IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 3, March 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104159 |

Suspect Criminal Face Generation

R. C. Suryawanshi, Gaurav Karkhile, Rohit Kalase, Tejas Karanjavkar

Department of Computer Engineering, A C Patil College of Engineering, Kharghar, Navi Mumbai, India

ABSTRACT: In today's society, there has been a rise in the number of crimes committed, and sketch artists have been hired by law enforcement agencies to create sketches of suspects based on eyewitness accounts. However, these sketches are often inaccurate due to the limitations of the sketch artist or the eyewitness. To address this issue, Generative Adversarial Networks (GANs) have been used to train a neural network to generate pictures that fit a specific category. Nonetheless, GANs are unable to produce high-resolution images, and StyleGAN has been implemented to improve the quality of the generated images. To further refine the generated images to match the suspect's description, TL-GAN is used. Our project seeks to create a Suspect Face Generation System that enables users to create accurate sketches of suspects without the need for a forensic sketch artist or special training. The accuracy of sketches produced by sketch artists is generally low, with only around 8% of sketches being accurate. Our system can help law enforcement reduce the risk of misidentifying suspects and contribute to a safer society with reduced crime rates.

KEYWORDS: GAN, TL-GAN, CNN, RES-NET

I.INTRODUCTION

The rising crime rate in modern times has made it essential to identify criminals using face sketches generated from eyewitness information. However, the accuracy of sketches produced by sketch artists is often inadequate, resulting in misidentification and potentially causing problems for innocent people in society. Additionally, the current systems can only match faces registered in the criminal database, making it challenging to create sketches for new faces.
To overcome these challenges, our system will generate high-quality face sketches based on the initial eyewitness description. Users can further refine and modify the generated images according to their needs and preferences using the system's options, improving the accuracy of the sketches while reducing the need for human intervention. Our system aims to produce better face sketches that will contribute to reducing misidentification and promoting a safer society.

II.LITERATURE SURVEY

2.1 Survey of Existing System/SRS

I. Forensic Face Sketch Artist System

The Forensic Face Sketch Artist System is a forensic application that allows users to construct accurate composite face sketches using predefined facial feature sets provided as tools that can be resized and repositioned as required. The system has been developed and tested keeping real-world scenarios in mind, ensuring security, privacy, and accuracy as the key factors in every scenario. The platform has demonstrated good accuracy and speed while face sketch construction and recognition processes. [2]

II. A Realistic Image Generation of Face from Text Description using the Fully Trained Generative Adversarial Networks

This system works in two sections. The first section describes how to encode text into semantic vectors, while the second section explains how to decode the semantic features of text into realistic natural images. The network architecture is designed based on two streams, with text being encoded in the first portion and images being decoded using the encoded text embedding in the second part. The generator contains the trainable encoder and decoder and serves as the backbone of the proposed architecture for the text-to-image task. [1]

III. Towards Criminal Sketching with Generative Adversarial Network

In this work, They made a first step to criminal sketching with a novel GAN based architecture, allowing user to synthesize a real-like portrait of a criminal suspect with a description vector. The generation procedure into two steps

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104159 |

each achieved by using a GAN: Sketch Prediction GAN and Portrait Prediction GAN. Given the description vector of a face, the sketch generation GAN will predict an outline, containing rough information of image has been predicted, the portrait generation GAN tries to complete details and colors on that sketch image by reading the description again. [3]

Paper Name	Author	Limitations
Forensic Face Sketch Artist System	Srujan Mahajan, Vipul Humbe, Advait Raorane, Asmita Deshmukh	❖ In this application predefined face,eyes,hairs types are given. ❖ It gives black and white Output image so some time it will not match with criminal.
A Realistic Image Generation of Face from Text Description using the Fully Trained Generative Adversarial Networks	Muhammad Zeeshan Khan , Saira Jabeen , Muhammad Usman Ghani Khan , Tanzila Saba, Asim Rehmat, Amjad Rehman , Usman Tariq	❖ It is only Desktop application We cannot use it online ❖ In this system text description is given as input. ❖ Image quality generated by image is low
Towards Criminal Sketching with Generative Adversarial Network	Hanzhou Wu1 , Yuwei Yao , Xinpeng Zhang and Jiangfeng Wang1	❖ This system does not provide option for edit the image.

Survey of existing system

2.2 Limitations of Existing System or Research Gap:

The Forensic Face Sketch Artist System has limitations such as predefined facial features that limit the accuracy of the generated sketches, and the generated sketches being in black and white, which may not match the actual criminal. The second paper, A Realistic Image Generation of Face from Text Description using the Fully Trained Generative Adversarial Networks, is a desktop application with low-quality image generation. The third paper, Towards Criminal Sketching with Generative Adversarial Network, lacks an edit option after image generation.

2.3 Mini Project Contribution:

In our system, verbal description is taken as input, and the system is web-based, allowing for online use on any application at any time. Our system generates faces without predefined facial features, and users are not required to do any manual work. The sketches generated by our system are not in black and white, making identification easier, and the quality of the generated sketches is high.

III. PROPOSED SYSTEM

3.1 Introduction:

This system aims to generate high-quality images based on the initial verbal description provided by users. The system will also allow for further alterations based on user requirements. Additionally, the system will check the generated images against a criminal database to determine if the criminal is already in the system.

The system uses a ResNet-50 CNN model to extract features from the generated images and map them to the desired image description. A new encoder-decoder algorithm is used to convert the description into a latent vector. The initial face image is generated using the StyleGan model, and the required modifications are made using the TL-GAN model. By reducing human intervention and utilizing machines, this system aims to generate better output images and reduce the time required for face generation. The system will also strive to provide high-resolution images using a regression model.

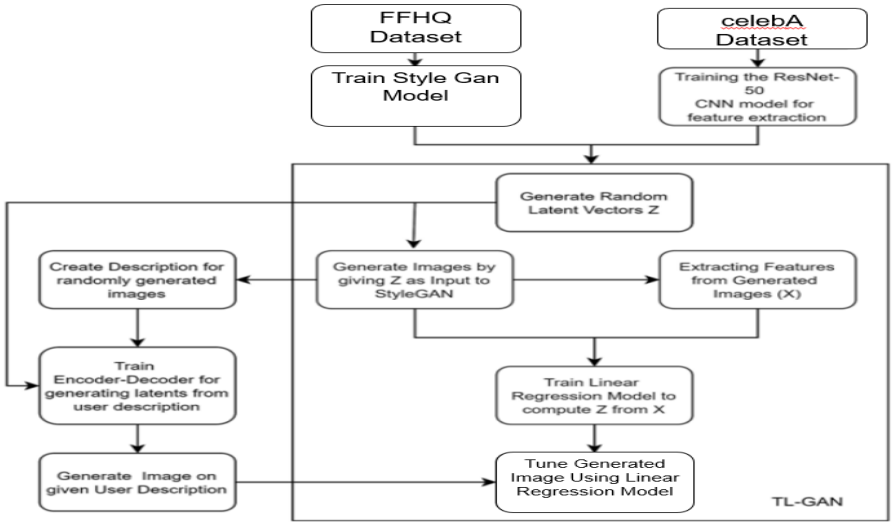
International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104159 |

3.2 Architecture and Framework



1...System Architecture

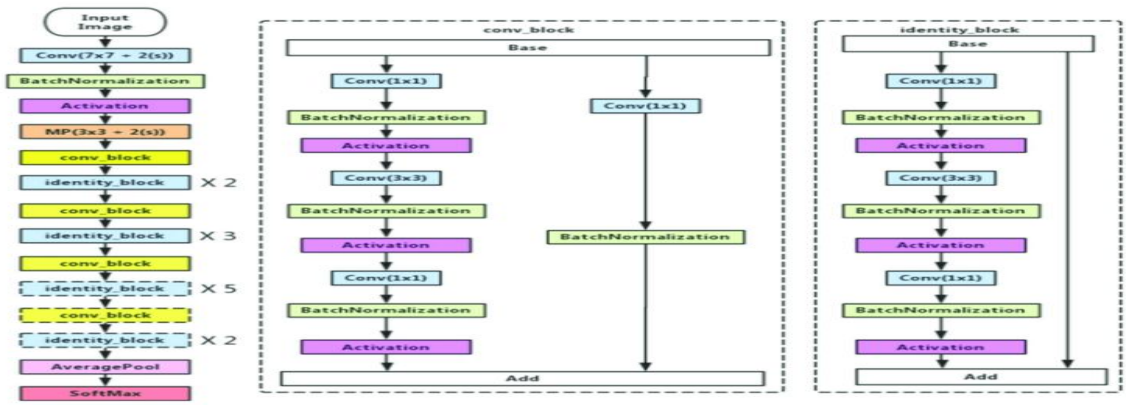
3.3 Process Design

A new latent space untangling method has been proposed that uses an encoder-decoder architecture that allows natural language input. This input is provided by an eyewitness and describes the suspect. The description is then converted to a set of latent representations using the Encoder-Decoder architecture. These latent fingerprints are then fed into StyleGAN, which generates an initial image. To refine the generated image, regression is used on the sampled latencies of the generated images and their feature scores, which are generated using CNN Feature Extractor (Resnet-50 CNN). This method is known as Transparent Latent-Space GAN (TL-GAN) and is used to tune faces holistically to create a more accurate image of a suspect.

3.4 Details of the Architectures

I. ResNet-50 CNN Architecture:

The ResNet-50 CNN model architecture includes a convolution block and an Identity block. Each block consists of three convolution layers, with the main difference being their skip connections. The convolution block includes a convolution layer in its skip connection, while the Identity block carries the output of previous layers directly to the next layer through its skip connection. The purpose of these skip connections is to address the vanishing gradient problem and allow the higher layers to perform as well as the lower layers by letting the network learn the identity function. The ResNet-50 CNN model is formed by repeating these blocks.



2.ResNet-50 CNN Architecture

International Journal of Innovative Research in Computer and Communication Engineering

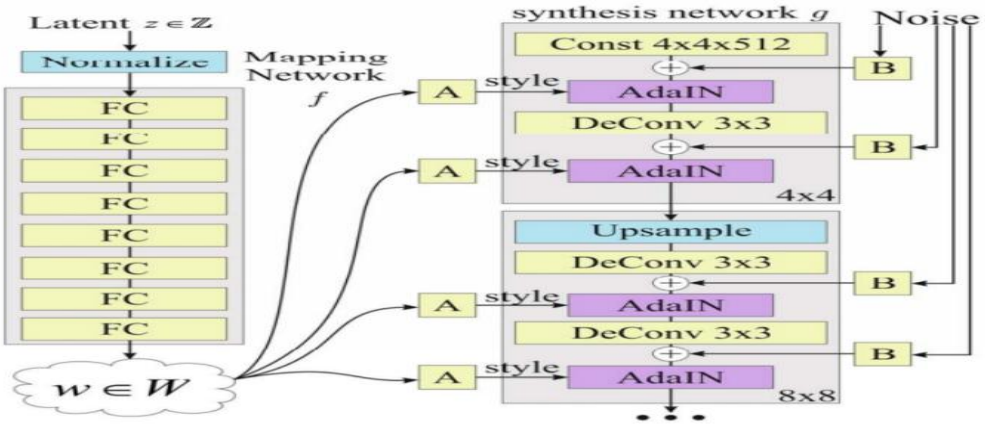
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104159 |

II. StyleGAN:

A standard GAN utilizes random noise, which is a vector of n floats, as its input and transforms this noise into a meaningful output that resembles the dataset it was trained on. The GAN consists of two neural networks: a generator and a discriminator. The generator generates fake samples of data, such as images or sounds, and attempts to deceive the discriminator. The discriminator, on the other hand, tries to distinguish between real samples from the dataset and fake samples generated by the generator. Through an adversarial process where they compete against each other, the generator and discriminator are trained. This process is repeated multiple times, resulting in improved performance for the generator and discriminator after each repetition. [4]



3.ResNet-50 CNN Architecture

IV.DESCRIPTION TO LATENTS CONVERTOR (ENCODER-DECODER)

A novel Encoder-Decoder Architecture has been proposed for converting a natural language description of a face into a latent vector z. This latent vector is then used as input for StyleGAN to generate an initial image(s), which can be modified using the TL-GAN algorithm. The encoder architecture consists of RNN units with long short-term memory that reads and encodes the facial description provided by the eyewitness. This encoded description is passed to the decoder, which consists of fully connected layers that learn to produce latents capable of generating a face that matches the eyewitness description. To train this architecture, a set of random latent vectors z is sampled, which are used to generate images using StyleGAN. These generated images are then tagged with a face description. By using descriptions as input to the encoder and latent vectors z as output, the encoder-decoder network can be trained to identify latents capable of generating a face that matches the eyewitness description.

V. TRANSPARENT LATENT-SPACE GAN (TL-GAN)

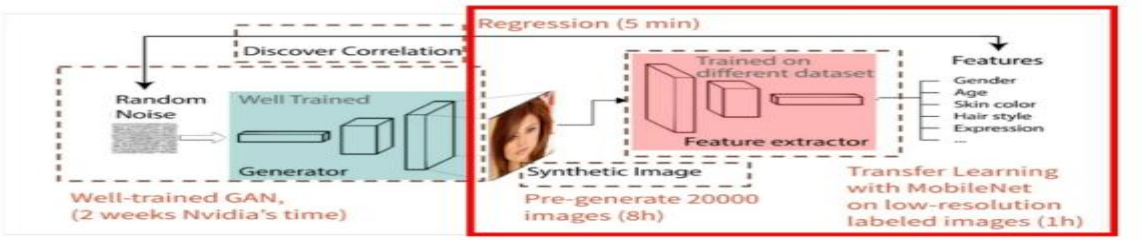
TL-GAN is a supervised learning model that generates images by taking fake inputs from the generator and real images from the domain or a higher resolution version provided by the discriminator. To compare the predicted output from the TL-GAN model with the actual output, we need to evaluate the model's performance using various metrics. One commonly used metric is the mean squared error (MSE), which measures the average squared difference between the predicted output and the actual output. Another metric is the peak signal-to-noise ratio (PSNR), which measures the ratio between the maximum possible value of the image and the noise present in the image. A higher PSNR indicates a more accurate model output. Additionally, the Structural Similarity Index (SSIM) can be used to evaluate the quality of the generated images by measuring the structural similarity between the predicted and actual images. To achieve accurate output, the TL-GAN model should be trained using a large and diverse dataset and the hyperparameters of the model should be fine-tuned to optimize its performance. [5]

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104159 |



1h + (generating 8h) : extract image label

TL-GAN

Requirement Analysis

Hardware Requirements:

- Laptop / Desktop
- GPU: MINIMUM 16GB
- Processor: i5 or Higher Processors
- Internet Connection

Software Requirements:

- Operating System: Windows 7,8,10 or higher
- Languages: Python
- Platform: VS code
- Web Browser: Chrome

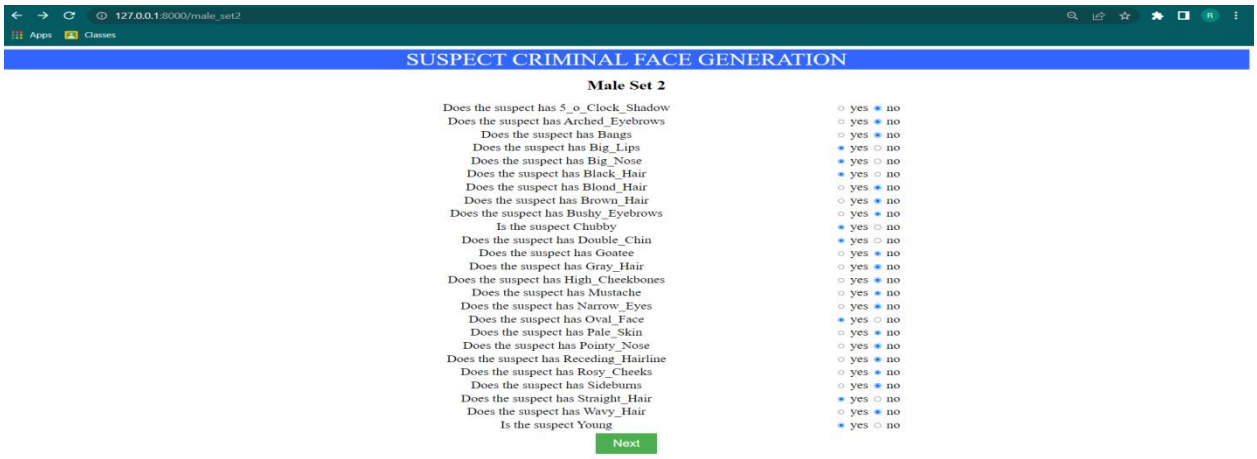
VI.RESULTS

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 4, April 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1104159 |



VII.CONCLUSION

This system will reduce human intervention and generate better output images. Also it will reduce the time consumption of making a face by the artist as machines are considered to be faster than the human being. The accuracy of the system to provide an initial face is 50%. After that the face can be edited to the desired attribute to obtain an accuracy of up to 75%. The total time taken to train the StypeGan and TL-GAN model is about 14 days using a 16GB GPU.

REFERENCES

[1] Muhammad Zeeshan Khan, Saira Jabeen, Muhammad Usman Ghani Khan, Tanzila Saba, Asim Rehmat, Amjad Rehman, and Usman Tariq. A realistic image generation of face from text description using the fully trained generative adversarial networks. IEEE Access, 9:1250–1260, 2020.

[2] SRUJAN MAHAJAN, ADVAIT RAORANE, VIPUL HUMBE, and ASMITA DESHMUKH. Forensic face sketch artist system.

[3] Hanzhou Wu, Yuwei Yao, Xinpeng Zhang, and Jiangfeng Wang. Towards criminal sketching with generative adversarial networks. In 2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP), pages 1–5. IEEE, 2020.

[4] Karras, T., Laine, S., & Aila, T. (2018, December 12). A style-based generator architecture for generative adversarial networks. arXiv.Org. https://arxiv.org/abs/1812.04948

[5] Kim, Dongyoung, et al. “TLGAN: Document Text Localization Using Generative Adversarial Nets.” arXiv.Org, 22 Oct. 2020, https://arxiv.org/abs/2010.11547.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details