Midterm proposal

D11921B15 程煒倫

In the field of image compression, handling high-frequency information is crucial. The Neural Tangent Kernel (NTK) theory offers a new approach to deeply understand the behavior of neural networks, particularly during network initialization and early training phases. This is vital for designing efficient neural networks for image compression, enhancing their capability to process high-frequency information, and thereby optimizing the quality of compression. My midterm proposal aims to explore the application of NTK theory in the field of image compression. My goal is to optimize neural networks through NTK theory to more effectively handle high-frequency information in images, thereby improving the overall quality of image compression.

Effectiveness and Purpose of the Core Paper

- ✓ **Key Focus of the Paper:** The paper, "Fourier Features Let Networks Learn High Frequency Functions in Low Dimensional Domains" (NeurIPS 2020 spotlight), demonstrates how using Fourier feature mapping enables a multilayer perceptron (MLP) to learn high-frequency functions in low-dimensional problem domains. This is critically important for the latest advancements in computer vision and graphics, where MLPs are used to represent complex 3D objects and scenes, achieving state-of-the-art results.
- **Technical Principle:** Utilizing tools from the Neural Tangent Kernel (NTK) literature, the authors show that standard MLPs struggle to effectively learn high-frequency information. By applying Fourier feature mapping, the effective NTK is transformed into a stationary kernel with tunable bandwidth, thereby enhancing MLP's performance in low-dimensional regression tasks.
- **Application of NTK:** NTK theory facilitates a deeper understanding and design of neural network architectures, especially in scenarios where network width tends toward infinity.

It provides a theoretical basis for learning dynamics during the network's initialization phase, helping to optimize the network's training and generalization capabilities.

Methods to be Employed and Their Rationale

- **Discrete Cosine Transform (DCT):**
    - **Reason for Selection:** DCT is particularly effective in image processing for highlighting low-frequency components of images, making it highly useful for video compression.
    - **Expectation:** Employing DCT could enhance the efficiency of video compression, especially for videos with substantial low-frequency information.
- **Discrete Wavelet Transform (DWT):**
    - **Reason for Selection:** DWT can capture information across different scales, providing a good representation of both high and low-frequency information.
    - **Expectation:** DWT could offer a more balanced frequency domain representation, aiding in more effective video compression, particularly in maintaining high-frequency details.
- **Combination with Convolutional Neural Networks (CNN) or Autoencoders:**
    - **Reason for Selection:** These neural network structures excel in image and video processing, particularly in feature extraction and compression.
    - **Expectation:** Integrating CNNs or autoencoders might enhance the performance of video compression, especially in terms of feature preservation and reconstruction.

Expected Results and Experiment Planning

- Objective: The main goal of the experiment is to explore the application of the aforementioned methods in video compression and their impact on compression efficiency and video quality.
- **Experimental Design:**
    - **Comparative Analysis:** Compress the same video dataset using different feature mappings and network architectures

to compare their performance in terms of compression ratio, reconstruction quality, and processing time.

- **Performance Evaluation:** Assess the performance differences in video compression using DCT, DWT, and traditional Fourier mappings, with a particular focus on the quality and compression ratio of the compressed videos.

In this study, we also aim to explore the application of NTK theory in the field of video compression. We anticipate using NTK theory to optimize the initialization and early training of neural networks for more effective processing of high-frequency information in video data. These theoretical insights could help enhance the performance of video compression algorithms, particularly in scenarios with larger network widths. By combining NTK theory with other compression techniques, we hope to develop more efficient and accurate video compression solutions.