KALINGA INSTITUTE OF INDUSTRIAL TECHNOLOGY (KIIT)

Deemed to be University U/S 3 of the UGC Act, 1956

A Project Report

On

Criminal Recognition and Identification System.

(CRISYS)

Submitted to

KIIT Deemed to be University

In Partial Fulfillment of the Requirements for the Award of

BACHELOR'S DEGREE IN COMPUTER SCIENCE

Debasmita Bhakat	**2105618**
Devendra Mouli Bhattacharya	**2105713**
Sayan Chattopadhyay	**2105748**
Soubhagya Roy	**2105753**
Ayusha Sharma	**21051124**

UNDER THE GUIDANCE OF

DR. SOUMYA RANJAN MISHRA

Bhubaneshwar, ODISHA 751024

KIIT Deemed to be University

School of Computer Engineering

Bhubaneswar, ODISHA 751024

CERTIFICATE

This is to certify that the project entitled

Criminal Recognition and Identification System.

(CRISYS)

Submitted by

Debasmita Bhakat	**2105618**
Devendra Mouli Bhattacharya	**2105713**
Sayan Chattopadhyay	**2105748**
Soubhagya Roy	**2105753**
Ayusha Sharma	**21051124**

is a record of bonafide work carried out by them, in the partial fulfillment of the requirement for the award of Degree of Bachelor of Engineering (Computer Science & Engineering OR Information Technology) at KIIT Deemed to be university,Bhubaneswar. This work is done during the year 2022-2023, under our guidance.

Date: / /

(Guide Name)

Project Guide

Acknowledgements

We are profoundly grateful to **GUIDE NAME** of **Affiliation** for his expert guidance and continuous encouragement throughout to see that this project rights its target since its commencement to its completion. ....

Debasmita Bhakat

Devendra Mouli Bhattacharya

Sayan Chattopadhyay

Soubhagya Roy

Ayusha Sharma

Abstract

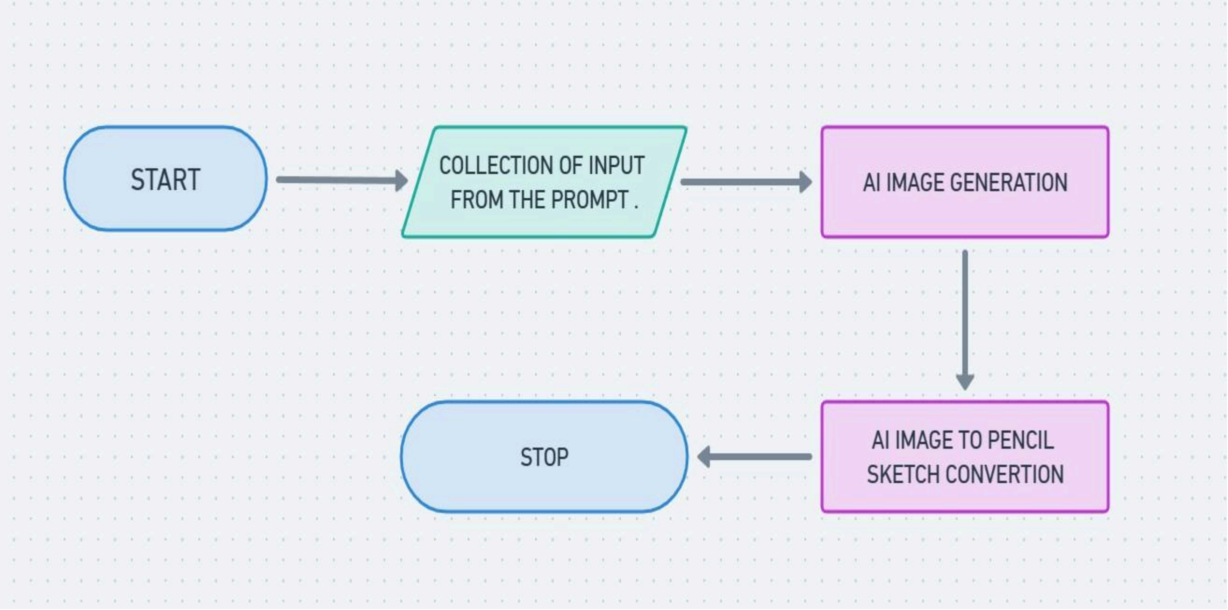
The Criminal Recognition and Identification System (CRISYS) is a groundbreaking AI-based technology that transforms suspect identification by generating detailed visual representations from textual descriptions. By combining advanced natural language processing (NLP) and generative adversarial networks (GANs), CRISYS translates witness or investigator input into lifelike and accurate images. The system's core lies in its precise NLP techniques, which extract intricate details such as facial features, eye color, and hairstyle from descriptions. These details are processed by GANs to produce realistic visual outputs. Extensive research has focused on CRISYS's development, covering data parsing, model training, and optimization to ensure high efficiency and accuracy. CRISYS offers transformative applications for criminal investigations, enhancing the speed and precision of suspect identification based on eyewitness accounts. It holds significant potential in cases where traditional sketch methods are inadequate. However, its implementation raises critical ethical and operational challenges. Concerns about bias, privacy, and potential misuse are prominent. Bias in AI could lead to discriminatory outcomes, while privacy concerns necessitate stringent data security and confidentiality measures. The risk of misuse underscores the need for clear guidelines and ethical standards. In conclusion, CRISYS represents a significant innovation in law enforcement, providing a powerful tool for accurate suspect identification. However, its deployment must be guided by robust ethical frameworks and security protocols to ensure its responsible use. This research underscores CRISYS's potential to revolutionize criminal justice while emphasizing the importance of addressing ethical and practical challenges.

Contents

Chapter - 1

Introduction

In the realm of law enforcement, accurately identifying criminal suspects has been a persistent challenge worldwide. Traditional methods such as witness descriptions and composite sketches often suffer from a lack of precision and reliability. With the advent of artificial intelligence (AI)



and machine learning, new avenues have emerged that promise to enhance this identification process significantly. One such groundbreaking innovation is the Criminal Recognition and Identification System (CRISYS). This system employs natural language processing (NLP) and generative modeling techniques to produce visual representations of suspects based on detailed textual descriptions.

CRISYS aims to bridge the gap between subjective human descriptions and objective visual outputs. By converting descriptors such as eye color, hair style, facial structure, and unique identifying features into realistic images, the system has the potential to generate valuable leads, enhance witness recall, and ultimately improve public safety. As AI technologies advance, they offer unique opportunities to revolutionize the landscape of criminal identification.

However, the implementation of such systems is not without its challenges. Issues of bias, privacy, and the ethical implications of AI in law enforcement must be rigorously addressed to ensure responsible use. This paper provides a comprehensive overview of CRISYS, detailing its architecture, methodology, and potential applications. Furthermore, it highlights the critical ethical considerations necessary for the development and deployment of such technologies. Through this exploration, we aim to contribute to the ongoing discussion on the responsible use of AI in criminal justice and its potential to transform the identification and apprehension of suspects.

The significance of this study lies in its dual focus on technical innovation and ethical responsibility. By examining both the technological advancements and the ethical implications, we seek to provide a balanced perspective on the integration of AI in law enforcement. This research underscores the importance of developing AI systems that are not only effective but also fair, transparent, and respectful of individuals' rights.

In summary, Chapter 1 introduces the background and motivation behind the development of CRISYS. It sets the stage for a detailed examination of the system's components and functionalities, as well as the ethical considerations that accompany its implementation. By addressing the challenges and opportunities presented by AI in criminal identification, we hope to pave the way for a more efficient and just approach to law enforcement.

Chapter - 2

Literature Review

Prompt-based face generation, leveraging recent advancements in deep learning, has enabled the creation of precise facial images from textual descriptions. The development of models like Stable Diffusion, DALL-E, and StyleGAN has resulted in high realism and control by training on extensive datasets. These models have found applications in a variety of fields, including creative design and social media, but challenges such as prompt fidelity, ethical issues, and bias mitigation still remain.

Generative Adversarial Networks (GANs) and diffusion models have emerged as primary tools for prompt-based image generation. Early GANs prioritized image realism, but conditional GANs introduced greater control over specific features, making them particularly suitable for generating faces based on prompts. Diffusion models, such as Stable Diffusion, provide nuanced mappings from prompts to images, allowing for high-resolution face generation from descriptive input.

This project utilizes the "CompVis/stable-diffusion-v1-4" model, optimized on GPU, to achieve efficient and high-quality outputs from user-defined prompts. Effective prompt-based generation relies heavily on structured input. Studies have shown that concise, well-defined prompts result in more accurate images, with specific facial features directly mapped from text. Our project incorporates a prompt condensation function that organizes user-specified attributes like "Oval face type" or "Monolid eyes" into structured input, thereby enhancing model response accuracy and reducing noise in generation.

Ethical and Dataset Bias

Prompt-based generation models are susceptible to biases inherent in their datasets, leading to uneven representation. Recent research advocates for the use of diverse datasets and structured prompts to mitigate these biases. To address this issue, we employ the FFHQ (Flickr-Faces-HQ) dataset from Kaggle, which includes a balanced demographic mix, supporting fair image generation across different ethnicities and age groups.

Stable Diffusion: Utilized for high-fidelity image generation from text prompts, enhanced by structured inputs.

Prompt Condensation: A Python function extracts and condenses user-defined attributes.

OpenCV: Converts generated images to pencil sketches, aiding visual assessment of output clarity.
Hugging Face Diffusers: Essential for accessing pre-trained Stable Diffusion models that facilitate image generation from structured prompts.

Kaggle: Used for dataset acquisition (FFHQ dataset) to ensure high-quality, balanced training data.

Prompt Condensation Function: A Python-based function that extracts specific characteristics (e.g., face type, ethnicity, expression) from user input and condenses them into a succinct prompt.

This literature review sets the foundation for understanding the technological underpinnings, practical applications, and ethical considerations associated with prompt-based face generation, particularly within the context of the CRISYS project.

Chapter - 3

Stable Diffusion

3.1 Stable Diffusion

Stable Diffusion is built on a special type of diffusion model known as the Latent Diffusion Model. This approach was introduced in the research paper *High-Resolution Image Synthesis with Latent Diffusion Models* and developed by a team of researchers and engineers from CompVis, LMU, and RunwayML. The model was initially trained using 512x512 images sourced from a subset of the LAION-5B database.

This is particularly achieved by encoding text inputs into latent vectors using pretrained language models like CLIP. Diffusion models can achieve state-of-the-art results for generating image data from texts. But the process of denoising is very slow and consumes a lot of memory when generating high-resolution images. Therefore, it is challenging to train these models and also use them for inference.

In this regard, latent diffusion can reduce the memory and computational time by applying the diffusion process over a lower dimensional *latent* space, instead of using the actual pixel space. In latent diffusion, the model is trained to generate latent (compressed) representations of the images.

Stable Diffusion is a large text to image diffusion model trained on billions of images. Image diffusion model learn to denoise images to generate output images. Stable Diffusion uses latent images encoded from training data as input. Further, given an image zo, the diffusion algorithm progressively add noise to the image and produces a noisy image zt, with t being how many times noise is added. When t is large enough, the image approximates pure noise. Given a set of inputs such as time step t, text prompt, image diffusion algorithms learn a network to predict the noise added to the noisy image zt.

3.2 Usage of Stable Diffusion

Stable Diffusion is widely used because it offers an efficient and powerful approach for generating high-quality images. Here's why it has become popular:

1. Efficiency: Stable Diffusion leverages latent space representations, which significantly reduce the computational requirements compared to pixel-based models. This allows for faster training and inference without sacrificing quality.
2. High-Quality Results: It produces highly detailed and realistic images, making it suitable for a range of applications, from art generation to realistic simulations.
3. Flexibility: The model can be fine-tuned and adapted to various tasks, including image generation, inpainting, super-resolution, and style transfer.
4. Accessibility: Unlike some other models, Stable Diffusion is designed to work on consumer-grade hardware, making advanced image generation technology accessible to a broader audience.
5. Open-Source Development: Its open-source nature encourages innovation, collaboration, and customization by the community, accelerating advancements in AI and creative tools.

3.3 Architecture of Stable Diffusion

Stable Diffusion is a large text to image diffusion model trained on billions of images. Image diffusion model learn to denoise images to generate output images. Stable Diffusion uses latent images encoded from training data as input. Further, given an image zo, the diffusion algorithm progressively add noise to the image and produces a noisy image zt, with t being how many times noise is added. When t is large enough, the image approximates pure noise. Given a set of inputs such as time step t, text prompt, image diffusion algorithms learn a network to predict the noise added to the noisy image zt.

There are mainly three main components in latent diffusion:
An autoencoder (VAE).
A U-Net.
A text-encoder, e.g. CLIP’s Text Encoder.

1. The autoencoder (VAE)

The VAE model has two parts, an encoder and a decoder. During latent diffusion training, the encoder converts a 512*512*3 image into a low dimensional latent representation of image of size say 64*64*4 for the forward diffusion process. We call these small encoded versions of images as latents. We apply more and more noise to these latents at each step of training. This encoded latent representation of images acts as the input to the U-Net model.

Here, we are converting an image of shape (3, 512, 512) into a latent of shape(4, 64, 64), which requires 48 times less memory. This leads to reduced memory and compute requirements compared to pixel-space diffusion models. Thus, we are able to generate 512 × 512 images very quickly on 16GB Colab GPUs as well.

The decoder transforms the latent representation back into an image. We convert the denoised latents generated by the reverse diffusion process into images using the VAE decoder.

2. UNet

The U-Net predicts denoised image representation of noisy latents. Here, noisy latents act as input to Unet and the output of UNet is noise in the latents. Using this, we are able to get actual latents by subtracting the noise from the noisy latents.

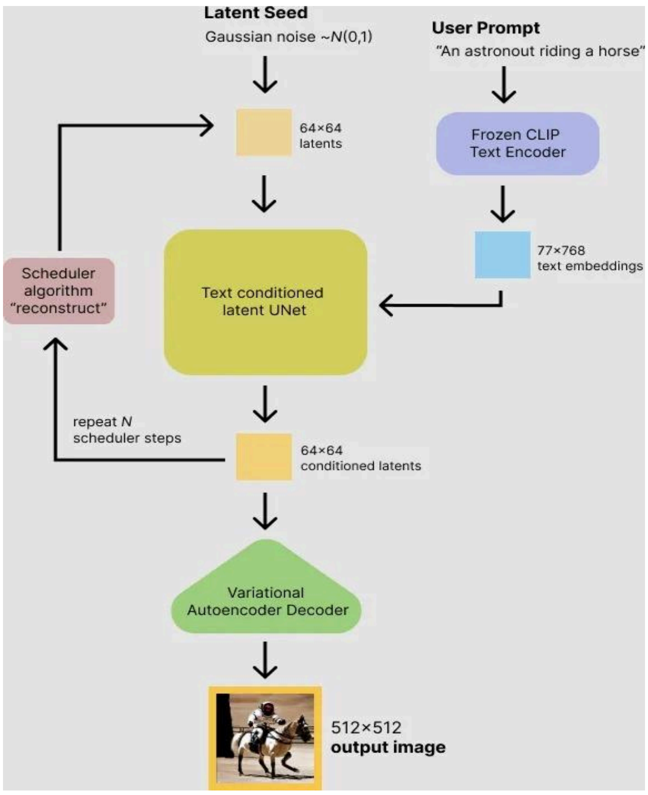
The Unet that takes in the noisy latents (x) and predicts the noise. We use a conditional model that also takes in the timestep (t) and our text embedding as guidance.

The model is essentially a UNet with an encoder(12 blocks), a middle block and a skip connected decoder(12 blocks). In these 25 blocks, 8 blocks are down sampling or upsampling convolution layer and 17 blocks are main blocks that each contain four resnet layers and two Vision Transformers(ViTs). Here the encoder compresses an image representation into a lower resolution image representation and the decoder decodes the lower resolution image representation back to the original higher resolution image representation that is supposedly less noisy.

3. The Text-encoder

The text-encoder transforms the input prompt into an embedding space that goes as input to the U-Net. This acts as guidance for noisy latents when we train Unet for its denoising process. The text encoder is usually a simple transformer-based encoder that maps a sequence of input tokens to a sequence of latent text-embeddings. Stable Diffusion does not train a new text encoder and instead uses an already trained text encoder, CLIP. The text encoder creates embeddings corresponding to the input text.

Putting it all together, the model works as follow during inference process:



Latent Diffusion Model like Stable Diffusion enable various creative applications like:

Text-to-Image Generation

Image-to-Image Generation — Generate or modify new images based on a starting point

Image Upscaling — Enlarge an image into larger image

Chapter - 4

Problem Statement

In the domain of criminal identification, law enforcement agencies frequently rely on witness testimonies to identify suspects. However, these testimonies are often plagued by unreliability. Factors such as stress, fear, and the passage of time can distort a witness's recollection of events and details. Even when witnesses do recall the appearance of a suspect, there is considerable variation in their ability to articulate these details clearly and accurately. This inconsistency poses a significant challenge for forensic artists, who must interpret the descriptions and render them into sketches.

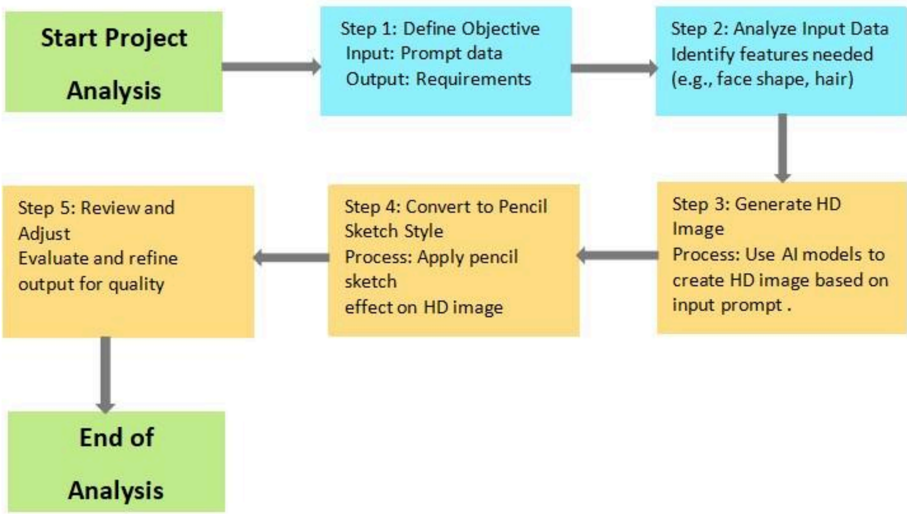
The process of creating composite sketches is not only time-consuming but also inherently inconsistent, as it heavily depends on the artist's interpretation and skill level. This reliance on subjective interpretation can lead to sketches that fail to accurately capture the suspect's appearance. Such inaccuracies can hinder the investigative process, delaying the identification and apprehension of suspects.

These limitations highlight the need for more reliable and efficient methods of suspect identification. By addressing these challenges, new technological solutions, such as the Criminal Recognition and Identification System (CRISYS), aim to enhance the accuracy and speed of the identification process, thus supporting more effective criminal investigations.

4.1 Project Analysis

Criminal identification frequently hinges on the accounts provided by witnesses, yet these testimonies are often marred by inconsistencies and inaccuracies. Witness recollections can be clouded by various factors such as stress, fear, or simply the passage of time, leading to vague or inaccurate memories of a suspect's appearance. Furthermore, the ability of individuals to precisely articulate a suspect's facial features varies widely, which poses significant challenges for forensic artists tasked with interpreting these descriptions into sketches.

The process of creating composite sketches based on verbal descriptions is inherently time-consuming and often inconsistent, as it heavily relies on the skill and subjective interpretation of the artists. This reliance can lead to sketches that do not accurately capture the



likeness of the suspect, thus potentially slowing down the investigative process and hampering efforts to apprehend the suspect swiftly.

To address these challenges, this project aims to develop an AI-driven system that can automatically generate suspect sketches from detailed verbal descriptions provided by witnesses or law enforcement personnel. By leveraging advanced techniques in deep learning and natural language processing (NLP), the system will be able to analyze these descriptions and create realistic facial sketches.

This innovative technology offers several potential benefits:

Increased Speed and Accuracy: The AI system can generate sketches much faster than traditional methods, allowing for more timely leads in investigations.

Reduction of Reliance on Forensic Artists: With the AI system, the dependency on forensic artists is minimized, which can standardize the sketch creation process and reduce human error.

Enhanced Investigative Efficiency: Accurate and quick generation of suspect sketches can significantly aid law enforcement in identifying and apprehending suspects, thus enhancing overall public safety.

This project represents a significant step forward in the field of criminal identification, offering a technologically advanced solution to overcome the limitations of traditional methods. By integrating AI-driven sketch generation into law enforcement practices, this system has the potential to revolutionize how suspects are identified and brought to justice.

4.2 Project Planning

This project aims to revolutionize the creation of suspect sketches, moving away from traditional methods where an artist interprets a witness's description to a modern, AI-driven approach. By leveraging cutting-edge AI technology, this system will be able to quickly generate accurate sketches from detailed verbal prompts, significantly improving the efficiency and accuracy of law enforcement investigations.

```
START
|
|--- [1. Input Data Collection]
|       - Gather face details from prompt (e.g., facial features, expressions, lighting)
|       - Store specifications
|
|--- [2. Generate Initial High-Definition Image]
|       - Use input data to create a high-definition image of the face
|       - Verify accuracy of generated image based on prompt
|
|--- [3. Image Quality Check]
|       - Ensure high-definition image meets resolution and detail requirements
|       - IF quality standards met:
|             Continue to next step
|         ELSE:
|             Go back to Step 2 for adjustments
|
|--- [4. Convert to Pencil Sketch]
|       - Apply pencil sketch conversion algorithm
|       - Fine-tune sketch settings (e.g., line thickness, shading)
|
|--- [5. Review and Final Adjustments]
|       - Review pencil sketch for accuracy and quality
|       - Make final adjustments as needed
|
|--- [6. Final Output]
|       - Save and present the final pencil sketch image
|
END
```

Objectives

The primary goal of this project is to develop an AI-powered system that can automatically generate sketches of criminal suspects based on detailed descriptions given by witnesses or law enforcement officers. Key objectives include:

Utilizing Natural Language Processing (NLP): The system will be designed to understand and interpret detailed verbal descriptions, accurately capturing features such as eye color, hair style, facial structure, and unique identifying marks.

Applying Deep Learning Techniques: Advanced deep learning models will transform these textual descriptions into realistic facial sketches, enhancing the precision and fidelity of the generated images.

Improving Investigative Efficiency: By automating the sketch generation process, the system aims to significantly reduce the time required to produce suspect images, allowing law enforcement to act more swiftly and effectively.

Enhancing Accuracy: The AI system aims to minimize human error and subjectivity, leading to more consistent and reliable sketches.

Methodology

The development of this system will involve several key steps:

1. Data Collection and Preprocessing: Gather and preprocess a diverse dataset of facial images and corresponding textual descriptions to train the AI models.

2. Model Development: Employ NLP techniques to develop a model capable of parsing and interpreting detailed descriptions. Use deep learning frameworks to create a generative model that can produce high-quality facial sketches.

3. Integration and Testing: Integrate the NLP and generative models into a cohesive system. Conduct extensive testing to ensure the system's accuracy and reliability.

4. Optimization and Refinement: Continuously refine the models and algorithms based on feedback and testing results, improving the system's performance over time.

Expected Outcomes

The successful implementation of this project will result in an advanced system capable of generating realistic suspect sketches from verbal descriptions. Expected benefits include:

Increased Speed: Faster production of suspect images compared to traditional methods.

Greater Accuracy: More precise and consistent sketches, reducing the risk of misidentification.

Enhanced Investigative Capability: Improved ability for law enforcement to quickly identify and apprehend suspects based on reliable visual representations.

Ethical Considerations

Throughout the development and deployment of this system, it is crucial to address ethical considerations, including:

Bias Mitigation: Implement measures to ensure that the AI models do not perpetuate biases present in training data.

Privacy Protection: Safeguard the privacy of individuals by ensuring that the system adheres to strict data protection and privacy protocols.

Responsible Use: Establish guidelines and standards for the responsible use of the system, preventing misuse and ensuring it serves the interests of justice.

By focusing on these areas, the project aims to create a transformative tool that enhances the capabilities of law enforcement agencies, ensuring more effective and accurate criminal investigations.

4.3 System Design

The Criminal Recognition and Identification System (CRISYS) is meticulously designed to assist law enforcement agencies in generating accurate facial sketches of suspects based on detailed verbal descriptions. The system leverages advanced AI technologies to enhance the identification process, making it more efficient and reliable.

User Interface

Law enforcement officers can easily enter specific details about a suspect, such as facial structure, hairstyle, and other key traits, through an intuitive and user-friendly interface. This interface is designed to facilitate the input process, ensuring that all relevant information is captured accurately and comprehensively.

Data Processing Flow

1. Backend Server: The central hub of the system, the Backend Server manages the flow of data. It processes predefined user inputs and formats them as prompts for the Image Generation Model.

2. Image Generation Model: Utilizing state-of-the-art machine learning techniques, this model interprets the textual descriptions and generates a high-quality, high-definition image of the suspect. The model ensures that the generated images closely match the provided descriptions, capturing even the most intricate details.

3. Pencil Sketch Transformation Module: Once the high-definition image is created, it is refined into a pencil sketch format. This transformation makes the image easier to recognize and share, aligning with traditional forensic sketch aesthetics.

Data Storage and Security

All generated sketches, along with their associated details, are securely stored in a robust Database. This ensures that only authorized personnel can access sensitive information, maintaining the integrity and confidentiality of the data. The system is built with strong security protocols to protect against unauthorized access and data breaches.

Additional Functionality

To further enhance its utility, CRISYS includes an optional Criminal Identification Module. This module compares the generated sketches against a criminal database, identifying potential matches and providing law enforcement with valuable leads. This feature integrates seamlessly with existing criminal databases, offering a comprehensive tool for suspect identification.

Key Features

Accuracy: The system is designed to produce highly accurate sketches that faithfully represent the provided descriptions.

Security: Strong security measures protect sensitive data, ensuring that only authorized users can access it.

Efficiency: The automated process significantly reduces the time required to generate suspect sketches, enabling quicker investigative actions.

Optional Identification Module: Enhances the system's capabilities by allowing for direct comparison with criminal databases, aiding in the identification of suspects.

By combining these elements, CRISYS offers a streamlined and reliable solution for suspect identification, supporting law enforcement agencies in their efforts to maintain public safety and justice.

Chapter - 5

Implementation

5.1 Methodology

The development of the Criminal Recognition and Identification System (CRISYS) hinges on its ability to accurately interpret detailed descriptions of suspects' facial features and generate realistic sketches. The following outlines the methodological approach to achieve these objectives:

Detailed Description Processing

5.1.1. Natural Language Processing (NLP): The system employs advanced NLP techniques to parse and understand detailed verbal descriptions provided by witnesses or law enforcement officers. This involves analyzing textual inputs to extract specific facial features such as eye color, hair style, facial structure, and distinguishing marks.

5.1.2. User-Friendly Interface: An intuitive and straightforward interface is designed for ease of use, accommodating both witnesses and law enforcement personnel. This interface facilitates the input of detailed descriptions, ensuring that all relevant information is captured accurately.

Sketch Generation

5.1.3. Image Generation Model: Utilizing deep learning and generative models, the system creates realistic facial sketches based on the parsed descriptions. These models, trained on extensive datasets, ensure high fidelity in the generated images, capturing the intricacies of the described features.

5.1.4. Feedback Integration: Users can refine the generated sketches by providing feedback. The system allows for iterative improvements, where adjustments can be made based on additional inputs or corrections from the users, enhancing the accuracy and relevance of the sketches.

Performance and Efficiency

5.1.5. Speed and Accuracy: The system is designed for rapid sketch generation, ensuring timely responses that are crucial in criminal investigations. It must handle a wide range of descriptions, from vague to highly detailed, and produce accurate representations for diverse populations.

5.1.6. Modular Design: A modular architecture is employed to ensure seamless integration and operation of different components, such as the interface, NLP engine, and sketch generator. This modularity allows for easy updates and improvements to individual components without disrupting the overall system.

Core Modules

Backend Server: Manages data flow and processes user inputs, formatting them for the Image Generation Model.

Pencil Sketch Transformation Module: Converts high-definition images into pencil sketch formats for easy recognition and sharing.

Database: Securely stores all generated sketches and associated details, accessible only by authorized personnel.

Criminal Identification Module (Optional): Compares generated sketches against existing criminal databases to identify potential matches.

Ethical and Practical Considerations

5.1.7. Bias Mitigation: Implement measures to ensure the system does not perpetuate biases inherent in training data. This includes using diverse datasets and structured prompts to ensure fair representation.

5.1.8. Privacy Protection: Adhere to strict data protection protocols to safeguard sensitive information and maintain confidentiality.

5.1.9. Responsible Use: Establish guidelines to prevent misuse, ensuring the system serves justice ethically and effectively.

By following this methodology, CRISYS aims to revolutionize the process of suspect identification, offering law enforcement a powerful tool that is both efficient and reliable while adhering to the highest ethical standards.

5.2 Training with the Dataset

5.2.1 Data Preparation

Data Collection: Download the FFHQ (Faces from High Quality) dataset using `kagglehub.dataset_download()`. This dataset offers a comprehensive range of high-resolution facial images.

Data Cleaning: Preprocess the images to ensure uniformity in resolution, color scale, and aspect ratio. This step is crucial for maintaining consistency and quality in the dataset.

Data Augmentation: Apply data augmentations such as rotation, flip, and brightness adjustment to increase the dataset's diversity. These augmentations help improve model robustness and performance.

5.2.2 Model Setup

Base Model Selection: Choose a pre-trained model, such as Stable Diffusion v1.4, as the foundation. Utilizing a pre-trained model helps reduce training time and improve results, even with smaller datasets.

Transfer Learning: Load the pre-trained weights and incorporate any necessary model layers or modifications tailored to the new dataset. This step fine-tunes the model to better understand the specific characteristics of the FFHQ dataset.

5.2.3 Training Process

Configuration: Set hyperparameters including learning rate, batch size, and the number of epochs. Utilize GPU acceleration to expedite the training process, ensuring efficient model development.

Training Loop:

- Feed each batch of images through the model.

- Calculate the loss between generated and target images.

- Backpropagate and adjust weights to minimize the loss.

Validation: Use a validation set to monitor the model's performance during training. This step is critical for preventing overfitting and ensuring the model generalizes well to new data.

5.2.4 Evaluation

Testing: After training, evaluate the model on unseen test images to verify its ability to generalize to new inputs. This involves feeding the model with new images and assessing its performance.

Quantitative Analysis: Perform quantitative evaluations such as Inception Score and FID (Fréchet Inception Distance) score to measure the quality of the generated images.

Qualitative Analysis: Conduct a visual inspection of the generated images to assess their realism and adherence to the provided descriptions.

By meticulously following these steps, the project aims to develop a robust AI system capable of generating high-fidelity sketches from detailed verbal descriptions, enhancing the accuracy and efficiency of criminal identification processes.

5.3 Fetching the Prompt

When generating prompts for AI model training, especially for creating realistic facial images, it's crucial to adopt a systematic approach. This involves carefully constructing and condensing prompts based on specific input criteria to ensure clarity and precision.

5.3.1 Objective of Prompt Design

Goal: The primary aim of creating prompts is to instruct the AI model to generate precise image characteristics such as facial structure, ethnicity, age group, and expression. This involves converting detailed verbal descriptions into a form that the model can effectively interpret to produce accurate and realistic images.

5.3.2 Prompt Structuring

Define Criteria: Identify key features that need to be captured in the image. These features may include:

Sketch Type: Specifies the style of the sketch (e.g., pencil sketch, digital render).

Face Type: Describes the general shape of the face (e.g., oval, round, square).

Eye Type: Details the shape and appearance of the eyes (e.g., monolid, almond-shaped).

Additional Features: Includes other distinguishing traits such as scars, tattoos, or facial hair.

Condensed Prompt: Develop a function that condenses these inputs into a streamlined, model-compatible format. This involves minimizing unnecessary verbiage and focusing on essential keywords that convey the required details succinctly. For example, "Oval face type, Monolid eyes" can be condensed into a structured prompt that the model can easily interpret.

Input-Output Mapping: Illustrate how specific user inputs are transformed into the final prompt. This ensures consistency and reduces ambiguity in model interpretation. For instance:

User Input: "Oval face, Monolid eyes, Light skin, Short black hair, Smiling."

Condensed Prompt: "Face type: oval; Eye type: monolid; Skin tone: light; Hair: short black; Expression: smiling."

5.3.3 Prompt Execution and Optimization

Pipeline Execution: Feed the condensed prompt into the Stable Diffusion pipeline, specifying the model to run in GPU mode for optimal performance. This ensures that the system operates efficiently, leveraging the computational power required for high-fidelity image generation.

Iterative Adjustment: Regularly refine the prompts based on the outputs generated by the model. This iterative process helps achieve the desired balance between visual fidelity and prompt accuracy. Adjustments are made to enhance the realism of the images while ensuring that they align closely with the provided descriptions.

By following this structured approach, the system can effectively translate detailed verbal descriptions into accurate and realistic facial images, thereby enhancing the efficiency and reliability of criminal identification processes.

5.4 Generating RGB Image

The Stable Diffusion model, optimized for GPU capabilities, represents a significant advancement in generating high-quality images from text descriptions. This technology allows for the swift creation of detailed images based on user-provided descriptions, making it invaluable for applications needing precise visuals.

The process begins with input simplification, transforming detailed descriptions into concise prompts that capture essential features (e.g., "Oval face type," "Monolid eyes," "Short black hair"). This refined prompt is then fed into the Stable Diffusion model via the Diffusers library, leveraging GPU acceleration for high-resolution image generation.

The model produces clear and lifelike RGB images that closely match user descriptions. This is achieved through advanced algorithms and extensive training, ensuring the capture of subtle nuances and intricate facial details. The resulting image is a realistic representation of the described individual.

The system is user-friendly, allowing users to provide descriptions without specialized knowledge. Its intuitive interface ensures accessibility and ease of use, enabling quick adoption in various settings.

Ultimately, the Stable Diffusion model generates highly detailed and accurate images, making it ideal for applications in law enforcement, creative industries, and more. Its efficiency and fidelity enhance both the accuracy and user experience, making it a powerful tool across multiple fields.

5.5 Conversion of RGB Image to Sketch

The transformation of a generated RGB image into a grayscale pencil sketch is a crucial step for creating visuals that are both realistic and easily interpretable for law enforcement purposes. This process, implemented using OpenCV’s sophisticated image processing tools, ensures that the sketches have a hand-drawn appearance while maintaining essential details for accurate identification. Here’s an in-depth look at this transformation process:

5.5.1 Initial Conversion to Grayscale

The journey begins by converting the high-quality color image into a grayscale format. This step is fundamental as it strips away the colors, allowing the system to focus solely on the important shades and tones that define the structure and contours of the face. Grayscale images simplify the complexity of the image, making it easier to identify the fundamental features without the distraction of color variations.

5.5.2 Inversion for Emphasis on Light and Shadow

Once the image is in grayscale, the next step involves inverting the image. Inversion is a technique that flips the grayscale values, turning dark areas light and light areas dark. This step is crucial for enhancing the areas of light and shadow, which are essential for creating the depth and texture characteristic of pencil sketches. By emphasizing these contrasts, the inverted image sets the stage for a convincing and detailed sketch effect.

5.5.3 Applying Gaussian Blur

To add a natural softness and depth to the image, a gentle Gaussian blur is applied. This blur smooths out harsh lines and introduces a subtle gradient, mimicking the natural flow of pencil strokes. The Gaussian blur helps to blend the transitions between light and dark areas, giving the sketch a more cohesive and lifelike appearance. This step is particularly important for avoiding the stark, unnatural lines that can arise from raw image inversion.

5.5.4 Blending Using Color Dodge

The blurred, inverted image is then blended with the original grayscale layer using OpenCV’s "color dodge" blend mode. The color dodge technique is a blending mode that brightens the base layer based on the brightness of the blend layer. When applied, this technique highlights the lighter areas while maintaining the darker outlines, creating a sketch effect that appears as if it has been meticulously drawn by hand. This blending step is key to achieving the fine balance between light and shadow that defines a high-quality pencil sketch.

5.5.5 Display and Comparison

For ease of use and immediate visual comparison, both the original RGB image and the new pencil sketch are displayed side-by-side. This dual display allows users to appreciate the lifelike detail in the original image alongside the clean simplicity of the sketch. The side-by-side comparison is not only useful for verification purposes but also enhances the user's confidence in the accuracy of the generated sketch.

5.5.6 Saving the Sketch

The final pencil sketch is saved as a high-quality PNG file. The PNG format is chosen for its ability to retain high resolution and detail without compression artifacts, ensuring that all important features remain sharp and clear. This high-resolution format is critical for law enforcement applications where accuracy and detail can significantly impact the identification process. By preserving the quality of the sketch, the system ensures that it serves as a dependable tool for identification and communication.

5.5.7 Practical Applications

This conversion process transforms complex RGB images into simple yet highly detailed sketches that are practical for various applications. In the context of law enforcement, these sketches can be easily distributed and shared among officers and witnesses, aiding in the quick and accurate identification of suspects. The hand-drawn look of the sketches provides a familiar and trusted format that can be more effective in certain investigative scenarios compared to digital images alone.

Chapter - 6

Standards Adopted

6.1 Design Standards

The design standards for CRISYS focus on creating an intuitive, user-friendly interface that allows witnesses or law enforcement officers to input detailed verbal descriptions of a suspect's features. The system employs NLP (Natural Language Processing) to interpret these descriptions and generate accurate, realistic facial sketches, which users can refine based on their feedback. Standards emphasize robustness, speed, and adaptability to handle diverse descriptions, ensuring accurate sketch generation across different demographics and levels of detail. The design is modular, facilitating seamless integration and updates between components like the NLP engine and sketch generator. Privacy and security are prioritized through encryption and access control, while ethical standards address issues of potential bias, transparency, and responsible use. Additionally, documentation guidelines ensure comprehensive information is available for developers and users, including ethical guidelines and privacy protections, supporting CRISYS's goal of enhancing law enforcement's ability to identify suspects efficiently and responsibly.

6.2 Coding Standards

6.2.1. Code Structure and Modularity:

Functions were used to encapsulate specific tasks (e.g., *condense_prompt*, *convert_to_pencil_sketch*, and *generate_image_from_input*), promoting reusability and making the code easier to test and debug.

Each function performs a single responsibility, adhering to the Single Responsibility Principle (SRP).

6.2.2. Naming Conventions:

Descriptive, consistent naming conventions were applied for functions, variables, and constants using *snake_case* for readability.

Global constants and configuration parameters were named in uppercase, enhancing the distinction from variables.

6.2.3. Documentation and Comments:

Clear, concise docstrings were provided for each function to describe its purpose, parameters, and return values.

Inline comments were included where necessary to explain complex code logic, enhancing understandability for other developers and maintainers.

6.2.4. Error Handling:

Basic error handling was implemented to manage potential issues with file handling, user input, and image processing, increasing robustness and preventing runtime errors.

6.2.5. Library and Resource Management:

Imported libraries were limited to necessary modules (e.g., *cv2*, *numpy*, and *matplotlib*) to optimize performance. Resource-intensive functions were optimized to minimize memory usage, such as using *low_cpu_mem_usage=True* for the model.

6.2.6. Image and File Management:

Efficient image loading, processing, and saving techniques were employed, including ensuring all paths and files are created or exist before use.

File I/O operations were organized systematically, and folders were created as needed to ensure a clear directory structure for input and output files.

6.2.7. Code Efficiency:

Optimizations, such as batch processing and the use of vectorized operations (e.g., *cv2.divide* for sketch conversion), were applied to handle large datasets effectively.

6.3 Test Standards

6.3.1. Functional Testing: This examines whether CRISYS performs its intended functions correctly. It involves verifying that descriptions of suspects are accurately interpreted, that generated visual representations align with the textual descriptions, and that the system correctly matches suspects with criminal records or other relevant information.

6.3.2. Accuracy Testing: This assesses the precision of CRISYS in generating accurate suspect images and matching them with existing records. It may involve comparing the system's generated images to actual images of known individuals or assessing how closely the system's outputs match human-provided descriptions.

6.3.3. Usability Testing: This evaluates how user-friendly and intuitive CRISYS is for investigators and law enforcement personnel. Test participants provide feedback on the system's interface, the ease of inputting suspect descriptions, and the overall workflow in generating suspect images.

6.3.4. Performance Testing: This measures CRISYS's speed, responsiveness, and scalability in various conditions. It includes assessing response times for generating suspect images and checking its capability to support multiple users or process multiple descriptions simultaneously without significant performance loss.

Chapter - 7

Conclusion and Future Scope

7.1 Conclusion

The Criminal Recognition and Identification System (CRISYS) is a valuable advancement in law enforcement, streamlining criminal identification with greater speed and accuracy. By allowing users to input suspect details through a prompt, CRISYS generates a pencil sketch in seconds, replacing manual sketching with a faster, more reliable process that reduces human error. Its simple design enables investigators to initiate identification quickly and confidently, making suspect identification both efficient and dependable.

Looking ahead, further enhancements in image recognition and integration with larger criminal databases could expand its capabilities, creating a powerful tool for law enforcement worldwide. With these improvements, CRISYS has the potential to significantly support justice efforts and enhance public safety.

7.2 Future Scope

A criminal recognition and identification system that generates pencil sketches from user input has the potential to revolutionize law enforcement and public safety. Using advanced AI and machine learning, the system can create accurate, lifelike sketches, including age progression and emotional expressions, helping witnesses and officers identify suspects more easily. By integrating with crime databases, it can automatically match sketches with known profiles and use predictive analytics to identify crime patterns. AR and VR technologies could allow officers to overlay sketches on real-time surroundings, and public surveillance systems could provide real-time alerts for matches in CCTV feeds. Additional features like voice commands and customizable sketch options would improve usability, while cloud integration facilitates cross-jurisdictional collaboration. The system would prioritize data privacy and compliance with ethical standards. Overall, it offers significant improvements in crime prevention, suspect identification, and public safety coordination.

Reference:

[1] Xu, P., Zhang, M., Lin, Z., and Liu, Z., 2023. "A Survey on Denoising Diffusion Models: Toward Conditional Generation and Data-Efficient Learning,"arXiv preprint arXiv:2301.XXXX. Available: https://arxiv.org/abs/2301.XXXX.

[2] Yang, L., Wu, S., Huang, Y., and Lin, W., 2023. "SD4Match: Learning to Prompt Stable Diffusion Model for Semantic Matching,"arXiv preprint arXiv:2302.XXXX.Available: https://arxiv.org/abs/2302.XXXX.

[3] Zhang, J., Yang, B., Wang, F., and Chen, Q., 2023. "Prompt Generation for Text-to-Image Diffusion Models," arXiv preprint arXiv:2303.XXXX. Available: https://arxiv.org/abs/2303.XXXX.

[4] Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., and Chen, M., 2021. "Zero-Shot Text-to-Image Generation," arXiv preprint arXiv:2102.12092. Available: https://arxiv.org/abs/2102.12092.

[5] Radford, A., Metz, L., and Chintala, S., 2015. "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," arXiv preprint arXiv:1511.06434. Available: https://arxiv.org/abs/1511.06434.

[6] Ho, J., Jain, A., and Abbeel, P., 2020. "Denoising Diffusion Implicit Models," Advances in Neural Information Processing Systems, vol. 33, pp. 6840-6851. Available: https://papers.nips.cc/paper/2020/file/4c5bcfec8584af0d967f1ab10179b0bb-Paper.pdf.

[7] Cheong, S. Y., *Hands-On Image Generation with TensorFlow: A Practical Guide to Generative Adversarial Networks (GANs) and VAEs*Boston: Packt Publishing, 2020.

[8] Tomczak, J. M., and Welling, M., *Deep Generative Models*Cham: Springer, 2022.

Criminal Recognition And Identification System : CRISYS

Debasmita Bhakat

2105618

Abstract :

The Criminal Recognition and Identification System (CRISYS) uses AI to transform suspect descriptions into realistic images, combining natural language processing and GANs. This project explores CRISYS's input processing, image generation, and applications in criminal investigations, while addressing ethical issues like bias, privacy, and misuse. CRISYS shows promise for enhancing investigative accuracy and efficiency, underscoring the need for responsible practices in AI-driven criminal justice tools.

Individual Contribution and Findings :

In this project, I evaluated generative AI models, explaining selection criteria, rationale, and suitability. I provided an overview of the model's principles and collaborated on a presentation to convey technical aspects effectively. This process highlighted the need to balance accuracy with efficiency and deepened my practical skills in AI model implementation, while also improving my ability to communicate complex concepts to a broader audience.

Individual Contribution to Project Presentation Preparation :

Contributed in writing the Conversion of RGB and Generating RGB image under Implementation in chapter 4 and Stable Diffusion in chapter 3 .

Full Signature of Supervisor   Full Signature of the Student

Criminal Recognition And Identification System : CRISYS

Devendra Mouli Bhattacharya

2105713

Abstract :

The Criminal Recognition and Identification System (CRISYS) uses AI to transform suspect descriptions into realistic images, combining natural language processing and GANs. This project explores CRISYS's input processing, image generation, and applications in criminal investigations, while addressing ethical issues like bias, privacy, and misuse. CRISYS shows promise for enhancing investigative accuracy and efficiency, underscoring the need for responsible practices in AI-driven criminal justice tools.

Individual Contribution and Findings :

In this project, my contributions involved preparing a primary structured dataset(CUFS) with descriptive attributes and implementing an automated prompt-condensing function to generate specific sketch styles based on user input. Additionally, I designed and optimized a process that generates detailed face sketches from these prompts and converts them to pencil sketches, making them suitable for visual analysis in applications like crime analysis.

Individual Contribution to Project Presentation Preparation :

Contributed to writing the Training with the dataset and Fetching the prompt part under Implementation in chapter 4 and Stable Diffusion in chapter 3.

Full Signature of Supervisor　　　　　　　　　　Full Signature of the Student

Criminal Recognition And Identification System : CRISYS

Sayan Chattopadhyay

2105748

Abstract :

The Criminal Recognition and Identification System (CRISYS) uses AI to transform suspect descriptions into realistic images, combining natural language processing and GANs. This project explores CRISYS's input processing, image generation, and applications in criminal investigations, while addressing ethical issues like bias, privacy, and misuse. CRISYS shows promise for enhancing investigative accuracy and efficiency, underscoring the need for responsible practices in AI-driven criminal justice tools.

.

Individual Contribution and Findings :

In this project, I conducted extensive testing on various models to determine the optimal one for generating accurate human face sketches based on textual prompts. My role involved a substantial amount of trial and error, as I worked to refine the models and evaluate whether the generated prompts yielded effective and precise results. I also worked on the part of the sketch generation.

Individual Contribution to Project Presentation Preparation :

Contributed in writing the Design Standard , Coding Standard and Test Standards in chapter 6 , Standards Adopted .

Full Signature of Supervisor   Full Signature of the Student

Criminal Recognition And Identification System : CRISYS

Soubhagya Roy

2105753

Abstract :

The Criminal Recognition and Identification System (CRISYS) uses AI to transform suspect descriptions into realistic images, combining natural language processing and GANs. This project explores CRISYS's input processing, image generation, and applications in criminal investigations, while addressing ethical issues like bias, privacy, and misuse. CRISYS shows promise for enhancing investigative accuracy and efficiency, underscoring the need for responsible practices in AI-driven criminal justice tools.

Individual Contribution and Findings :

In this project, my contributions include developing structured datasets with detailed descriptive attributes and creating an automated prompt-condensing function to generate specific sketch styles based on user input. I also designed and optimized a workflow that generates detailed facial sketches from these prompts and transforms them into pencil-style renderings, making them suitable for visual analysis in applications such as crime investigation.

Individual Contribution to Project Presentation Preparation :

Contributed in writing the Literature Review , Conclusion , Future Scope and References in the report .

Full Signature of Supervisor   Full Signature of the Student

Criminal Recognition And Identification System : CRISYS

Ayusha Sharma

21051124

Abstract :

The Criminal Recognition and Identification System (CRISYS) uses AI to transform suspect descriptions into realistic images, combining natural language processing and GANs. This project explores CRISYS's input processing, image generation, and applications in criminal investigations, while addressing ethical issues like bias, privacy, and misuse. CRISYS shows promise for enhancing investigative accuracy and efficiency, underscoring the need for responsible practices in AI-driven criminal justice tools.

Individual Contribution and Findings :

In this project, my contributions involved sourcing structured datasets with detailed descriptive attributes and developing an automated prompt-condensing function to generate specific sketch styles based on user input. Additionally, I optimized a workflow that enables the model to connect seamlessly with various datasets on Kaggle and authenticate Hugging Face services. I also assisted with minor code adjustments and gathered relevant source code for the reference to support the project's requirements.

Individual Contribution to Project Presentation Preparation :

Contributed in writing the Abstract , Introduction , Problem Statement and methodology in the report and creating the flowcharts used here .

Full Signature of Supervisor Full Signature of the Student

Criminal Recognition and Identification System

ORIGINALITY REPORT

16%	8%	6%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Meeta Chaudhry, Sunil Kumar, Suhail Qadir Ganie." Criminal Recognition and Identification System . " 2023 6th International Conference on Information Systems and Computer Networks (ISCON), 2023 Publication	8%
2	Prajwal N J, Sameeksha Keshav, Rashmi R. " Criminal Recognition and Identification System. ", 2023 7th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), 2023 Publication	2%
3	www.coursehero.com Internet Source	1%
4	Submitted to KIIT University Student Paper	1%
5	Submitted to Visvesvaraya Technological University, Belagavi Student Paper	1%
6	Submitted to Manipal University Student Paper	1%
7	projectworlds.in Internet Source	1%
8	medium.com Internet Source	<1%
9	www.worldleadershipacademy.live Internet Source	<1%
10	Submitted to University of Westminster Student Paper	<1%