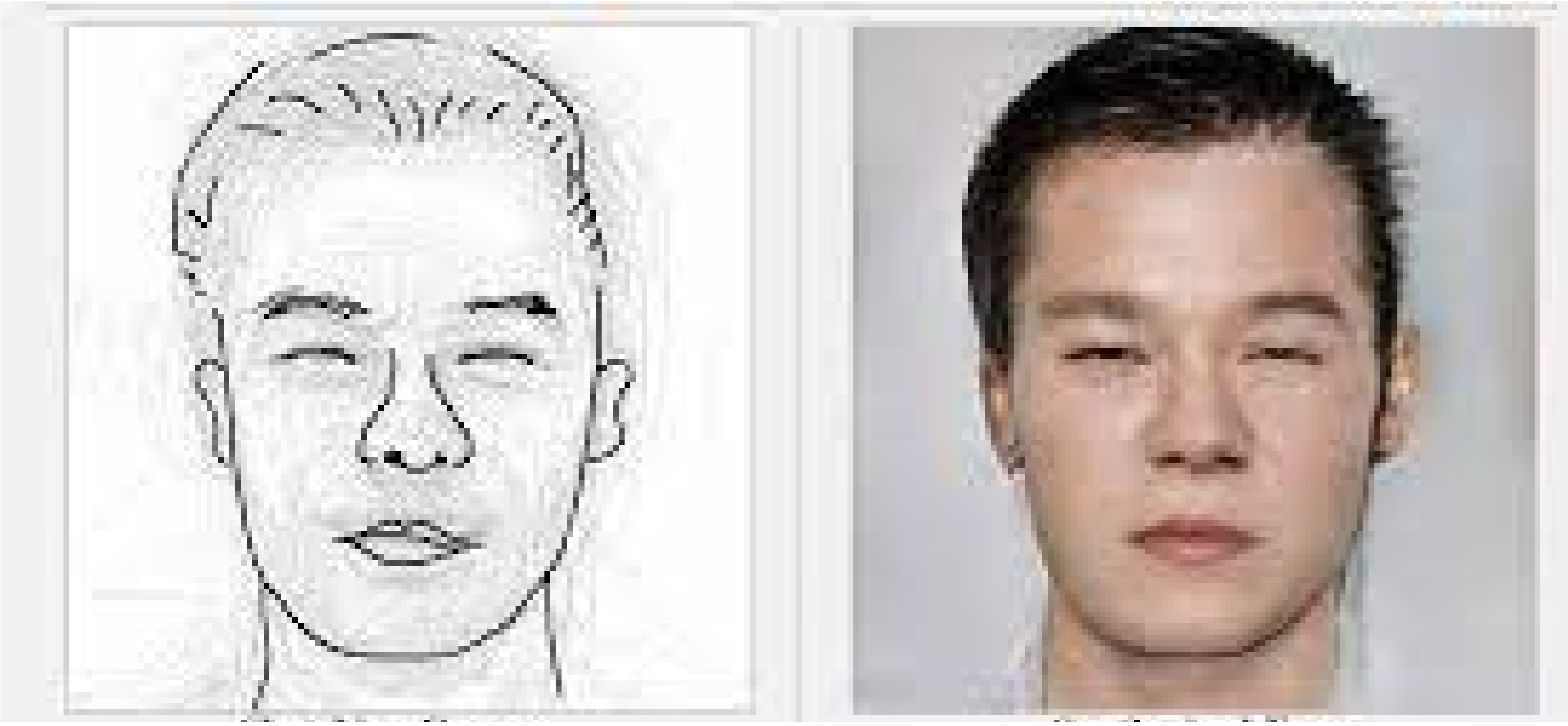
Group - 6



ReFaceIt

Team Members

Advika Lakshman
Aronnya Pal
Dhruv Bhagat
Pritha Mishra
Sushyanth Mohan
Vignesh RV

CORPORATE GURUKUL

Mentored by : Kanav Sabharwal
Faculty In-Charge : Dr. Amir Hussain

ReFaceIt - Sketch to 2D face reconstruction

- Developed a deep learning model that converts simple **hand-drawn sketches** into **realistic facial photos**.
- Trained on over **1,000 paired images of sketches** and their corresponding real photos.
- The model learns to add texture, shading, and color to enhance sketch outlines.
- Implemented using **Deep Convolutional GAN (DCGAN)** architecture and **trained from scratch.**

Problem Statement

- Facial sketches are often the only visual evidence in **law enforcement** when photos aren't available.

- Our system uses deep learning to convert sketches into realistic faces, aiding recognition and investigations.

Realistic reconstructions. Sharper identifications. Safer communities.

showing example of a sketch to the actual image of the person

Business Use Case

- Convert **eyewitness sketches** into **realistic faces** to improve suspect identification.
- Create **realistic faces** from sketches to help find **missing** persons when photos are unavailable.
- Use **AI** to help forensic experts recreate faces from **bones or incomplete descriptions.**

- Driven by the need for faster and more accurate suspect identification, this solution uses AI to convert forensic sketches into photorealistic images.
- It enhances facial recognition accuracy, shortens investigation time, and offers strong value for law enforcement, security agencies, and forensic tech providers.

CORPORATE GURUKUL

Data Overview & EDA

CORPORATE GURUKUL

Data Overview

- In this project, we explored datasets for training facial **sketch-to-photo synthesis** models.
- **The dataset which is considered:**

- **CUHK Dataset Overview:**

1. Contains hand-drawn photo-sketch pairs in a single realistic style , sourced from public datasets like CelebA, but lacks facial annotations or metadata.
2. Requires manual preprocessing and augmentations; evaluated using basic metrics like SSIM and L2 norm.
3. Due to its limited style and diversity, the dataset may not generalize well across different sketching techniques or facial conditions (like varying lighting or accessories).
4. Its smaller size and lack of structured metadata make it less suitable for training deep models that rely on multi-attribute supervision or stylistic adaptation.

EDA

EDA Aspect	CUHK Dataset
Image Distribution	Approximately 1K–2K paired images and sketches
Sketch Variety	Single consistent hand-drawn sketch style
Image Quality	Moderate resolution; well-aligned and usable
Facial Attributes	Not labeled or annotated
Visual Diversity	Consistent poses and lighting; limited variation
Augmentation Used	Flip, rotation, noise added to enhance training
Evaluation Metric	SSIM used to assess structural similarity
Sketch Complexity	Visual inspection confirms clear, clean sketches

CORPORATE
GURUKUL

Methodology

CORPORATE GURUKUL

1. **Data Preparation**
    - Used paired sketch and photo images from the **CUHK dataset**.
    - Preprocessing included resizing to **128×128** and normalizing image tensors.
    - Matched filenames to align sketch-photo pairs accurately.
    - Rebalanced the dataset by moving extra pairs from testing to training to ensure sufficient training data.

CORPORATE GURUKUL

2.**Generative Adversarial Networks (GANs)**

- **Adversarial Learning:** GANs consist of two networks—a generator that tries to produce realistic images, and a discriminator that learns to tell generated images apart from real ones.

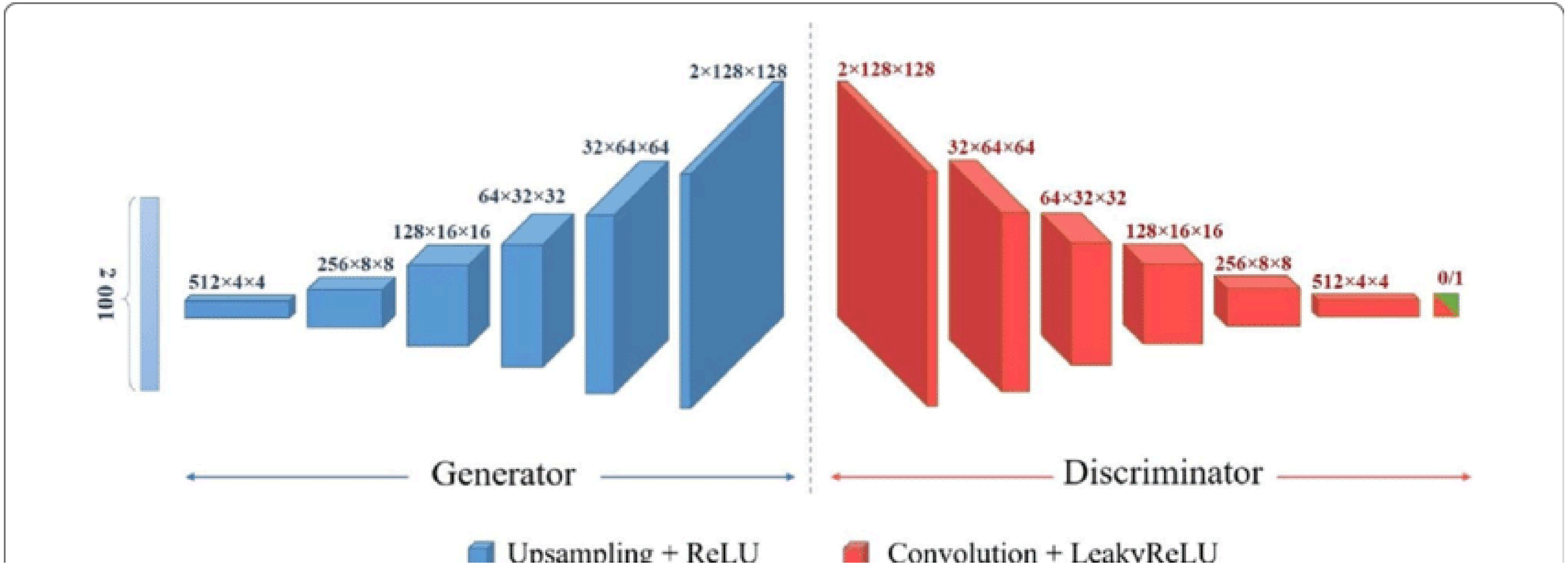
- **Mutual Improvement:** As training proceeds, the generator improves its outputs to fool the discriminator, while the discriminator sharpens its ability to spot fakes.

- **Stabilized Training:** Using mean-square adversarial loss (instead of the original logarithmic loss) helps avoid situations where the generator's gradients vanish too early.

3. **Deep Convolutional GAN (DCGAN) Design**

- **Downsampling (Discriminator):** Strided convolutions reduce image size while capturing deep features.
- **Upsampling (Generator):** Transposed convolutions reconstruct high-res images from feature maps.
- **Batch Normalization:** Stabilizes training by normalizing feature distributions.
- **LeakyReLU Activation:** Used in the discriminator to maintain gradient flow and avoid dead neurons.

CORPORATE GURUKUL



CORPORATE GURUKUL

4. **Encoder–Decoder Structure for Sketch→Photo**

- **Encoder:** Learns a compact representation of the input sketch by repeatedly halving spatial dimensions and increasing feature depth.

- **Decoder:** Reconstructs a full-resolution face image from that representation, striving to match both global structure and local detail.

5. **Patch-Level Discrimination**

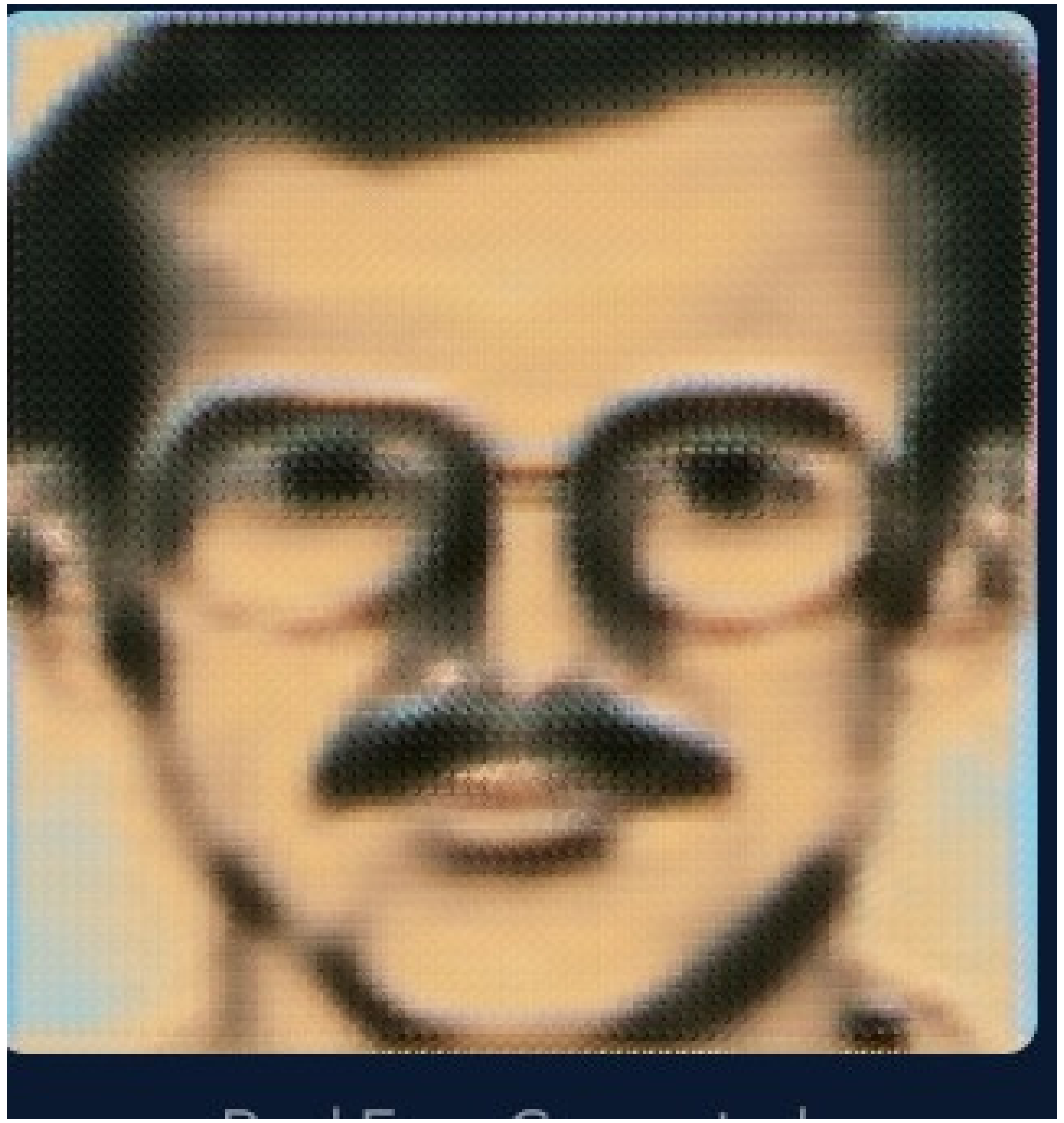
- Instead of judging an image as a whole, the discriminator evaluates many small overlapping patches.

- This **“PatchGAN”** approach focuses on local realism, encouraging sharper textures and more coherent fine details.

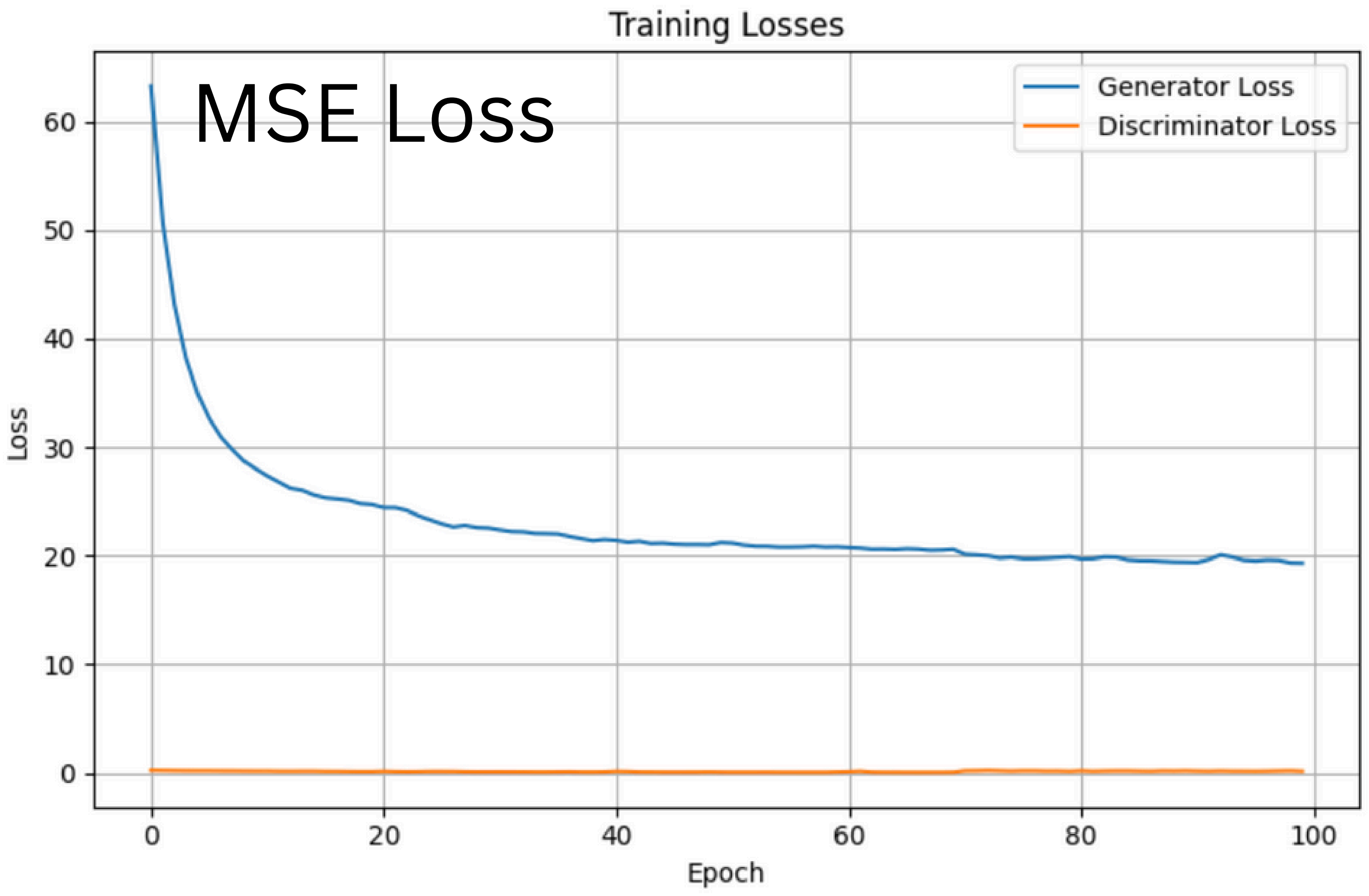
Result
&
Analysis

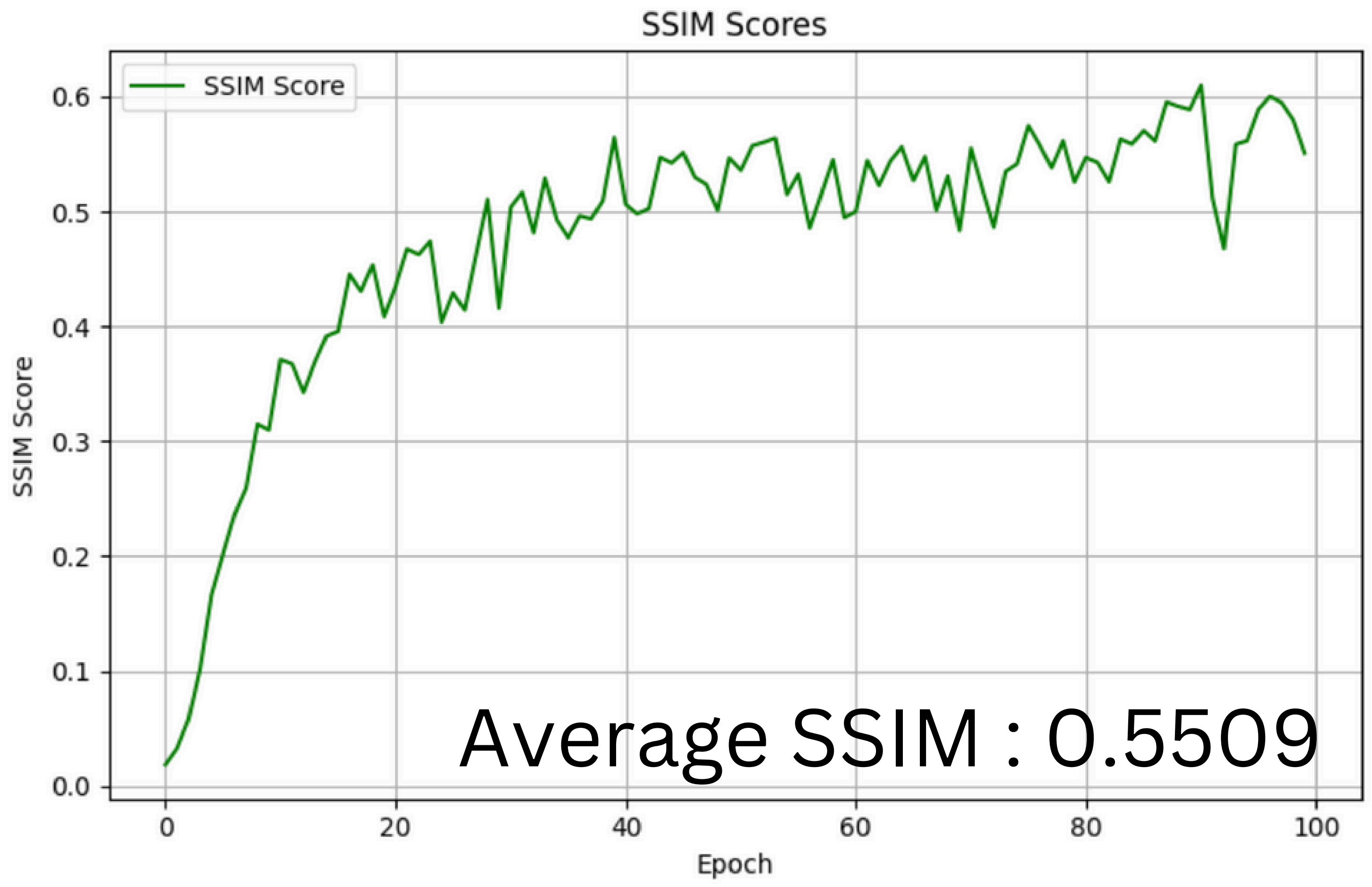
Real Face Generated

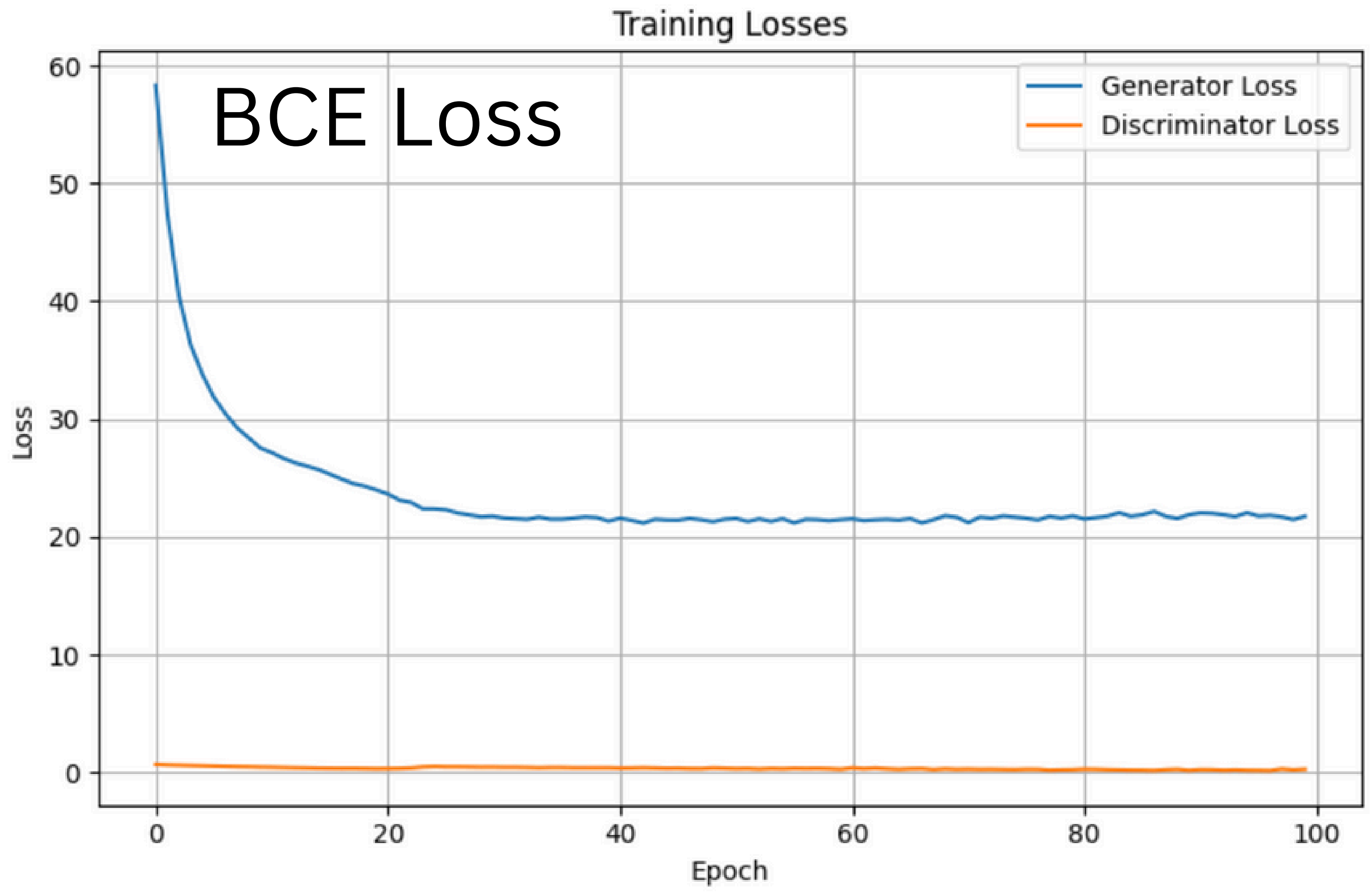
CORPORATE GURUKUL

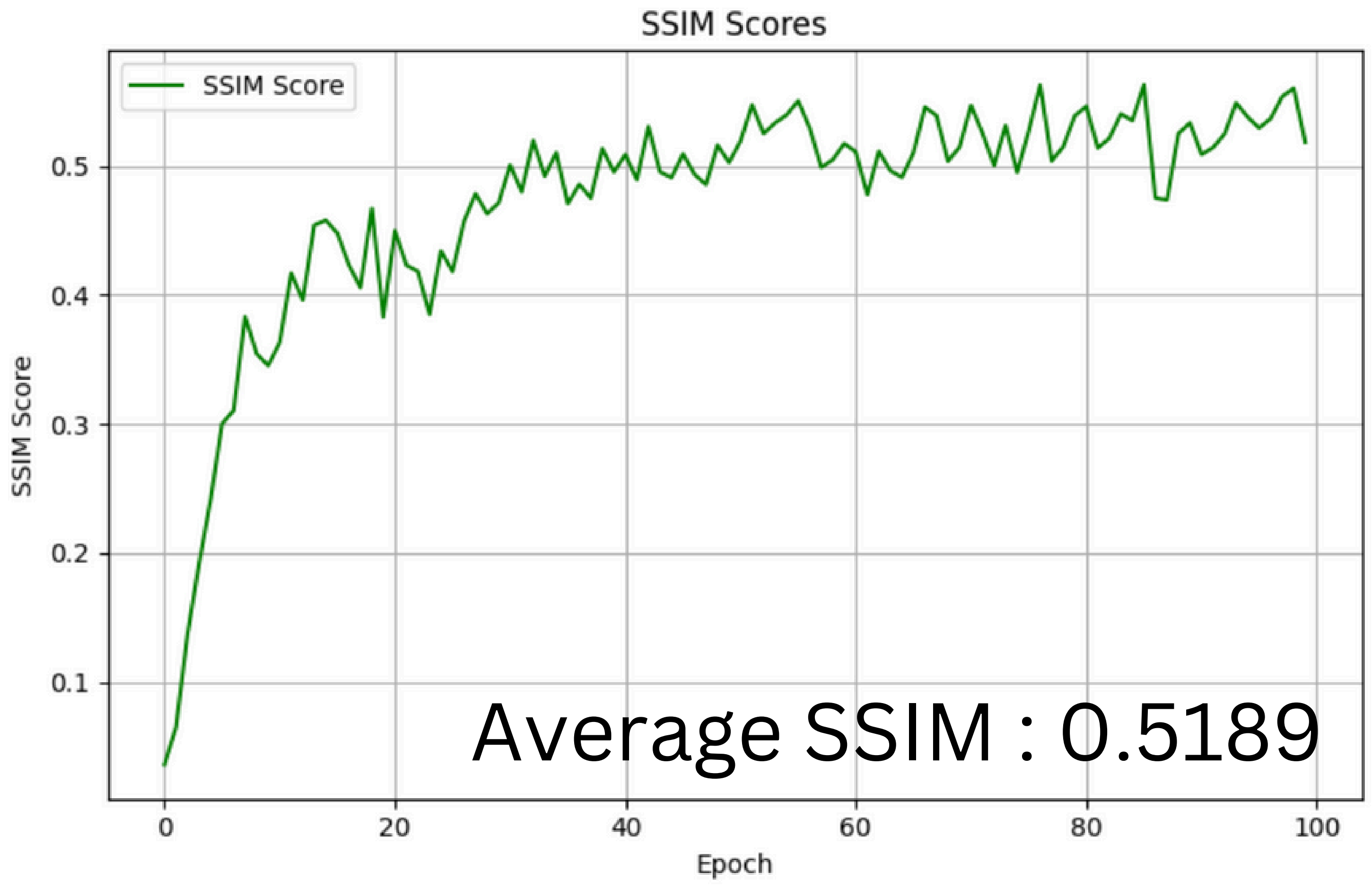


Loss Curves And SSIM Curve Over Epochs



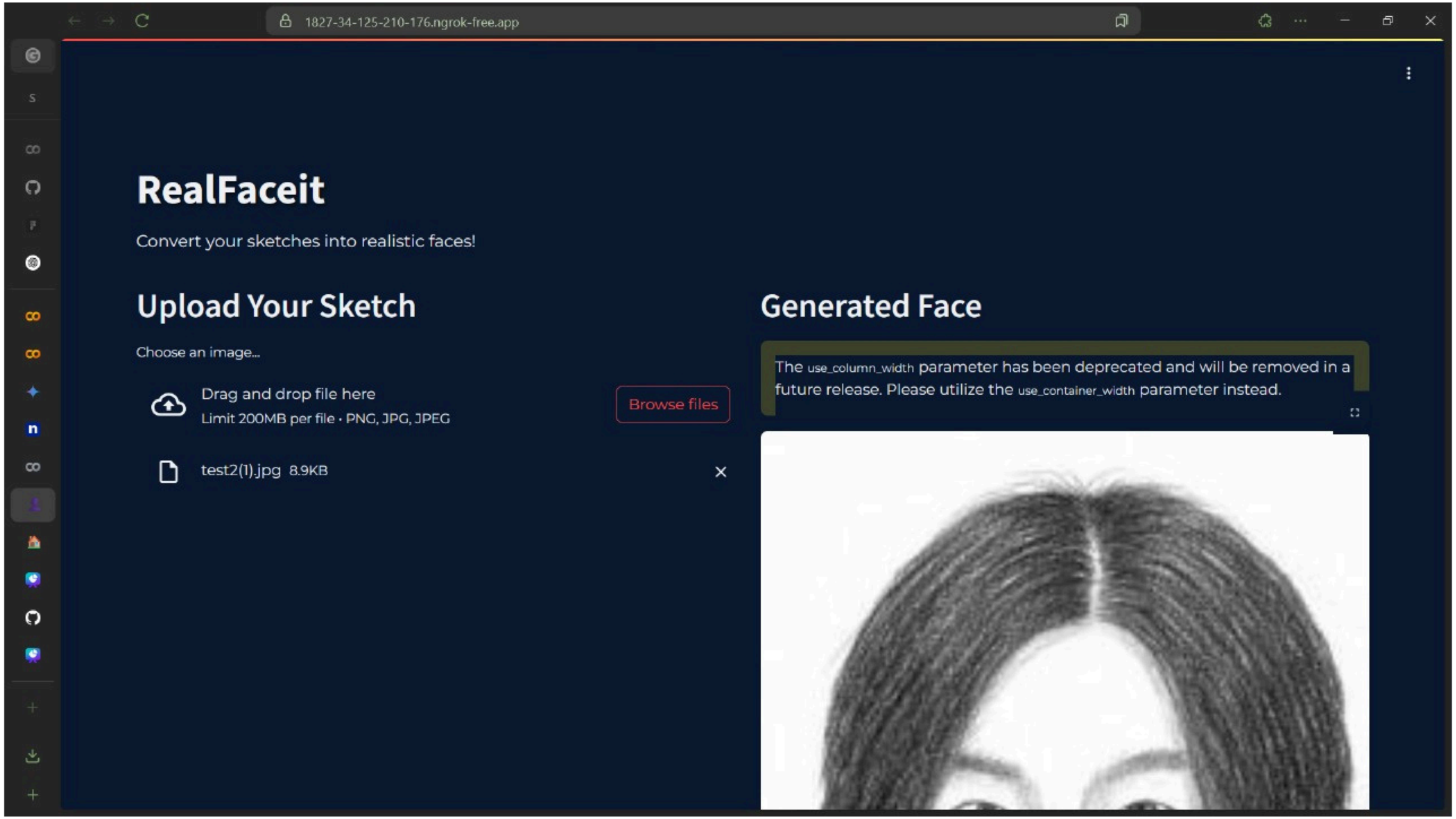






Demo
&
Challenges

CORPORATE GURUKUL



CORPORATE GURUKUL

Challenges

Sketch Variability:
Eyewitness sketches differ greatly in style and accuracy, making it hard for models to generalize across inputs.

Data Scarcity
A lack of large, high-quality paired sketch-to-photo datasets limits the training of robust deep learning models.

Computational Demands
Training and inference require high-performance GPUs and optimized architectures for real-time application.

CORPORATE GURUKUL

Future Scope

- **Multi-Modal Image Generation**

To enhance user experience and provide flexibility, we plan to extend the current model to generate multiple plausible face reconstructions from a single sketch

- **Side-Profile Reconstruction Capability**

The current system is primarily trained on frontal face images. Future work will involve extending the model to handle side-profile sketch-to-photo translation.

We have also fine-tuned a Stable Diffusion v1.5 model augmented with ControlNet (Canny) using LoRA to generate images (e.g., realistic faces) conditioned on structure-based inputs (sketches)

References

- https://github.com/Malikanhar/Face-Sketch-to-Image-Generation-using-GAN

- https://www.researchgate.net/publication/341631538_Face_Sketch_Recognition-An_Overview