Face Sketch-to-Image Generation Using GANs

Manahil Sarwar
Department of Artificial Intelligence,
FAST-NUCES, Islamabad, Pakistan
Roll No: 21I-0293, Section: AI-K

***Abstract*—This report discusses the implementation of a GAN-based model for generating realistic photos from face sketches. The project aims to enhance the capabilities of neural networks in converting hand-drawn sketches into high-quality images, using a combination of convolutional layers and advanced loss functions.**

I. Introduction

The task of generating realistic images from sketches is a challenging problem in computer vision. Generative Adversarial Networks (GANs) have shown impressive results in image generation tasks by learning to map input sketches to realistic images. This assignment aims to apply GANs to convert sketches of human faces into their corresponding photorealistic images. We utilize a custom dataset of paired sketches and photos to train a model capable of high-quality image synthesis.

II. Methodology

A. Dataset

The dataset consists of paired face sketches and their corresponding photos. The images were split into training, validation, and test sets. The sketches serve as the input, while the photos act as ground truth labels for the training process. Fig. 1 shows examples from the dataset.

B. Preprocessing

We performed data augmentation by resizing images to 64x64 pixels, normalizing pixel values to [-1, 1], and converting them to tensors. This preprocessing ensured the model learned from diverse input images while maintaining computational efficiency.

C. Model Architecture

We designed a GAN-based architecture consisting of a Generator and a Discriminator. The Generator takes a random noise vector and a face sketch as input and generates a photorealistic image. The Discriminator classifies images as real or fake and employs a minibatch discrimination layer to promote diversity in the generated images.

Generator: The Generator consists of transposed convolution layers with ReLU activation and batch normalization, ensuring smooth gradient flow and effective learning. It accepts both noise vectors and sketch embeddings as input.

Discriminator: The Discriminator is composed of convolutional layers with LeakyReLU activations and minibatch discrimination. This network attempts to classify real images

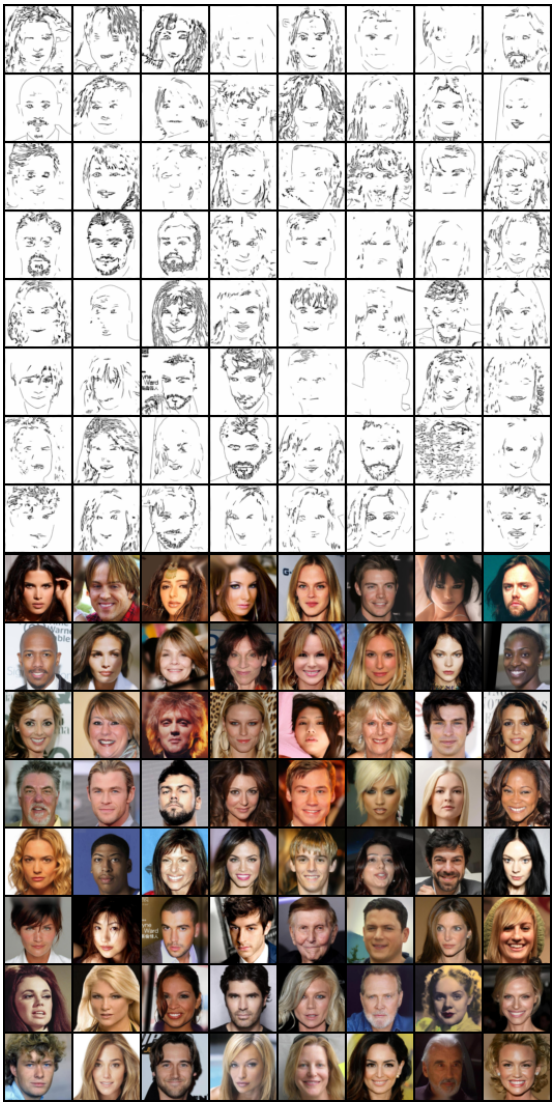

Fig. 1. Examples from the dataset: Left - Sketch, Right - Corresponding Photo.

and sketches against the generated ones, thereby training the Generator to improve over time.

III. Results

The model was trained for 15 epochs using the Adam optimizer with different learning rates for the Generator and Discriminator.

Generated images for the provided sketches are shown in Fig. 2.

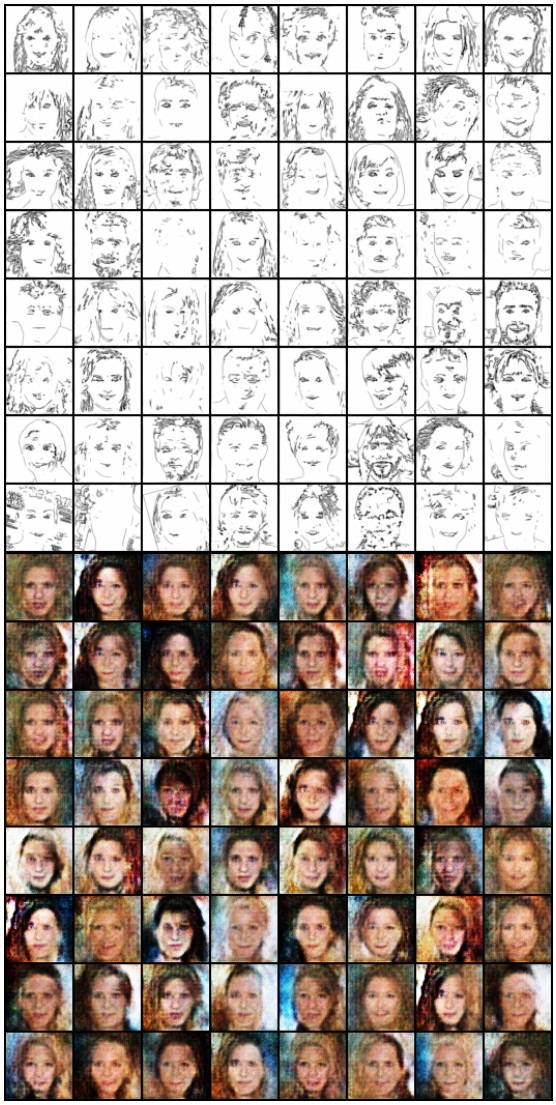

Fig. 2. Generated images based on input sketches: Left - Sketch, Right - Generated Image.

IV. Discussion

Over the course of training, we observed significant improvements in the quality of generated images. Initially, the model struggled to produce coherent images, but as training progressed, it learned to generate sharper and more realistic faces. One challenge encountered was ensuring diversity in generated outputs. The inclusion of minibatch discrimination helped mitigate this issue.

Another challenge was the sensitivity of GANs to hyperparameters. Careful tuning of the learning rates for the Generator and Discriminator was required to stabilize training.

V. Conclusion

In this project, we successfully implemented a GAN model capable of generating realistic face images from sketches. The combination of convolutional layers, feature matching, and minibatch discrimination led to stable training and high-quality image generation. Future work could involve scaling the model to higher-resolution images or experimenting with more advanced GAN architectures.

VI. References

References