly popular in computer vision and graphics. In particular, these fully-connected networks can represent 3D scenes more compactly than voxel grids but are still easy to optimize with gradient-based methods. We presented Neural Radiance Fields (NeRF) last year [1], a technique for achieving photorealistic view synthesis of complex objects and scenes. Since then, we have continued to further investigate and explain the capabilities of these coordinate-based networks. Below are two extension projects. The first investigates the positional encoding used in NeRF using tools from the neural tangent kernel literature. The second project uses meta learned initialization for fast optimization of NeRFs on new scenes of a given class.

Fourier Feature Mapping [3]

We show that passing input points through a simple Fourier feature mapping enables a multilayer perceptron (MLP) to learn high-frequency functions in low-dimensional problem domains. These results shed light on recent advances in computer vision and graphics that achieve state-of-the-art results by using MLPs to represent complex 3D objects and scenes. Using tools from the neural tangent kernel (NTK) literature, we show that a standard MLP fails to learn high frequencies both in theory and in practice. To overcome this spectral bias, we use a Fourier feature mapping to transform the effective NTK into a stationary kernel with a tunable bandwidth. We suggest an approach for selecting problem-specific Fourier features that greatly improves the performance of MLPs for low-dimensional regression tasks relevant to the computer vision and graphics communities.

Meta Learned Initialization [2]

Coordinate-based neural representations have shown significant promise as an alternative to discrete, array-based representations for complex low dimensional signals. However, optimizing a coordinate-based network from randomly initialized weights for each new signal is inefficient. We propose applying standard meta-learning algorithms to learn the initial weight parameters for these fully-connected networks based on the underlying class of signals being represented (e.g., images of faces or 3D models of chairs). Despite requiring only a minor change in implementation, using these learned initial weights enables faster convergence during optimization and can serve as a strong prior over the signal class being modeled, resulting in better generalization when only partial observations of a given signal are available.

References

[1] 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 *ECCV*, 2020.

[2] Matthew Tancik, Ben Mildenhall, Terrance Wang, Divi Schmidt, Pratul P Srinivasan, Jonathan T Barron, and Ren Ng. Learned initializations for optimizing coordinate-based neural representations. In *CVPR*, 2021.

[3] 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 *NeurIPS*, 2020.