

Sketch-to-Image Colorization Using Ensemble U-Nets

Team Name
Tensor

October 18, 2024

Abstract

This project presents an ensemble architecture for sketch-to-image translation, aimed at converting hand-drawn criminal sketches into same sized fully colorized image. It integrates a U-Net-based generator (inspired by pix2pix) for image translation and a discriminator for enhancing output quality. The system is designed for use by law enforcement and security officers to generate realistic images from sketches, potentially assisting in criminal identification and investigations.

1 About the Project

1.1 Motivation

Sketch artists are often required to draw suspects' portraits based on witness descriptions. This system is designed to assist in transforming these sketches into realistic color images that can be used for identification, media releases, or further investigations. This application bridges the gap between traditional criminal sketching and modern AI-assisted image generation, thus reducing reliance on manual sketch coloring.

1.2 Key Components

1. **Custom designed Ensemble U-Net architecture:** The generator is inspired from the U-Net architecture, where the input sketch is progressively downsampled and then upsampled via the Encoder and Decoder block respectively, capturing fine details while translating the sketch into a fully colored image.
2. **Discriminator:** A patchGAN discriminator, adapted from pix2pix, is employed to distinguish between real and generated images, ensuring the output image quality is as close to real photographs as possible.

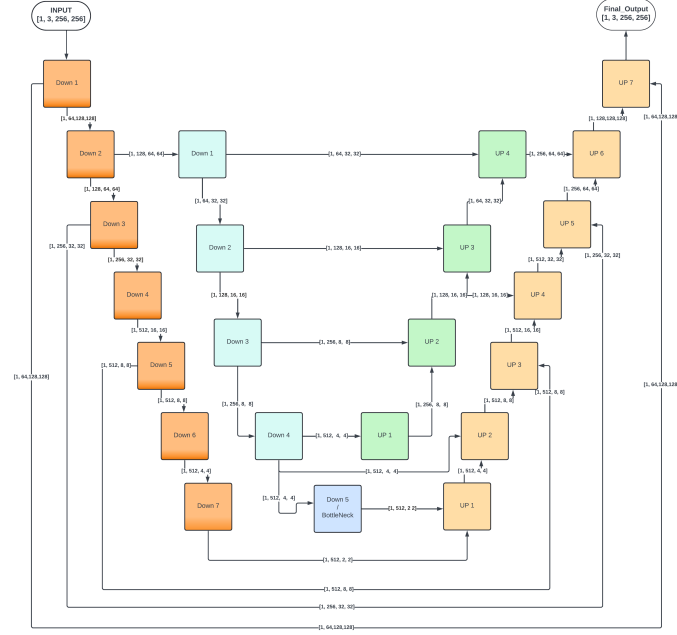


Figure 1: Current version of custom designed Generator architecture

3. **Application Use Case:** This system is built for police officers, sketch artists, and other law enforcement agents who often require criminal sketch colorization during criminal investigations.

2 Current Progress

2.1 Tasks Completed:

- **Model Architecture Design:** Completed the design and implementation of the U-Net generator for image translation. Integrated a **patch-GAN**-based discriminator to assess the realism of the generated images. The current architecture is near-finalized but minor modifications might be made in near-future for optimizing the performance. The current version of architecture, we propose can be sited at in Figure ??.
- **Dataset Preparation:** We have successfully prepared a dataset of human sketches and their corresponding colorized images to train the model. For this purpose, we have used the **CelebA** dataset, and reverse engineered using Sobel and Canny filters for edge detection and then inverted the grayscale output to generate the sketches.
- **Model Training:** Conducted preliminary training sessions on smaller

datasets to validate the architecture's functionality.

- **Initial Testing:** Achieved satisfactory colorization results with mediocre accuracy on sampled test data.

3 Future Enhancements

3.1 Pending Work

- **Training the Final Model:** The current model needs to be trained on the complete CelebA dataset we preprocessed to generate corresponding sketch input with paired images.
- **Improved Discriminator:** Future iterations will incorporate advanced discriminators for better image realism and feature learning.

3.2 Upcoming Features

- **Model Deployment:** Deploy the model on the cloud for scalable and accessible use by police departments globally.

4 Real-World Implementations

Built to be used by professional sketch artists who work in the domain of Security wherein this model can be used for better suspect identification.

5 Team Contributions

- **Vedant Marodkar:** Lead developer, model architecture design.
- **Rajeev Goel:** Secondary developer , training and testing.
- **Raj Mohammad:** Data Collection, augmentation and pre-processing.
- **Arpan Goswami:** Hyperparameter tuning and model evaluation .
- **Ashwin Jaiswal:** Hyperparameter tuning and metric analysis.