# Image-to-Image Translation Using CycleGAN for Sketches and Photos

Manahil Sarwar
Department of Artificial Intelligence,
National University of Computer and Emerging Sciences (FAST),
Islamabad, Pakistan

***Abstract*—This report outlines the development of a CycleGAN model for image-to-image translation between sketches and photos. The model aims to generate realistic photos from sketches and vice versa. The process involves training two generators and two discriminators to perform the image transformation while maintaining the quality of the outputs. We also developed a front-end interface for visualizing the generated results.**

## I. Introduction

The goal of this project is to create an image-to-image translation model based on the CycleGAN architecture to convert images between two domains: sketches and real-life photos. This has practical applications in areas such as sketch-based image retrieval and art-to-image translation. Our system is designed to generate high-quality results by employing adversarial training using convolutional neural networks (CNNs) for both the generator and discriminator models.

Fig. 1. Example of Dataset

## II. Methodology

### A. Dataset

The dataset used consists of paired sketches and corresponding real-life photos. It was split into training, validation, and test sets, each containing a balanced number of images from both domains.

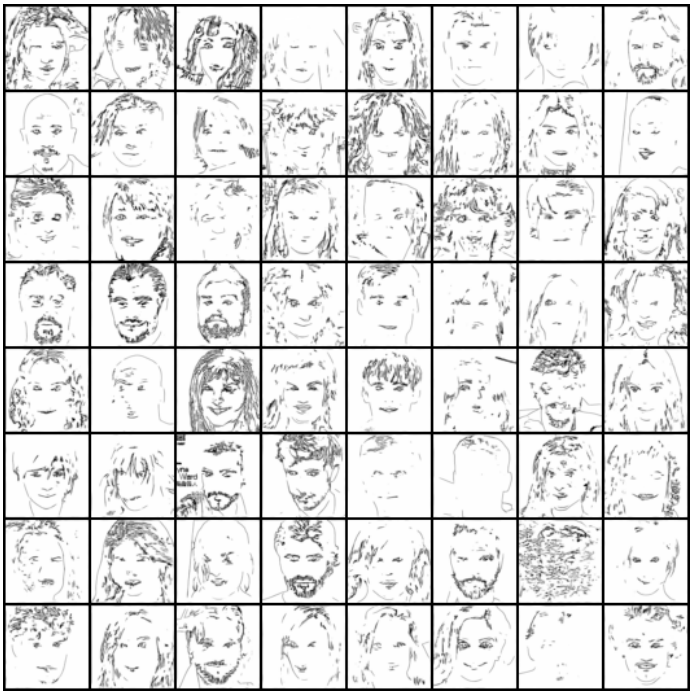

Fig. 2. Example of Dataset

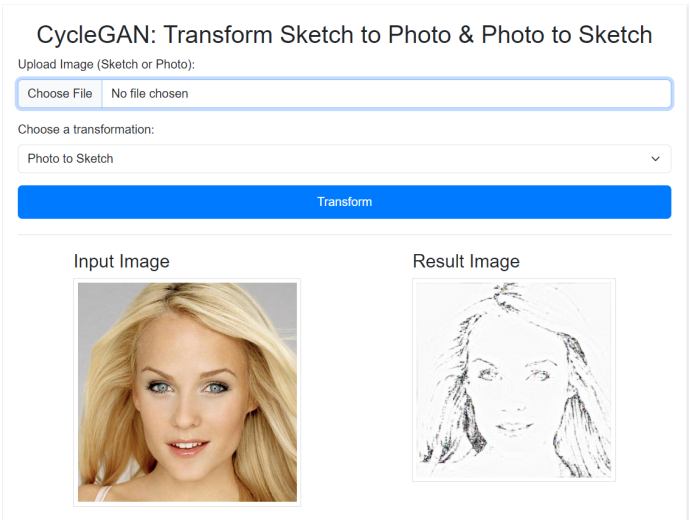


Fig. 3. Frontend interface for input and output visualization.

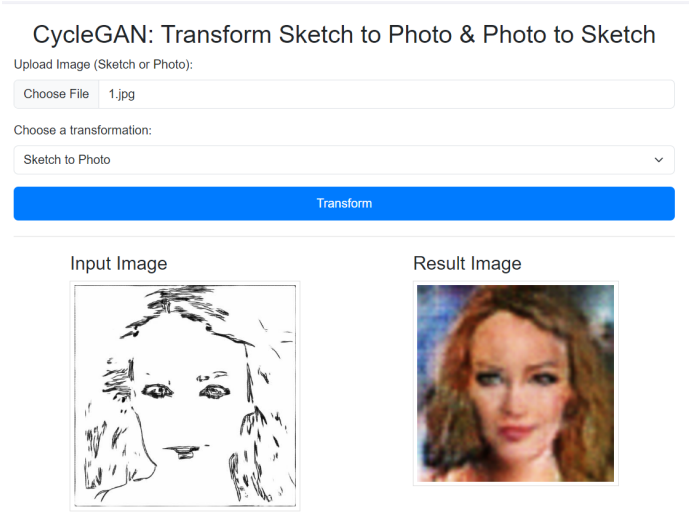


Fig. 4. Frontend interface for input and output visualization.

### *B. Preprocessing*

Images were resized to 64x64 pixels and normalized with a mean and standard deviation of 0.5. The preprocessing also involved converting the images to tensors using PyTorch's `transforms` package.

### *C. Model Architecture*

The CycleGAN architecture was implemented with two generator networks and two discriminator networks. The generator network consists of convolutional and residual blocks to facilitate image transformation, while the discriminator network identifies real and fake images. Each generator was trained to map an input image from one domain (sketch) to another domain (photo) while ensuring the inverse transformation yields a cycle-consistent image.

### *D. Frontend Interface*

A web-based frontend was created using HTML, CSS, and JavaScript to visualize the generated images. It allows users to upload their own sketches or photos and see the translated outputs. Figure 4 shows the front-end interface.

## III. Results

After 20 epochs of training, the model generated realistic photo images from sketches and vice versa. Examples of the generated results are displayed below.

### *A. Generated Images*

Figures 5 show the input sketch, the generated photo, the original photo, and the generated sketch from our model.

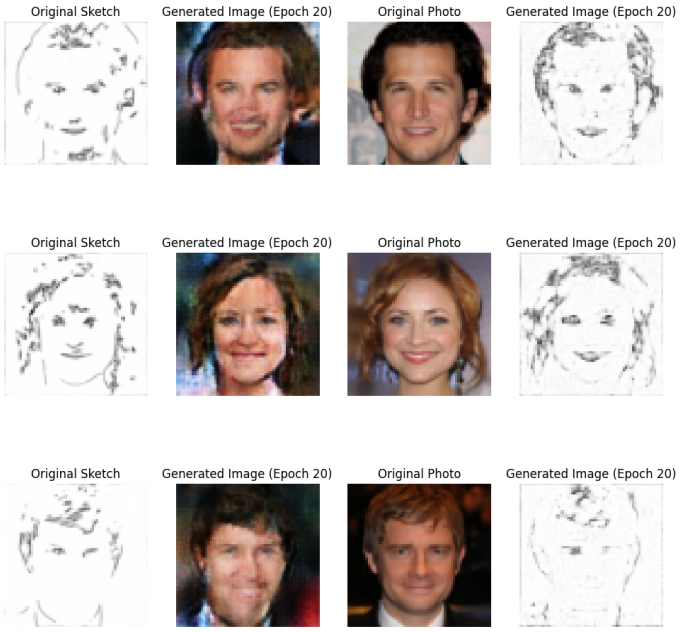


Fig. 5. Examples of Generated Photos from Sketches (Epoch N).

## IV. Discussion

Throughout the training process, the model's predictions improved noticeably as the number of epochs increased. Initially, the generated images were blurry and lacked clear detail. However, by epoch 10, both the generators began producing clearer outputs, and the cycle-consistency loss helped preserve fine details between transformations.

One of the challenges encountered was ensuring that the generators did not collapse into a trivial mapping, which was mitigated by fine-tuning the learning rate and implementing instance normalization.

## V. Conclusion

In conclusion, the CycleGAN model successfully performed image-to-image translation between sketches and photos. The results improved over time, and the adversarial loss, combined with cycle-consistency loss, enabled the generation of high-quality images. Future work can focus on improving the resolution of generated images and exploring other types of image pairs.