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

Milestone 3

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

1. Overview

In Milestone 3, our focus shifted to utilizing transfer learning with pre-trained weights from the Real-ESRGAN model to address hardware constraints and achieve improved results in image super-resolution. We attempted to integrate and train these weights but faced challenges due to limited computational resources. As a result, we plan to leverage open-source pre-trained models to meet project goals effectively.

2. Task Breakdown

2.1 Model Implementation and Training

- **Model Architecture**: We implemented the Real-ESRGAN model components, including:
    - **Residual Block**: A critical component consisting of convolutional layers with Batch Normalization and PReLU activations.
    - **Generator**: A network with an initial convolutional block, multiple residual blocks, and a final convolutional layer to generate high-resolution images.
    - **Discriminator**: A network to distinguish between real and generated high-resolution images, with layers that include convolutional operations and LeakyReLU activations.
- **Training Process**: Despite using pre-trained weights, the results were suboptimal due to:
    - Limited epochs (only 4 successfully run on TPU V2).
    - Poor performance metrics due to hardware limitations preventing extended training.

2.2 Data Handling

- **Dataset**: The dataset was processed and used for training and validation. Training was conducted with a split of 684 images for training and 169 images for validation.

2.3 Challenges Encountered

- **Hardware Limitations**: The Google Colab TPU and GPU resources were insufficient for extended training. Training for 1000+ epochs was unfeasible, resulting in limited model performance.
- **Training Duration**: Despite using memory-efficient techniques and lighter models like StyleGan-ADA, only 5-6 epochs were manageable, leading to subpar results.

3. Next Steps

3.1 Use Open-Source Pre-trained Models

- We plan to explore and utilize open-source pre-trained models available for image super-resolution tasks. This approach is expected to provide better performance without requiring extensive training resources.

3.2 Model Evaluation

- **Metrics**: Evaluate the performance of the open-source pre-trained models using metrics such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). Visual inspection will also be conducted to assess the quality of enhanced images.

3.3 Fine-Tuning

- **Adjustment**: If needed, we will fine-tune the pre-trained models to adapt them to our specific dataset and project requirements.

3.4 Deployment

- **Web Application**: Integrate the best-performing model into a web application using Flask or Streamlit, ensuring user-friendly access to the image enhancement tool.

4. Conclusion

Milestone 3 highlighted the constraints of training deep learning models under hardware limitations. Despite challenges in achieving satisfactory results with transfer learning and pre-trained weights, the decision to utilize open-source pre-trained models is expected to overcome these limitations and enhance the quality of our image super-resolution approach. Moving forward, the focus will be on leveraging these models to achieve our project goals and deploying the solution effectively.

5. References

- StyleGAN-ADA GitHub Repository
- Real-ESRGAN GitHub Repository
- Fine-tune Real-ESRGAN

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