2024S-T3 BDM 3035 - Big Data Capstone Project 01
(DSMM Group 1 & Group 3)

Milestone 2

AI Image Remastering

Guide
Meysam Effati

Group E
Kuncheria Tom C0900973
Prince Thomas C0894907
Remya Kannan Kandari C0895293
Shanmuga Priyan Jeevanandam C0889053
Sravya Somala C0907007

.

Overview

In Milestone 2, our primary objective was to develop and test a deep learning model for super-resolution of images. Initially, we focused on training an Enhanced Super-Resolution Generative Adversarial Network (ESRGAN) from scratch using our dataset. However, due to hardware limitations and unsatisfactory results, we shifted our focus to exploring pre-trained models for super-resolution tasks, specifically SRGAN and ESRGAN.

Task Breakdown

1. **Model Implementation and Training**:
    - **Implementation**: We implemented the ESRGAN architecture which includes a Generator and Discriminator. The Generator consists of an initial convolutional block, multiple residual blocks, and final output layers. The Discriminator follows a sequence of convolutional blocks with LeakyReLU activations.
    - **Training**: The model was trained using BCEWithLogitsLoss for the GAN loss and L1Loss for content loss. Adam optimizer was used for both Generator and Discriminator with a learning rate of 1e-4. Training was conducted for 100 epochs.
2. **Data Handling**:
    - We created a custom ImageDataset class to handle low-resolution (LR) and high-resolution (HR) images. The dataset was split into training and validation sets with 684 images for training and 169 images for validation.
3. **Challenges Encountered**:
    - **Hardware Limitations**: Training the ESRGAN model on Google Colab with TPU and T4 GPU proved challenging. The TPU frequently crashed after 7-8 epochs with a batch size of 16. Moreover, training on T4 GPU was not feasible due to extended training times and resource constraints.
    - **Results**: The generator model's output after the final epoch was significantly poor when evaluated using Peak Signal-to-Noise Ratio (PSNR) and visual inspection compared to simpler CNN models. This indicated a need for more training data and extended training duration, which was not possible given our hardware limitations.
4. **Decision to Use Pre-trained Models**:
    - Given the hardware constraints and subpar results from our trained ESRGAN model, we decided to explore pre-trained models. We specifically focused on pre-trained models like SRGAN and ESRGAN, which are known for their efficiency and quality in single-image super-resolution tasks.

Next Steps

1. **Explore Pre-trained Models**:
    - We plan to experiment with pre-trained models available for SRGAN and ESRGAN. These models are trained on extensive datasets and optimized for performance, which can potentially overcome the limitations faced with our custom model.
2. **Model Evaluation**:
    - We will evaluate the performance of the pre-trained models on our dataset using PSNR and Structural Similarity Index (SSIM) metrics. Visual inspection will also be used to assess the quality of the generated high-resolution images.
3. **Fine-tuning**:
    - If necessary, we will fine-tune the pre-trained models on our dataset to further enhance their performance and adapt them to our specific requirements.

Conclusion

Milestone 2 was pivotal in understanding the challenges associated with training deep learning models for image super-resolution given limited hardware resources. The shift towards using pre-trained models is a strategic move to achieve our project goals efficiently. We anticipate that leveraging pre-trained models will significantly improve the quality of super-resolution outputs, making our approach more feasible and effective.

Project GitHub Link: https://github.com/shanmugapriyan357/Big-Data-Capstone-Project-AI-Image-Remastering-/tree/Dev