Image Restoration Using Transformers

Muahmmad Moiz Uddin Farooqui
Roll No: 23K-8024
Supervisor: Dr. Maria Siddiqua
Department of Computer Science, National University of Computer and Emerging Sciences, 22-G, Block-6, PECHS, Karachi, Pakistan.
k238024@nu.edu.pk

1. Abstract

Sand image restoration is vital in deep learning and computer vision, especially for autonomous vehicles and satellite applications like environmental planning and disaster response. Traditional methods using CNNs and GANs perform well but struggle with capturing global context. This thesis explores transformer-based models for sand image restoration, leveraging their self-attention mechanism for improved feature learning. A synthetic dataset simulating sand degradation is created, and a Pix2Pix GAN baseline model is implemented. The proposed work involves applying vision transformer models and evaluating their performance against the baseline, with detailed methodology and future plans outlined in this report.

2. Introduction

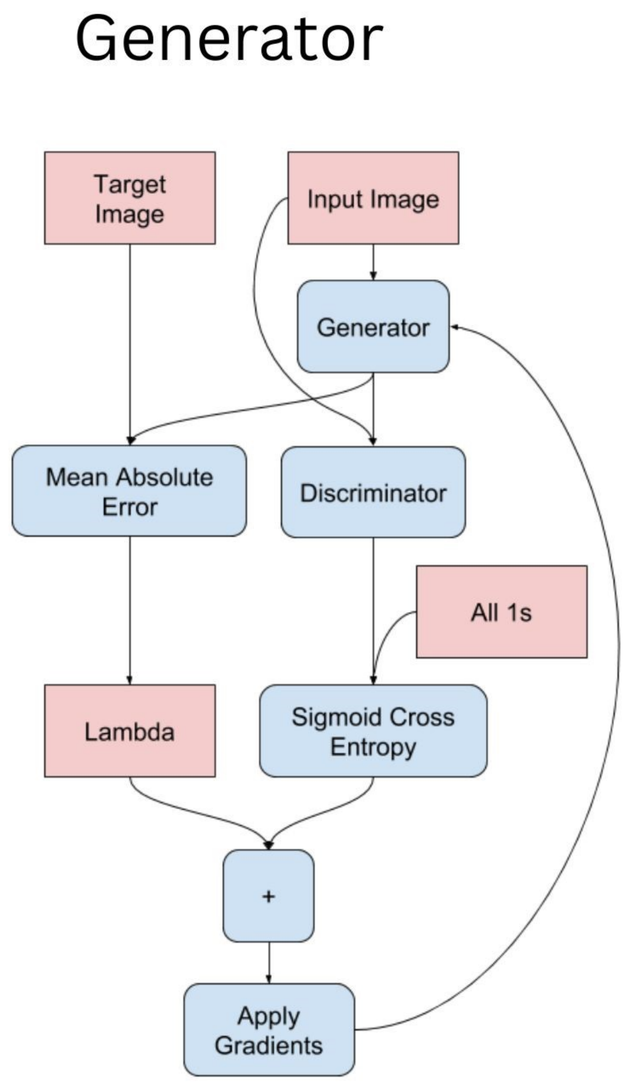
With the rise of autonomous vehicles and satellite imaging, the need for effective image restoration is growing rapidly. Environmental factors like sand and dust degrade image quality, impacting systems that rely on visual data for decision-making. This research aims to develop an efficient model to restore sand-degraded images, enhancing the reliability of autonomous systems. Traditional methods like GANs (e.g., Pix2Pix, CycleGAN) have shown promise but struggle with long-range dependencies and often require paired datasets. Transformer-based models, particularly Vision Transformers (ViTs), address these limitations using self-attention mechanisms for better feature learning. Our work involves three stages: creating a synthetic sand-degraded dataset, implementing GAN-based baseline models, and applying ViT-based models for improved restoration. This research has applications in autonomous driving, satellite imaging, drone vision, and security surveillance.

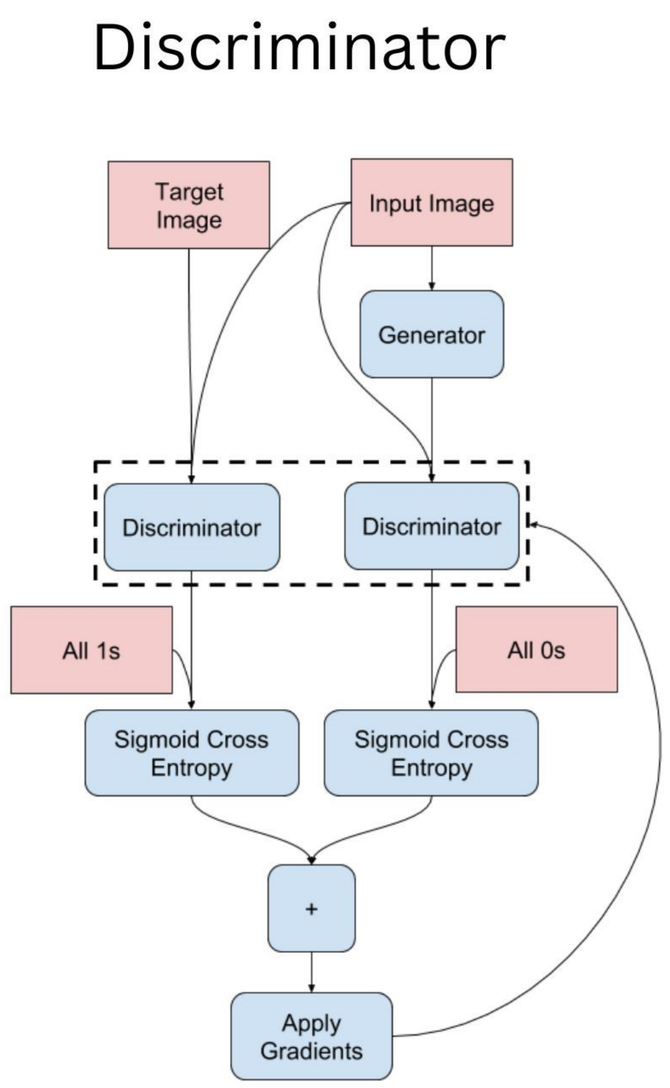
3. Dataset Preparation

Google Streetview Dataset has been used for implementation of our models . It consists of around 23000 images. The data was divided into training(70) and test(30) dataset . The data has been synthesized by adding sand effect to make the images look like sand degraded

4. Implementation of Baseline Model (Pix2Pix)

In this model we have trained conditional generative adversarial network (cGAN) called pix2pix that learns a mapping from input images to output images, as described in [11] by Isola et al. (2017). pix2pix is not application specific—it can be applied to a wide range of tasks, including synthesizing photos from label maps, generating colorized photos from black and white images, turning Google Maps photos into aerial images, and even transforming sketches into photos.

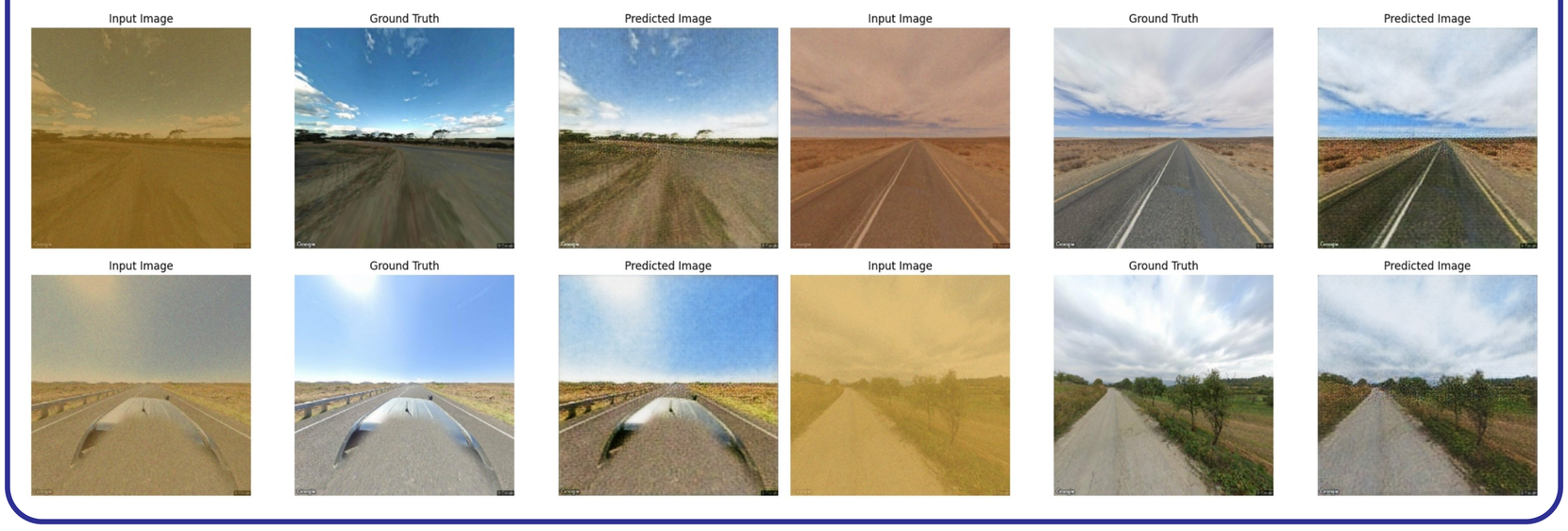




5. Experiment Results

Here are three experimental observations from the model showing the results of sand image restoration

- Significant Dust Removal and Sky Recovery
- Preservation of Structural Details
- Remaining Grain Artifacts and Sharpness Issues



7. Conclusiom

In conclusion, the baseline model shows promising results when trained and tested on the sand-degraded image dataset. It effectively reduces the visual noise and restores key features demonstrating the potential of GAN-based architectures for image restoration tasks. However, certain limitations still persist. These shortcomings highlight the need for more advanced models. This sets the foundation for our proposed research—to implement and evaluate Vision Transformer-based models, which leverage self-attention mechanisms to enhance restoration.