Face Sketch to Image Generation using Hybrid GAN

Shantanu Kharade[1], Pranav Asane[2], Shubham Tapale[3], Neeraj Kalambe[4], Priti Malkhede[5]
[1, 2, 3, 4, 5]*Dept. of Artificial Intelligence & Data Science, PES's Modern College of Engineering, Pune, India.*

Abstract*— Creating lifelike facial images from hand-drawn sketches is a complex task with broad applications, ranging from artistic and design purposes to criminal investigations. This paper presents an innovative web application powered by a hybrid Generative Adversarial Network (GAN) model, combining the strengths of DCGAN and Cycle-GAN for superior image synthesis. Users can seamlessly upload sketches, generate high-quality images, and download results, supported by real-time processing feedback. The architecture, featuring an intuitive interface and advanced GAN training methods, is discussed alongside evaluation metrics for image fidelity. This approach opens new avenues for digital creativity and practical use in law enforcement.

Indexed Terms- Sketch-to-image generation, Generative Adversarial Networks, Hybrid GANs, Image synthesis, Web- based tools.

I. INTRODUCTION

Creating lifelike facial images from hand-drawn sketches has immense value in various domains, including portrait creation, character design, and forensic investigations. This technology enables artists to produce realistic visuals quickly and supports law enforcement in creating suspect profiles from witness-provided sketches. Traditional approaches like digital painting and photo editing demand significant time and expertise, limiting their practicality and adaptability. In contrast, modern advancements in artificial intelligence, particularly Generative Adversarial Networks (GANs), have introduced more efficient methods. GANs, which consist of a generator and a discriminator, have demonstrated exceptional capabilities in producing realistic imagery. This paper introduces a web-based application that employs a hybrid GAN model, merging the strengths of DCGAN and Cycle GAN, to convert face sketches into photorealistic images. Designed with ease of use in mind, the platform features an intuitive.

This method utilizes the combined strengths of DCGAN and Cycle GAN to enhance image generation capabilities. DCGAN is adept at producing high-quality visuals from random noise, while Cycle GAN specializes in domain translation without requiring paired datasets. The resulting hybrid model is versatile, capable of processing a wide range of hand-drawn sketches while preserving image accuracy and detail. Beyond its utility for artists and designers, this application offers significant value to law enforcement, providing a faster and more precise solution for generating facial images from rough sketches. By merging artificial intelligence with a user-friendly interface, the system seamlessly integrates advanced technology with creative workflows, unlocking new opportunities across artistic and practical domains.

II. LITERATURE REVIEW

[1] presented a study on the use of DCGAN for converting forensic sketches into real images. Their approach employs a generator and two discriminators to achieve high-resolution outputs. However, the model faces challenges with significant geometric deformations and textural alterations. Chao feng Chen et al. [2] introduced a semi-supervised Cycle-GAN that addresses the issue of small paired datasets and steganography in face photo-sketch translation. While effective, it still relies on a limited reference set and may not perform well on diverse or unseen data, with noise-injection strategies that might not completely mitigate overfitting. Heng Liu et al. [3] developed Sketch2Photo, which synthesizes photo-realistic images from sketches using Fast Fourier Convolution (FFC), Swin Transformer, and Improved Spatial Attention Pooling (ISAP). This method improves image quality by capturing both local and global features but faces challenges with the computational complexity of self-attention for large-size feature maps, as well as potential artifacts from misaligned or

incomplete sketches. Kaushal Rathore et al. [4] proposed an unsupervised domain adaptation technique for synthesizing face photos from sketches using adversarial networks without requiring paired training data. Although innovative, the method lacks sharpness in finer details, particularly for complex sketches. Guang can Liu et al. [5] presented a Conditional GAN based model that translates face sketches into photo-realistic images by conditioning the generator on input sketches. The approach struggles with large variations in facial features and complex images.

Yuki Tanaka et al. [6] utilized Least-Squares GAN (LSGAN) to convert rough, hand-drawn facial sketches into photo- realistic images, reducing gradient instability for better synthesis. However, the model encounters issues with misaligned sketches and potential distortions in facial features. Jin Han Lee et al. [7] introduced a Cycle GAN model with multi-scale discriminators to enhance the translation of facial sketches to realistic images. This approach faces challenges with fine-grained details and significant pose variations. Lingzhi Zhang et al. [8] proposed a Conditional GAN-based approach for generating photo-realistic images from sketches by learning the conditional dependencies between face sketches and photos. However, performance diminishes with abstract or incomplete sketches. Eiji Yonekura et al. [9] focused on generating photo-realistic face images from sketches using StyleGAN, employing a pre-trained encoder to map sketch features into StyleGAN's latent space. The model struggles with incomplete or poorly drawn sketches and extreme variations. Jing Zhang et al. [10] utilized a Dual GAN framework that enables two GAN

Rachel Johnson et al. [11] proposed a cross-domain GAN architecture for generating face photos from sketches, effectively handling differences in representation across various domains. Nonetheless, performance deteriorates with highly incomplete or distorted sketches. Lin Wang et al. [12] presented a cascaded GAN framework to enhance the quality of generated images from sketches through sequential refinement processes. The model may require extensive training data and faces challenges with complex sketches. Zhang et al. [13] developed a Multi-Scale Generative Adversarial Network (MS-GAN) for face sketch to photo synthesis, leveraging multiple scales of input to capture both global and local features effectively. This method enhances image realism but may struggle with high-resolution outputs due to increased computational demands. Lee et al. [14] explored the application of a Progressive Growing GAN (PGGAN) for generating realistic face images from sketches. By progressively increasing the resolution during training, the model produces high-quality images. However, it requires a substantial amount of data and training time, which can be a limitation for smaller datasets. Kim et al. [15] proposed an Attention-Guided GAN (AGGAN) that employs attention mechanisms to improve the focus on important regions in sketches, enhancing the quality of generated images. While effective, this method may introduce artifacts if the attention mechanism misaligns with the actual sketch features.

Patel et al. [16] introduced a Dual-Path GAN (DPGAN) for translating sketches into photos, utilizing two distinct pathways to capture both the high-level semantics and low- level details. Despite its innovative architecture, the model can encounter difficulties with highly complex sketches and may not generalize well to unseen data. Singh et al. [17] presented a Hybrid GAN model that integrates features from both CycleGAN and Pix2Pix, aiming to improve sketch-to-image synthesis by leveraging paired and unpaired data. However, the model may still face challenges with generating realistic images from rough sketches due to the inherent variability in input quality. Wang et al. [18] focused on a Semantic-Aware GAN (SAGAN) that incorporates semantic segmentation maps to guide the image generation process. By providing additional contextual information, the model improves output quality. Nonetheless, it may still struggle with occlusions and missing details in the original sketches.

III. PROPOSED SYSTEM

A. Introduction

The task of creating realistic facial images from hand-drawn sketches has a longstanding history, blending elements of art and computer science. This system seeks to harness advanced technologies to transform the simplicity of sketches into detailed, photorealistic images. Traditional methods have provided a solid

foundation for contemporary techniques, which can be grouped into several distinct strategies.:

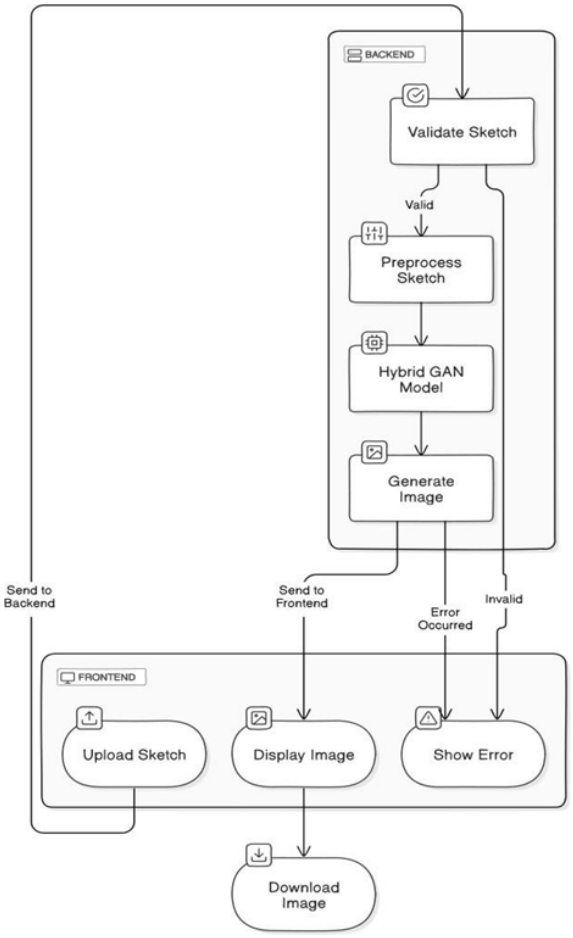

Fig. 1. Proposed System

B. Key Components

1. User Interface (UI):
A user-friendly interface that allows users to upload hand- drawn sketches and view the generated face images. The UI also provides options for users to refine their sketches and visualize different generated outputs.
2. Sketch Input Processing:
This component handles the preprocessing of the input sketches, including resizing, normalization, and noise reduction, to ensure that they are suitable for analysis by the GAN model.
3. Generator Network (DCGAN):
The core component of the system that generates realistic face images from random noise and the processed sketch inputs. The DCGAN is trained on a dataset of real face images to learn the complex patterns and features of human faces.
4. Discriminator Network (Cycle GAN):
This component evaluates the authenticity of the generated images, ensuring they are indistinguishable from real images. It helps improve the quality of the generated outputs by providing feedback to the generator.
5. Admin Dashboard:
A control panel that displays statistics on user interactions, generated outputs, and system performance. It allows for tracking user engagement and refining the model based on feedback and usage patterns.

C. Key Features
1. High-Quality Image Generation:
The system utilizes advanced GAN architectures to produce high-quality face images that closely resemble real human faces.

2. Sketch Refinement Options:
Users can modify their sketches interactively, allowing them to see how changes impact the generated images in real-time.

3. Performance Analytics:
The admin dashboard provides insights into user participation, the number of images generated, and overall system effectiveness, helping improve the user experience.

This system leverages state-of-the-art deep learning techniques to convert sketches into realistic images, enhancing applications in digital art, law enforcement, and entertainment. The integration of user-friendly features and comprehensive analytics ensures a smooth experience for both users and administrators.

CONCLUSION

This paper introduces a novel web application that leverages a hybrid Generative Adversarial Network (GAN) architecture to synthesize photorealistic images from hand- drawn facial sketches. The proposed model integrates the strengths of DCGAN and Cycle GAN to enhance image generation quality. The user-friendly interface facilitates intuitive sketch

uploads and image generation, promoting accessibility for users with varying levels of technical expertise. This paper investigates the application's functionalities, including real-time processing feedback and efficient image download mechanisms. Furthermore, it delves into the intricacies of the hybrid GAN architecture, encompassing the training process and the evaluation metrics employed to assess the quality of the synthesized images. This innovative approach has the potential to significantly impact various domains, including digital art, law enforcement, and entertainment, by enhancing the expressive power of hand-drawn sketches.

REFERENCES

[1] S. Devakumar and G. Sarath, "Forensic Sketch to Real Image Using DCGAN," International Conference on Machine Learning and Data Engineering, 2023. Available: www.sciencedirect.com.

[2] H. Liu, Y. Xu, and F. Chen, "Sketch2Photo: Synthesizing Photo-Realistic Images from Sketches via Global Contexts," in Proceedings of the International Conference on Machine Learning and Data Engineering, 2023.

[3] C. Chen, W. Liu, X. Tan, and K. Y. K. Wong, "Semi- Supervised Cycle-GAN for Face Photo-Sketch Translation in the Wild," in Proceedings of the International Conference on Machine Learning and Data Engineering, 2023.

[4] K. Rathore, M. Narang, and N. S. Karnik, "Unsupervised Domain Adaptation for Face Sketch-Photo Synthesis Using Adversarial Networks," in Proceedings of the International Conference on Machine Learning and Data Engineering, 2023.

[5] Shikang Yu, et al., "Improving Face Sketch Recognition via Adversarial Sketch-Photo Transformation," in Proceedings of the 2019 14th IEEE International Conference on Automatic Face Gesture Recognition (FG 2019), 2019.

[6] Y. Liu, et al., "Domain-Adaptive Generative Adversarial Networks for Sketch-to-Photo Inversion," in Proceedings of the 2017 IEEE 27th International Workshop on Machine Learning for Signal Processing (MLSP), 2017.

[7] C. Galea and R. A. Farrugia, "Forensic Face Photo- Sketch Recognition Using a Deep Learning-Based Architecture," IEEE Signal Processing Letters, vol. 24, no. 11, pp. 1586–1590, 2017.

[8] Y. Lin, et al., "An Identity-Preserved Model for Face Sketch-Photo Synthesis," IEEE Signal Processing Letters, vol. 27, pp. 1095–1099, 2020.

[9] R. Yeshaswy, et al., "Design and Analysis of a 6 Watt GaN-Based X-band Power Amplifier," in Proceedings of the 2016 International Conference on Communication and Signal Processing (ICCSP), 2016.

[10] Q. Li, et al., "AF-DCGAN: Amplitude Feature Deep Convolutional GAN for Fingerprint Construction in Indoor Localization Systems," IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 5, no. 3, pp. 468–480, 2019.

[11] E. Schonfeld, B. Schiele, and A. Khoreva, "A U-Net Based Discriminator for Generative Adversarial Networks," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020.

[12] A. Tohokantche Gnanha, et al., "The Residual Generator: An Improved Divergence Minimization Framework for GAN," Pattern Recognition, vol. 121, p. 108222, 2022.

[13] U. Demir and G. Unal, "Patch-Based Image Inpainting with Generative Adversarial Networks," arXiv preprint arXiv:1803.07422, 2018.

[14] S. S. Channappayya, A. C. Bovik, and R. W. Heath, "Rate Bounds on SSIM Index of Quantized Images," IEEE Transactions on Image Processing, vol. 17, no. 9, pp. 1624–1639, 2008.

[15] B. Klare, Z. H. Li, and A. K. Jain, "Matching Forensic Sketches to Mug Shot Photos," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 3, pp. 639–646, 2010.

[16] V. M. Mohan and V. Menon, "Measuring Viscosity of Fluids: A Deep Learning Approach Using a CNN-RNN Architecture," in Proceedings of the First International Conference on AI-ML-Systems, 2021.

[17] V. Mohan, “Detection of COVID-19 from Chest X-Ray Images: A Deep Learning Approach,” in Proceedings of the 2021 Ethics and Explainability for Responsible Data Science (EE-RDS), IEEE, 2021.

[18] N. Aloysius and M. Geetha, “A Review on Deep Convolutional Neural Networks,” in Proceedings of the 2017 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2017.

[19] V. Balachandran and S. Sarath,” A Novel Approach to Detect Unmanned Aerial Vehicle using Pix2Pix Generative Adversarial Network,” 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), 2022, pp. 1368-1373, Doi: 10.1109/ICAIS53314.2022.9742902.