Project Report

Title: Sketch-to-Image generation

Reasearch Paper Details

Title: Unsupervised Sketch-to-Image Translation Network for Single Category Image Synthesis.

Publisher: IEEE

Published In: August, 2024

Summary: This paper presents an **unsupervised two-stage framework** for translating sketches into realistic images **without needing paired training data**. The goal is to synthesize high-quality and diverse images from abstract sketches, a task made difficult due to the lack of structured data and the domain gap between sketches (sparse lines) and real images (rich color and texture).

Introduction

Sketch-based image synthesis is a challenging task in the field of computer vision due to the large domain gap between sparse, abstract sketches and realistic colored images. In this project, we present a supervised sketch-to-image translation framework based on Conditional Generative Adversarial Networks (cGANs) that generates facial images conditioned on both input sketches and facial attributes.

Methodology

The proposed pipeline consists of:

-Generator Network

The generator is designed to take two inputs:

- A **1-channel sketch image** (grayscale).
- An **18-dimensional attribute vector** (e.g., gender, smile, eyeglasses).

The sketch and attribute data are processed through separate convolutional paths and later fused. A series of downsampling and upsampling layers are used to generate a realistic **3-channel face image (RGB, 128×128)**.

-Discriminator Network

The discriminator receives a face image (either real or generated) along with the corresponding attributes. It outputs a probability indicating whether the image is real or fake, using multiple convolutional layers and attribute fusion.

- Loss Function

We use **Binary Cross-Entropy Loss (BCELoss)** to train both the generator and discriminator in a typical adversarial setup. The generator is optimized to fool the discriminator, while the discriminator is trained to distinguish between real and fake images.

Data Preparation

- Sketches were loaded from a pre-processed pickle file containing sketch tensors.
- Real face images were extracted from the **CelebA dataset**, resized to 128×128, and normalized.
- Facial attributes were loaded from a CSV file and reduced to the most relevant 18 features.
- All data was converted to PyTorch tensors and loaded into a custom dataset class.

Training Process

- The model was trained using the Adam optimizer for **2 epochs** with a batch size of 8.
- During each iteration, the generator and discriminator were updated alternately.
- Checkpoints were saved after each epoch to preserve progress.

Inference & Output

After training, the model was tested with randomly generated sketches and attribute vectors. The output images were visually evaluated and saved using `torchvision.utils.save_image`.

Novelty

Make model compatible for design related sketch like designing some clothes, so that we can provide it the cloth outline(design) and it’ll automatically color it.