Project Proposal: Efficient Real-Time Super-Resolution

Pichsorita Yim, Jan Remennik, Shu Zhang

Problem

For this project, we aim to train models for real-time single-image super-resolution using two architectures: ResNet-based CNN architecture and Transformer-based architecture. Our goal is to optimize and compare the performances of these two models on the same dataset.

Motivation

A real-time creation of super-resolution single images is a prerequisite for upscaling video that is transmitted via low-bandwidth streaming. Upscaling images also has the potential to enhance perception for computer vision systems. Additionally, this technology could be applied to improve image quality for old or compressed pictures.

Dataset

We will train and evaluate on the DIV2K Dataset, which contains a split of 800 high-resolution and 4x bicubically downscaled training images and a testing split of 100 images. It is widely used as the standard benchmark for single-image super-resolution.

Methods

1. **Residual CNN**: Build a ResNet-CNN model employing convolutional layers with residual connections.
2. **Transformer Model**: Implement a lightweight Transformer-based model to capture global dependencies. We plan to use the SwinIR architecture by Liang et al.

Outcome / Deliverable

1. Quantitative analysis of the resulting upscaled images, comparing inference time, PSNR or similar metrics.
2. Qualitative zoom-in examples of upscaled images.
3. Source code, project final report, and video presentation.