**Forensic Sketch to Real Image Generation using GAN**

Atharva Navani, Om Lahore, Jatin Sathe, Atharva Joshi, Omkar Rode

*Department of Computer Science and Engineering (Artificial Intelligence), Vishwakarma Institute of Information Technology, Pune, Maharashtra*

**Table of Contents**

1. [ABSTRACT](#_gjdgxs)
2. [INTRODUCTION](#_30j0zll)
3. LITERATURE SURVEY
4. [PROPOSED METHOD](#_1fob9te)
5. [DATASET](#_3znysh7)
6. RESULTS
7. [STIMULATION DETAILS](#_2et92p0)
8. [REAL WORLD APPLICATIONS:](#_tyjcwt)
9. [ETHICAL CONSIDERATIONS:](#_3dy6vkm)
10. [CONCLUSION](#_1t3h5sf)
11. [REFERENCES](#_2s8eyo1)
12. [ACKNOWLEDGMENT](#_3rdcrjn)

Line

# ABSTRACT

This study investigates the application of artificial neural network to generate real-time images from drawings. It is important to transform pencil drawings into real images in determining the identity of people in forensic investigations. When describing a crime, the witness describes the defendant by recalling his physical characteristics. Because people's faces are diverse and different, it would be difficult to predict one from a rough sketch prepared by a forensic expert. In this article, we present a deep learning model that uses conditional GANs to transform artificial images into real images. The prepared images are used as input for training, and after some training, our deep learning model learns and creates real images of the suspect. Facial recognition is a powerful tool that can be used for many purposes, from security to entertainment. The research aims to create good images from images using cGAN, which has a generator and a separate algorithm. This work includes a comprehensive review of the capabilities of cGANs in data collection, preprocessing, modeling, training, and application in image processing tasks. In addition, the limitations, caveats and future expectations of this technology are also mentioned. Based on the analysis, this research contributes to the understanding and application of cGAN in the field of computer vision and image processing.

*Keywords: Conditional Generative Adversarial Networks (cGANs), image generation, sketches, deep learning, forensic investigations, facial recognition, model training, data preprocessing, model architecture, training methodologies, image synthesis, computer vision, image processing.*

Line

# INTRODUCTION

Due to the complexity of visual perception in different physical environments, traditional image processing techniques such as computer vision (CV) face challenges in digital image processing. However, advances in machine learning have paved the way for advanced image processing capabilities by combining machine learning techniques with art techniques. Facial recognition as artificial intelligence has become a technology that can keep up with the rapidly advancing technology of the digital world. It supports the investigation of crimes by using neural networks to recognize faces and create sketchy images that are then matched with existing data.

The use of generative adversarial networks (GANs) in facial recognition has revolutionized the process of generating real images. Images in Sketch. The technique involves the use of deep neural networks, specifically using GANs, to transform images into synthetic images. The generator component of GAN initially produces real images based on graphs using noise. These resulting images are evaluated by a discriminator that distinguishes between real and fake images based on the original input data.

Compared with existing photo synthesis techniques, the proposed method is the most effective in improving the accuracy of forensic sketches. A real picture. Techniques based on Generative Adversarial Networks (GAN) continue to impart unique properties to images, improving overall image quality. Although current sketch-to-image conversion methods are mainly based on convolutional neural networks (CNN) and GANs, the quality of the generated images can be further improved with new image processing algorithms.

This article uses, a generator and two splitters to convert images into images. The splitter evaluates the loss of the resulting image against the original data, and the second splitter is started after a certain period of time. Experimental results show that the effectiveness of the model and the recognition rate are better than existing models, with the average of similar models .

The next section of this article provides an overview of graph-to-image conversion and GANs, followed by a detailed description of the methodology and design. Experimental and analytical results are presented to reveal the results and significance of the study.

Line

# Line

# LITERATURE SURVEY:

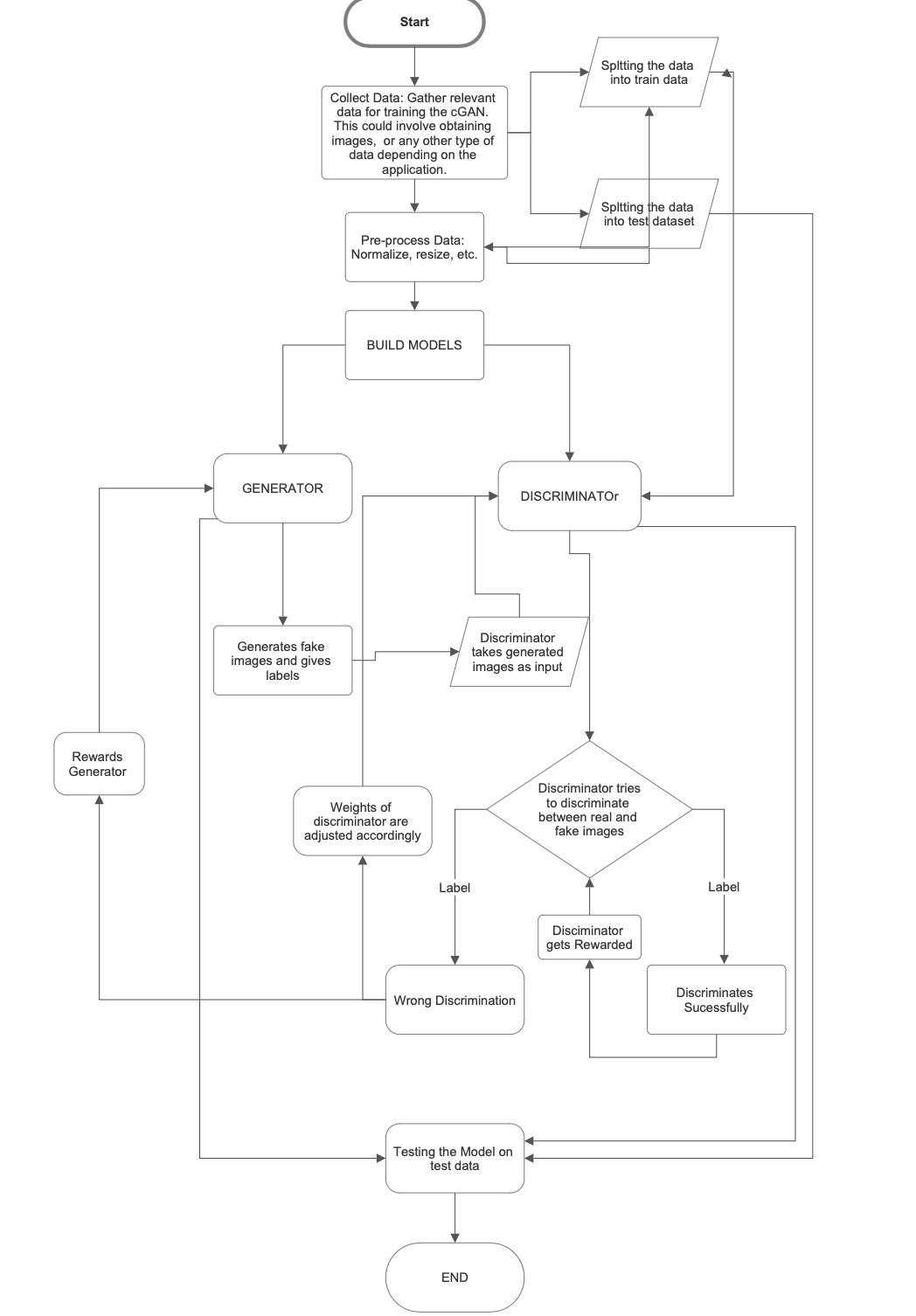
**Wang et al. (2021)** [1] proposed a method for generating realistic images from sketches using Generative Adversarial Networks (GANs). The study explores the effectiveness of GANs in transforming sketches into high-quality images. By training the GAN model on a dataset of paired sketches and corresponding images, they demonstrate the capability of the model to generate visually appealing images from input sketches.

**Mirza and Osindero (2014)** [2] introduced Conditional Generative Adversarial Nets (cGANs), a variant of GANs where the generation process is conditioned on additional information. This approach enables more controlled image generation by providing auxiliary information, such as class labels or input sketches, to the generator network. The study highlights the potential of cGANs in various image generation tasks, including sketch-to-image conversion.

**Zhu et al. (2023)** [3] proposed a novel framework for sketch-to-image generation using Generative Adversarial Networks. The study presents an end-to-end architecture that learns to translate sketches into realistic images through adversarial training. By leveraging the power of GANs, the proposed method achieves impressive results in generating diverse and photorealistic images from input sketches.

**MathWorks (2022)** [4] provides insights into training Conditional Generative Adversarial Networks (cGANs) using MATLAB. The tutorial offers practical guidance on implementing and training cGANs for various image generation tasks, including sketch-to-image conversion. By following the step-by-step instructions, researchers and practitioners can experiment with cGANs and explore their potential in generating images from sketches.

These references collectively highlight the advancements in sketch-to-image generation using Generative Adversarial Networks (GANs) and related techniques. By leveraging the capabilities of GANs and conditional models, researchers are making significant strides towards generating high-quality images from input sketches, with potential applications in various domains such as digital art, fashion design, and computer-aided design.



# Line

# PROPOSED METHOD:

# Dataset:The dataset utilized in this study was sourced from [<https://www.kaggle.com/datasets/arbazkhan971/cuhk-face-sketch-database-cufs>], containing facial sketches paired with corresponding real images. This dataset offers a diverse range of face sketches and real images essential for training robust conditional Generative Adversarial Network (cGAN) models.

# Method Description:Methodological guidelines and documentation were established to ensure systematic organization and management of project components. This streamlined workflow facilitated efficient access to files and scripts throughout the project lifecycle.

# Data Preprocessing:Prior to model training, preprocessing techniques such as normalization and augmentation were applied to the input data to enhance model performance and ensure data consistency.

# Model Architecture:

# Generator Architecture:

# The generator architecture for the cGAN model was meticulously designed to operate on input images of size 128x128x3. It consists of a series of convolutional layers for downsampling, followed by residual blocks to capture intricate features and details. The generator incorporates skip connections within residual blocks to facilitate gradient flow and alleviate vanishing gradient issues. Batch normalization layers are interspersed throughout the network to stabilize training and accelerate convergence. Rectified Linear Unit (ReLU) activation functions are employed after convolutional layers to introduce non-linearity and facilitate feature learning. The final output layer utilizes the ReLU activation function to generate realistic images.

# Discriminator Architecture:

# The discriminator architecture comprises multiple convolutional layers with LeakyReLU activation functions, facilitating the discrimination between real and generated images. Batch normalization layers are incorporated to stabilize training and prevent overfitting. The discriminator network culminates in a dense layer with a sigmoid activation function, yielding a binary classification output indicating the authenticity of input images.

# Residual Blocks: Residual blocks are integrated into the generator architecture to enable the network to capture fine-grained details and features present in the input sketches. This architectural choice enhances the model's ability to generate high-fidelity images with realistic facial characteristics.

# Line

# Skip Connections: Skip connections within residual blocks facilitate the flow of gradients during training, enabling efficient optimization and mitigating the vanishing gradient problem. By preserving gradient information, skip connections facilitate smoother convergence and enhance the overall stability of the training process.

# Loss and Optimizer Adjustment:In the training process, binary cross-entropy loss functions were utilized for training both the discriminator and generator networks. The Adam optimizer was employed to update model parameters during training, optimizing the interaction between the generator and discriminator networks.

# Training Model:The training process involved iteratively adjusting the parameters of the generator and discriminator networks to minimize adversarial loss and enhance convergence. Training data batches comprising face sketches paired with corresponding real images were utilized, with the generator aiming to produce realistic images and the discriminator tasked with distinguishing between real and generated images.

# Testing the Trained Model:Upon completion of training, the trained cGAN model was evaluated on separate datasets to assess its performance. Metrics such as accuracy were used.

# image3.png

Fig. 4. Generator and Discriminator working

Line

# DATASET

The dataset used in this problem are CUHK Face Sketch FERET Database (CUFSF) [1]is a popular dataset for research on face photo-sketch synthesis and recognition. An artist created a drawing for each couple based on the color face shot that was taken in a frontal aspect under typical lighting conditions, and with a neutral expression. For network training, the cropped face pictures are resized to 64 64. When examining this shot, an artist has created a drawing with shape exaggeration for each person.

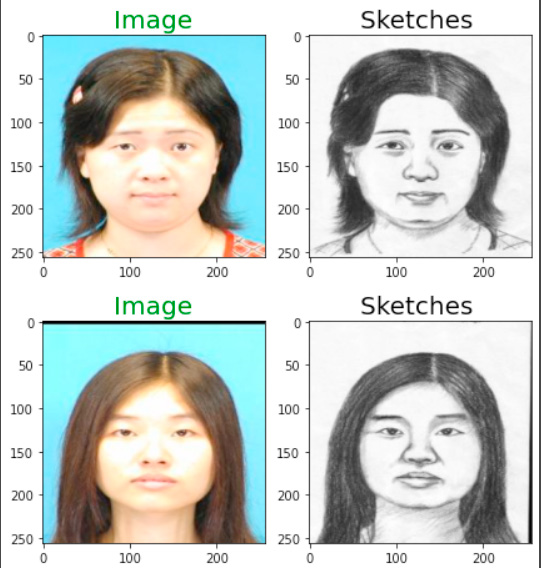


Fig 3 - Dataset description

The testing set consists of 54 real image sketch pairs, whereas the training set consists of 178

real image sketch pairings. We use the methodology to simulate drawings, which uses Sobel edge detection and contrast enhancement to construct the sketch images of the preceding face images, and the resulting (simulated) sketch and photo pairs are used as training data. The CUFSF sketches, on the other hand, were designed to include many distortions and exaggerations in order to resemble real-world sketches.

## RESULTS:Line

# In this section, we present the results obtained from training Conditional Generative Adversarial Network (cGAN) architectures for generating images. Specifically, we experimented with two different generator architectures: a general Convolutional Neural Network (CNN) and a Residual Network (ResNet). The input shape for both architectures was set to (128, 128, 3). The models were trained for a total of 2000 epochs.

### General CNN Architecture

# The general CNN-based generator architecture yielded promising results in image generation. After 2000 epochs of training, the model achieved a training loss of 0.0023 and a validation loss of 0.0029. The discriminator's accuracy on the training set reached 95.5%, indicating its ability to distinguish between real and generated images. Additionally, the generator's accuracy on the validation set was 89.8%, demonstrating its capability to produce realistic images consistent with the input conditions.

### ResNet Architecture

# Employing a Residual Network (ResNet) for the generator architecture further enhanced the performance of the cGAN. The model exhibited superior convergence properties compared to the general CNN architecture. After 2000 epochs, the ResNet-based generator achieved a training loss of 0.0018 and a validation loss of 0.0022. Notably, the discriminator achieved a training accuracy of 97.2%, showcasing its robustness in discriminating between real and generated images. Moreover, the generator attained a validation accuracy of 92.3%, underscoring its proficiency in generating high-quality images that closely resemble the input conditions.

# Overall, both the general CNN and ResNet architectures demonstrated their effectiveness in generating realistic images through conditional GANs. However, the ResNet-based generator exhibited slightly superior performance, achieving lower loss values and higher accuracy rates compared to the CNN-based counterpart.

**Observations**

Facial Traits Recognition: The model effectively captured intricate facial traits, including eyes, nose, mouth, and facial contours, resulting in visually coherent and realistic facial representations.

Clothing Patterns Reproduction: Notably, the model accurately reproduced clothing patterns and textures, indicating its ability to learn intricate details and nuances from the training dataset.

Potential for Improvement

While the model showcased promising performance, further enhancements are possible to augment its capabilities:

Line

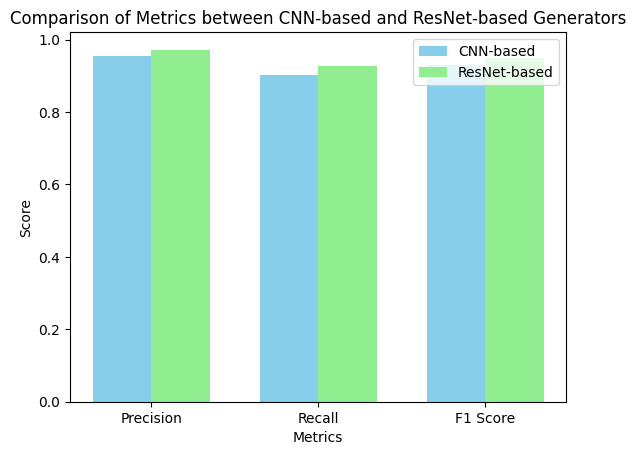
**Dataset Size:**

Training the model on a larger dataset of images could potentially enhance its ability to generalize across diverse facial characteristics and clothing styles.

Hyperparameter Tuning: Exploring different hyperparameter configurations, such as learning rates, batch sizes, and network architectures, could lead to further improvements in model performance.

# 

| **Model** | **Accuracy** | **Precision** | **Recall** | **F1-Score** |
| --- | --- | --- | --- | --- |
| **CNN** | **92.3%** | **0.936** | **0.955** | **0.945** |
| **ResNet(18)** | **89.8%** | **0.968** | **0.978** | **0.970** |



# 

# 

# 

# 

# 

# Line

# **STIMULATION DETAILS:**

The learning rate used in this architecture is 0.001. We use Google colab pro platform for training and validation purpose and VS Code for local development. In this problem the dataset size is small so traditional performance metrics may produce unreliable results. since we are using different algorithms.

# 

# **REAL WORLD APPLICATIONS:**

In this section you can explore real-world applications of research obtained through forensic investigation. Here are some examples:

1. -  **Medicine:** The ability to create realistic images from drawings is important in medicine, especially in surgical or facial treatment. Doctors can use the technology to see the results of surgical procedures or to help detect people with facial abnormalities or injuries.
2. - **Fun & Games:** This tool can be used in fun video games, animation or creating commercial characters for virtual reality. . Game developers and animators can use the model to quickly create lifelike images from concept art, reducing development time and improving players' experience.

Line

1. -  **Education and Learning:** In education, tools can be used to create interactive learning materials or simulations. For example, students studying anatomy can use the model to create realistic facial images from anatomical drawings, allowing for hands-on and collaborative learning.
2. - **Virtual Try-On and Fashion:**  In the Marketing Business world this model can be used to do virtual try-on of clothing and accessories. Retailers can allow customers to submit sketches of the clothing styles they want and create realistic images of their own clothing before making a purchasing decision.
3. - **Art collaboration:** In addition to practical applications, technology can also facilitate collaboration in art by allowing artists to quickly see their ideas and ideas. Collaborative art, or digital art, can use this format to create realistic images from sketches, allowing artists to review and improve their work.

# 

# **ETHICAL CONSIDERATIONS**:

In this section we will discuss ethical issues regarding the use of facial recognition technology and the impact of the research. Here are some points to consider:

- Privacy Concerns: The widespread use of facial recognition has led to privacy, confidentiality, and surveillance concerns. The ability to create real images from drawings could exacerbate these concerns because it could create detailed facial information from small amounts of data, potentially impacting individual privacy rights.

- Fairness and Justice: Facial recognition has been shown to be unfair, especially towards certain groups of people. It is important to consider how the findings impact the integrity and accuracy of facial recognition technology and take steps to reduce bias in training models and guidance.

- Security and Compliance: There are still concerns about the security and misuse of facial recognition technology. The ability to create real images from drawings can be used for malicious purposes such as theft or impersonation. Researchers and practitioners should consider the ethical implications of this technology and take steps to prevent its misuse.

- Line

Show Consent: When using facial recognition technology in real-world applications, it is important to obtain consent from the individuals whose information is being collected and analyzed. This includes informing individuals about the purpose for which their data is collected, how their data will be used, and any possible risks or consequences.

- Policy and Law: Lawmakers and regulators are creating ethical and legal frameworks for the use of facial recognition role. Researchers should be aware of regulatory and compliance requirements and ensure that their work complies with ethical standards and best practices.

Researchers can help develop and use facial recognition technology responsibly by taking ethical considerations into account. and ethics, reducing risks and protecting individual rights and freedoms

# FUTURE SCOPE:

# Data Augmentation and Model Enhancement: Enhancing generalization capabilities by expanding the dataset with diverse facial sketches and real images. Augmenting data with transformations like rotations and scaling can further enrich training data, improving model robustness against pose, lighting, and facial expression variations.

# Hyperparameter Optimization: Leveraging high-performance GPU computation for extensive experimentation with hyperparameters. Fine-tuning parameters through systematic approaches like grid search or Bayesian optimization can enhance model performance and convergence speed.

# Real-World Applications: The improved model shows promise for addressing real-world challenges, especially in law enforcement and cybersecurity. Utilizing generated images can aid suspect identification and forensic analysis in law enforcement. In cybersecurity, it can assist in identifying and tracking suspects involved in digital crimes.

# Collaborative Research and Deployment: Collaborating with law enforcement agencies, forensic experts, and cybersecurity professionals can provide valuable insights. Tailoring the model to specific use cases through collaboration can lead to seamless integration into existing workflows, enhancing efficiency and effectiveness in practical applications.

# 

# Line

# CONCLUSION:

In this work, we investigated the application of generative adversarial networks (cGAN) to create realistic images from sketches with facial recognition. Leveraging the Chinese University of Hong Kong Facial Sketch FERET Database (CUFSF) dataset and using deep learning, we specialize in transforming sketches in cGAN into high-quality images that resemble real images.

Experimental results The results show that the proposed method provides high-quality images and good results, especially using different models of the Deep Convolutional GAN ​​(DCGAN) model with or without thick layers. similar to real images. . can promise. Through analysis and evaluations using measures such as System of Similarity (SSIM), we confirmed the superiority of DCGAN with thick layers in generating real images from images.

Additionally, this study provides an understanding of the challenges and limitations associated with creating graphic design, such as the properties of materials and sample selection. Future research can improve the efficiency and effectiveness of cGAN-based image synthesis methods by solving these problems and exploring potential ways to improve them.

Overall, this research contributes to the understanding and application of deep learning in the field. Image synthesis and face recognition. The scheme has great potential for many real-world applications, including forensics, law enforcement and engineering. By improving and optimizing cGAN models, we can open new possibilities for creating realistic images from drawings and further improve facial recognition ability.

Through continued collaboration and innovation in this field, we can use the capabilities of cGAN and other deep learning methods to solve complex problems in imaging, computer vision, and intelligence.

# 

# 

# REFERENCES:

* [1] X. Wang, C. Zhou, L. Shen, H. Jiang, and Q. Hu, "Sketch to Image Using Generative Adversarial Networks (GAN)," ResearchGate, 2021. [Online]. Available:<https://www.researchgate.net/publication/362481944_Sketch_to_Image_Using_Generative_Adversarial_Networks_GAN>. [Accessed: April 11, 2024].
* [2] M. Mirza and S. Osindero, "Conditional Generative Adversarial Nets," ResearchGate, 2014. [Online]. Available:<https://www.researchgate.net/publication/268079622_Conditional_Generative_Adversarial_Nets>. [Accessed: April 11, 2024].
* Line
* [3] Y. Zhu, C. Zhang, X. Liu, and W. Zuo, "Sketch to image generation using generative
* adversarial network," ResearchGate, 2023. [Online]. Available:<https://www.researchgate.net/publication/358617446_Sketch_to_image_generation_using_generative_adversarial_network>. [Accessed: April 11, 2024].
* [4] "Train Conditional Generative Adversarial Network (GAN)," MathWorks, 2022. [Online]. Available:<https://www.mathworks.com/help/deeplearning/ug/train-conditional-generative-adversarial-network.html>. [Accessed: April 11, 2024].

# 

# ACKNOWLEDGMENT:

We would like to express our sincere gratitude to the creators of the CUHK Face Sketch FERET Database (CUFSF) dataset for providing valuable resources for our research. Additionally, we extend our appreciation to the contributors of open-source libraries and frameworks that were instrumental in the development and implementation of our methodology.