Project Work Phase-I on

**INTELLIGENT POLICE SKETCH SYSTEM USING VAE**

Submitted in partial fulfillment of the requirements for the award of the

## Bachelor of Technology

in

## Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning)

by

|  |  |
| --- | --- |
| **Nikhil Garimella** | **21241A6623** |
| **Sai Prakash Reddy** | **21241A6622** |
| **Shanmukh Challa** | **21241A6616** |
| **J.Yuva Teja** | **21241A6629** |

Under the Esteemed guidance of

## Ms. P. Shirisha

## Assistant Professor



**Department of Computer Science and Engineering**

(**Artificial Intelligence and Machine Learning)**

**GOKARAJU RANGARAJU INSTITUTE OF ENGINEERING AND TECHNOLOGY**

**(Approved by AICTE, Autonomous under JNTUH, Hyderabad) Bachupally, Kukatpally, Hyderabad-500090**



**GOKARAJU RANGARAJU INSTITUTE OF ENGINEERING AND TECHNOLOGY**

**(Autonomous)**

**Hyderabad-500090**

## CERTIFICATE

This is to certify that the major project entitled “**Intelligent Police Sketch System Using VAE**” is submitted by **Nikhil Garimella (21241A6623), Sai Prakash Reddy (21241A6622), Shanmukh Challa (21241A6616) and J. Yuva Teja (21241A6629)** in partial fulfillment of the award of degree in BACHELOR OF TECHNOLOGY in Computer Science and Engineering (Artificial Intelligence and Machine Learning) during Academic year 2024- 2025.

|  |  |
| --- | --- |
| **Internal Guide** | **Head of the Department** |
| **Ms. P. Shirisha** | **Dr. G. Karuna** |
| **External Examiner** | |

## ACKNOWLEDGEMENT

There are many people who helped us directly and indirectly to complete our project successfully. We would like to take this opportunity to thank one and all. First, we would like to express our deep gratitude towards our internal guide **Ms. P. Shirisha, Assistant Professor,** Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning), for her support in the completion of our dissertation. We are thankful to major project coordinator **Dr. R. P. Ram Kumar,** Professor, for his valuable suggestions and comments during this project period.

We wish to express our sincere thanks to **Dr. G. Karuna,** Head of the Department**,** and to our principal **Dr. J. Praveen,** for providing the facilities to complete the dissertation. We would like to thank all our faculty and friends for their help and constructive criticism during the project period. Finally, we are very much indebted to our parents for their moral support and encouragement to achieve goals.

|  |
| --- |
| **Nikhil Garimella (21241A6623)** |
| **Sai Prakash Reddy (21241A6622)** |
| **Shanmukh Challa (21241A6616)** |
| **J. Yuva Teja (21241A6629)** |

## DECLARATION

We hereby declare that the major project titled **“Intelligent Police Sketch System Using VAE”** is the work done during the period from **1st  August 2024 to 26th November 2024** and is submitted in the partial fulfillment of the requirements for the award of degree of Bachelor of Technology in Computer Science and Engineering (Artificial Intelligence and Machine Learning) from Gokaraju Rangaraju Institute of Engineering and Technology (Autonomous under Jawaharlal Nehru Technology University, Hyderabad). The results embodied in this project have not been submitted to any other University or Institution for the award of any degree or diploma.

|  |
| --- |
| **Nikhil Garimella (21241A6623)** |
| **Sai Prakash Reddy (21241A6622)** |
| **Shanmukh Challa (21241A6616)** |
| **J. Yuva Teja (21241A6629)** |

## ABSTRACT

The Sketch Generation in law enforcement process entails creating a sketch of the suspect through the descriptions given by the witnesses as they simplify the process of investigating. Essentially, the following are the research questions that the project expects to achieve: Improving the general appearance to assist in aligning the picture with that contained in the criminal repository with the intention of expediting investigation duration and; improving efficiency and effectiveness of drawing with use of slider features. The following methodology is used where first we, conduct a research and take large facial database then train VAE model and realistically generate sketches with the use of Facenet-512 or Facenet matching generated sketch to dataset and filter top ‘n’ (number of matches) and by using Decision tree, we design a GUI also with the feature of intensity, depth and breadth to enhance the result. The novelty in the present innovation is in using VAE driven sketch generators coupled with matching done by the Facenet based approach for precise identification of the suspect from a sketch. In addition, with the interactive slider element, we are certain that we will get a more accurate outcome and those established values can be fed back into the dataset for more training data.

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **Figure No.** | **Figure Name** | **Page No.** |
| 1.1 | Architecture Diagram | 5 |
| 3.1 | Connectivity Diagram | 23 |
| 3.2 | Class Diagram | 34 |
| 3.3 | Sequence Diagram | 35 |
| 3.4 | Use case Diagram | 36 |
| 3.5 | Activity Diagram | 37 |
| 4.1 | Representation of Dataset | 39 |
| 4.2 | Representation of Data | 41 |
| 4.3 | User Interface of the Sketch Generation System | 42 |
| 4.4 | Representation of the slider feature | 42 |
| 4.5 | Results of given inputs | 43 |

**LIST OF TABLES**

|  |  |  |
| --- | --- | --- |
| **Table No.** | **Table Name** | **Page No.** |
| 2.1 | Summary of the Existing Approaches | 11 |

**LIST OF ACRONYMS**

|  |  |
| --- | --- |
| **Acronym** | **Full Form** |
| VAE | Variable Auto Encoder |
| GAN | Generative Adversarial Networks |

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **Chapter No.** | **Chapter Name** | **Page No.** |
|  | Certificate | ii |
|  | Acknowledgement | iii |
|  | Declaration | iv |
|  | Abstract | v |
|  | List of Figures | vi |
|  | List of Tables | vii |
|  | List of Acronyms | viii |
| **1** | **Introduction** |  |
|  | 1.1 Introduction to the Project Work | 1 |
|  | 1.2 Objective of the Project | 2 |
|  | 1.3 Methodology | 2 |
|  | 1.4 Architecture diagram | 4 |
|  | 1.5 Organization of the Report | 5 |
| **2** | **Literature Survey** |  |
|  | 2.1 Summary of Existing Approaches | 7 |
|  | 2.2 Summary Drawbacks of Existing Approaches | 15 |
| **3** | **Proposed Method** |  |
|  | 3.1 Problem Statement & Objectives of the Project | 19 |
|  | 3.2 Architecture Diagram, Connectivity Diagram & Software and Hardware Requirements | 20 |
|  | 3.3 Modules and its Description | 24 |
|  | 3.4 Requirements Engineering | 30 |
|  | 3.5 Analysis and Design through UML | 33 |
| **4** | **Results and Discussions** |  |
|  | 4.1 Description about Dataset | 39 |
|  | 4.2 Detailed Explanation | 41 |
|  | 4.3 Significance of the Proposed Method with its Advantages | 45 |
| **5** | **Conclusion and Future Enhancements** | 47 |
|  | **Appendices** |  |
|  | 6.1 Exploratory Data Analysis | 50 |
|  | 6.2 Training CLF | 51 |
|  | 6.3 Training LLM | 52 |
|  | **References** | 55 |

## CHAPTER 1 INTRODUCTION

**1.1 Introduction to the Project Work**

In this project, the application of machine learning for improvement of police sketching for better operational efficiency in identification of suspects during police work. Previously police relied on forensic artists and the latter are given written descriptions of a suspect by the police under the directions of witnesses; the method is not only inaccurate due to human effort but also depends on the talent of a particular artist. Such cases shall also limit the extent to which the witnesses will be able to describe the facial features appropriately resulting in sketches that are less accurate. This project is a response to these challenges: Variational Autoencoders (VAEs) and FaceNet-512 are used in the semi-authoritative sketching process as well as in creating an additional interactive feedback loop to adjust the sketches.

The first process of this approach is to sample a massive facial image database for the learning purpose of the VAE model. VAEs are intuitively viewed as a type of deep-learning architecture particularly used for decoding new data that looks similar to the initial data. In this case, the VAE is trained to extract features of human faces and also the other related attributes of faces such as size and texture. The VAE is able to generate realistic face sketches from limited input and the parameters can also be real time controlled. This is advantageous over regular drawing styles, because it is more uniform and real life-like and can be adjusted in a second to accommodate for a witnesses change of mind.

After generating a facial sketch, it is then matched against a criminal database through the help of FaceNet-512. FaceNet-512 generic freshness deep learning model that is used for face recognition. This is done to convert the generated sketch and the images in the database to standard 512 feature vectors. It computes the distance between these vectors and based on which FaceNet-512 establishes how close the given sketch is to each image in the face database. The system then refines the matches to the top ‘n’ of the result list which actually seals the number of probable suspects, probably making it easier to solve the mystery. This means the investigators have a small number of people to investigate therefore speeding up the process.

Another improvement of this project is the inclusion of an adaptive Graphical User Interface (GUI) which is used for creating control sliders for the facial features, for instance, the intensity, depth, and width. This makes it possible for the users—whether the witnesses or investigators—to tune the generated sketch in a manner that it resembles the suspect’s likeness. This interaction aspect is important because facial images tend to be difficult to describe and memorize by witnesses, and simple alterations to some features such as the shape of the nose, or space between the eyes would make a lot of difference in the final sketch. These adjustments made at run time help make the generated sketch as resembling the witness’ memory as is possible.

Additionally, this process is constructive not only for creating the present sketch, but also for refining the system’s characteristics step by step. The final adjusted sketches can be plugged back into the model to improve the VAEs prediction ability for future cases of sketches generation. This feedback loop also makes it possible for the model to learn successively from new descriptions and enhance its capability of creating facial synthesis.

Furthermore, to help identify and refine-the best matches from the FaceNet-512 outcome, the Decision Tree algorithm is also utilized. A decision tree is a tree-like structure that is used in machine learning for preparing classifications. It divides decisions into a series of questions, which prioritize specific features as opposites to others in the context of facial images comparison. With reference to this context, it can be seen that the use of the decision tree assists the investigators to prioritize the highest possible matches about the suspect, consequently enhancing the effectiveness of the suspect identification process.

The usefulness of this project is based on the synthesis of several tools that as they are integrated make the police sketch much more accurate and faster drawn. Due to the use of VAEs when generating the sketches, the sketch can be set to be realistic, and the change in the sketches can be done on a real-time basis making it more flexible compared to other design methods. On the other hand, the use of FaceNet-512 allows the generated sketches to be aligned rapidly to a database of pre registered individuals. These generation and recognition steps provide a much higher improvement than traditional approaches to forensic sketching.

Many aspects are sped up by the system, while another key feature is the ability of users to refine the produced sketches by using the interactivity of the GUI. This feature is again a bonus to increase the accuracy of the final results and have the sketches look as the witness has described. The possibility to use slider controls to set such values as intensity, depth and amount gives users a great deal of freedom in the sketching which can have a great impact on the final result.

Hence, this project improves the existing police sketch generation by applying Variational Autoencoders (VAEs), FaceNet-512 and graphical user interface to increase suspicion’s accuracy as well as decrease the time required to identify the suspects. Besides the increase in time sparing, reliability experienced during the process of the sketches’ creation and real-time change, it is also worth mentioning that the system can be useful in the process of overcoming the challenges described. When using these advanced methods altogether, it is quicker to perform investigations, and added as a tool to the hand of the law enforcement to effectively narrow down the list of suspects to solve cases. Over time it simply becomes even more effective due to the learning from the users, and may just as easily become a most valuable method for investigation of crimes.

## 1.2 Objective of the Project

* To generate realistic facial sketches from witness descriptions, reducing reliance on manual sketch artists.
* To match the generated sketches with criminal databases, providing top 'n' suspect matches to accelerate investigations.
* To allow real-time adjustments to the sketch's intensity, depth, and breadth for higher accuracy.
* To Implement a feedback loop where user-modified sketches are added back into the dataset to improve the model's future performance.
* To Streamline the suspect identification process by reducing the time required for sketch generation and matching, improving overall investigative efficiency and accuracy.

## 1.3 Methodology

The work starts with data acquisition and data preprocessing of a large facial image data set which are used as the training data for the VAE. The VAE develops the way to encode faces into the compressed latent space and to reconstruct practically credible sketch based on the witness input. These sketches can be expanded with other variables, including intensity, depth and breadth variables.

Afterward, the VAE composes a facial sketch, the model passes it to FaceNet-512 to transform it into the 512-dimension which it desires; the criminal images stored are also translated into identical 512-dimension by FaceNet-512. Of the two FaceNet versions, FaceNet-512 determines the Euclidean distance between vectors to determine which ones are most similar to a sketch of a criminal. From this, one gets the list of the best ‘n’ matching candidates.

Facial reconstruction is drawn with a dynamic key-point model; an application program with Graphical User Interface (GUI) is created which allows the users (witnesses or investigators) to modify generated sketches using sliders. This make/sure that the sketch is fine tuned to the witness description in order to generate as accurate an image as possible. In a second step, the refined sketches can be used to update the VAE, so that the model is improved in the long term.

A Decision Tree classifier refines a list of matches obtained from FaceNet-512 and ranks them according to the metrics calculated from the witness statement, so it is more efficient. The prospect to define the case again and feed back new data means that the system improves its feedback as to the better cases.

In addition to this, the suggested methodology minimizes the time needed to identify a suspected candidate and generally provides a better and more effective remedy to the issue as a way of enhancing law enforcement tasks as opposed to the other techniques already developed.

## 1.4 Architecture Diagram

The Figure 1.1 shows that the VAE starts to create a rudimentary drawing or a scaffold of what a face looks like. From this sketch, the system goes to the Graphical User Interface (GUI) enables the users to modulate the intensity, depth, and breadth of the sketch using adjustable sliders to make a sketch much more precise depending on the feedback.

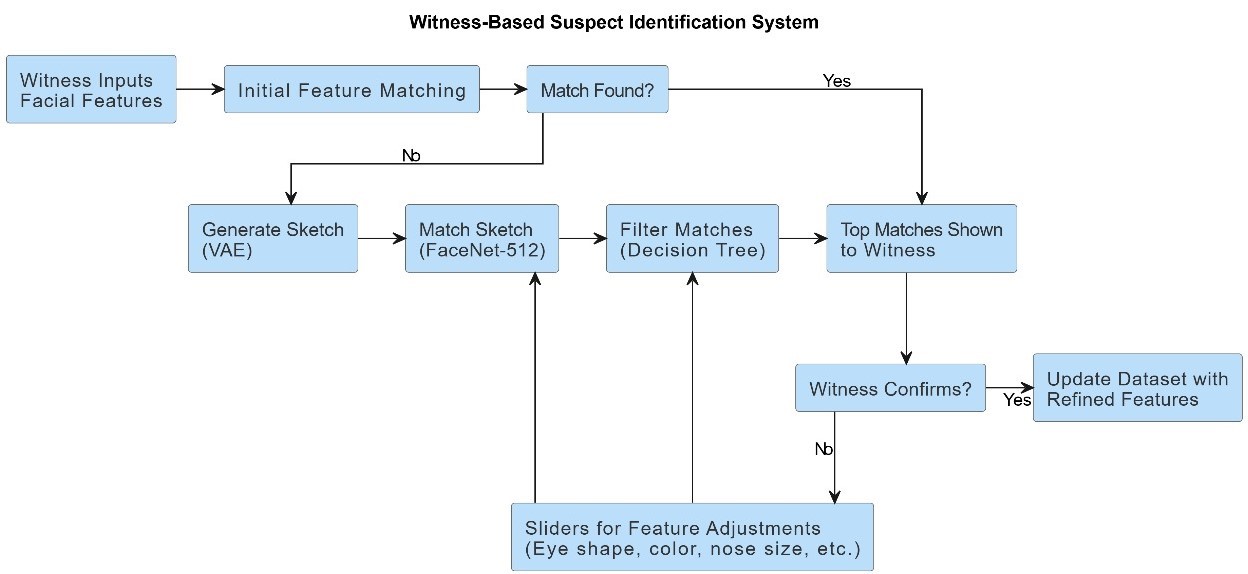
****

Figure 1.1 Architecture Diagram.

Final sketch is then put through FaceNet-512 to vectorize that particular sketch. It is then compared with criminals ‘that are in the record’ as well as serves to fit the ‘sketch’ to ‘the potential criminals’. The Decision Tree Algorithm organizes and chooses the involved ‘n’ numbers of matches to give the greatest number of suspects to the investigators.

This fast and effective approach allows the identification of the sketches as fast as possible and also iteratively, which improves the user interaction and also allows the models being improved constantly

## 1.5 Organization of the Report

This report offers a thorough summary of the subjects covered, arranged in a logical order to aid comprehension. From the introduction to the conclusion, every chapter addresses a different facet of the project and provides thorough explanations and evaluations.

**Chapter 1: Introduction**

In this chapter, we present the project and its goals, The Sketch Generation in law enforcement process entails creating a sketch of the suspect through the descriptions given by the witnesses as they simplify the process of investigating. Essentially, the following are the research questions that the project expects to achieve: Improving the general appearance to assist in aligning the picture with that contained in the criminal repository with the intention of expediting investigation duration and; improving efficiency and effectiveness of drawing with use of slider features.

**Chapter 2: Literature Survey**

In this section, we described the existing methods to solve the given problem of the Intelligent Police Sketch System and their pros and cons. These surveys in the literature created a background and pace necessary to accomplish the project efficiently

**Chapter 3: Proposed Methods**

Here, we describe the approaches used to tackle the problem of identifying and predicting tags. The described method starts with the Generation of Facial Sketches from Witness description from the Variational Autoencoder (VAE). These sketches are refined in a man-machine dialogue in which an interactive Graphical User Interface (GUI) contains slider controls for intensity, depth, and breadth. The last sketch is then fed to FaceNet-512 that brings it out to a feature space and compared with a database of criminals. The Decision Tree Algorithm sorts and narrows down the possible ‘match’ to ‘n’ and then leads the investigators to concentrate on the likely suspects, enhancing prospect accuracy and reduction of it as the criminals are identified.

**Chapter 4: Results and Discussions**The VAE is trained to encode facial images into a latent space and reconstruct them accurately. The training process is evaluated using two loss functions: reconstruction loss, which is computed using Mean Squared Error, and KL divergence loss that regulates the latent space regularization. It is expected that over 20 epochs, the total loss; hence the reconstruction ability as well as the feature encoding quality improves, drops for the VAE.

**Chapter 5: Conclusion and Future Enhancements**

To conclude, this chapter presents the summation of the project. This project offers a solution of using artificial intelligence in order “to automatically generate police sketches” using Tasks such as Variational Autoencoders (VAEs) and FaceNet-512 for better accuracy and faster suspect recognition. Real Time interaction of the sketches can be done on the Graphical User Interface and the decision tree part filters out the most likely culprits. Additional improvements include the additional use of multiple-input modes, fine-tuning of deep features, interface designed for mobile devices, inclusion of real-time databases where needed, and additional options for facial aging and emotional changes. These improvements would make the system more flexible with the police investigations and very efficient in their functions.

**Chapter 6: Appendices**

This chapter contains code and additional project-related resources.There would be descriptions of data, data cleaning, structures of the models such as the VAE and FaceNet-512, and the layout of the GUI. Other sections would include algorithm flow charts, the measures employed for measuring the performance of the system, and the result as well as case studies. Future enhancements would also be described with technical specification on possibilities of future advancements as much as what prospects such as, Multi-Modal input, mobile support, and advanced facial recognition among other

## CHAPTER 2

**LITERATURE SURVEY**

## 2.1 Summary of Existing Approaches

In this set of studies, a number of advanced approaches for constructing the police sketch and its derivatives have been proposed, and all of them address various issues inherent to the process of translating the witness testimonies into the readily recognizable and lifelike depictions. The first significant method involves multi-class sketch-based image generation using GAN inversion. This technique closes the domain gap between sketches and natural images by using a pre trained GAN generator, encoding sketches to a latent space, and incorporating a shape loss function to improve image realism. By using the latent space, the method reduces the challenge of directly map sketch to image and at the same time overcome the difficult task of generating realistic images. This approach is particularly effective in creating high quality and photo-realistic images at a higher quality than baseline methods and therefore provides a good solution for law enforcement in being able to identify suspects using abstract sketches by witnesses.

The sketch-based generation of photo-realistic images is specifically achieved by concentrating on GAN inversion for multi-class sketch-based image generation[1]. It attempts to bridge the domain gap between the abstract sketches and real natural images. It bypasses the complex image-to-image mapping by using a pretrained generator then again by mapping sketches to a latent space using a novel shape loss function that enhances the fidelity of generated images. This approach addresses retrieval using sketches for 3D CAD models[2]. Loop relation trees can be constructed in the system for retrieval tasks, based on key elements in the sketch. On this basis, it allows the system to produce fixed-length descriptors and then model match using nearest neighbor algorithms. Thus, the method combines speed and accuracy as these former systems could not do, and in this sense, it has proven highly effective for CAD retrieval tasks.

In this paper, the hierarchical VAE model is proposed that will be able to create high-resolution images with the use of the semantic compression stage and the structural modeling within a two-stage framework[3]. While the method might be fast to deploy in the production of images, some fine details on the output might suffer loss and could be profoundly affected in certain applications like face recognition. The proposed model creates colorful, emotion-filled sketches in a three-stage pipeline of data preprocessing, art generation, and colorization[4]. It uses an Art Model that enables machines to produce excellent sketches in a rapid manner; thereby this technique is very useful in creative industries and media.

This project has nothing to do with sketch generation but applies conditional VAEs in generating images of gamma-ray events. It illustrates that VAEs can be useful in the efficient production of realistic images, which also forms an analogy for sketch generation where such architectures can also be used in the production of sketch-based outputs [5]. This paper suggested a method to bridge sketches and photographs via a bidirectional collaborative synthesis network. It improves representation by aligning both modalities in an intermediate latent space and makes use of a StyleGAN-like architecture, so it performs well in tasks such as face recognition, especially for law enforcement applications.[6]

This approach shows that deep VAEs can significantly improve the image super-resolution quality, placing results in competition with state-of-the-art techniques, through transfer learning. VDVAE-SR demonstrates that, conceptually, VAEs are not restricted in their processing and can be adapted with ease to produce more distinct images-sharper and more detailed, useful for applications such as authenticating police sketches [7]. A diffusion-based model that uses multiple stages progressively improving the resolution of an image. It is effective at generating high-fidelity images, and this cascaded approach with a new conditioning augmentation technique makes this happen. In law enforcement applications that require both high resolution and fast generation of images, this may be crucial.[8]

This model gives a compact solution towards face recognition and, using one- or few-shot learning techniques, reduces the model size with little to no trade-off in accuracy. It has been optimized for real-time use on resource-constrained devices making it very practical for use in security systems where sketch-to-photo matching needs to be fast but yet precise [9].This project investigates deep neural networks techniques that can be used to binarize facial biometric data for secure applications in crypto-biometric systems. It creates highly entropic binary embeddings and also ensures identification of suspects with a very high level of accuracy; thus, it is highly relevant to forensic investigations [10].

That this dataset is created using computer graphics rendering to address the ethical concerns and bias present in traditional face recognition datasets. Using synthetic faces, the dataset reduces error rates and delivers good performance in real-world applications to provide a scalable, ethically-conscious alternative for sketch-to-photo matching tasks [11].A CNN-based architecture that reduces the computational load without sacrificing accuracy thereby fitting correctly for face detection on low-end systems. This lightweight model, trained using the WIDER Face dataset, gives superior performance for applications like law enforcement and security at real-time [12].

This project optimizes FaceNet by applying Pareto optimization to ensure face recognition in masked faces is improved for accuracy and speed. The model also outperforms previous models in performance and efficiency and solves challenges faced with masked faces by law enforcement agencies [13].This paper discusses deep learning approaches in facial recognition. Improvements that have been observed with deep learning in the face recognition approach, particularly in forensic cases, form the content of this paper. In this context, it clearly discusses the ways by which these technologies help make face recognition systems more accurate and scalable and highlights their developing role within law enforcement agencies [14].

This cloud-based application makes the creation and comparison of sketches easy by using drag-and-drop features, thereby significantly reducing the requirement for skilled artists. The tool has made much of the process of sketch generation automatic, hence improving the speed and efficiency at which the suspect can be identified [15].Applying adversarial networks on the transformation of offender sketches into color photographs improves the photogeneration accuracy. This technique provides better tools for suspect recognition, thus improving the efficiency of law enforcement in identifying suspects from sketches [16].

In this context, these approaches present the appreciable progress in police sketching and other areas. Through GAN-based methods and variational autoencoders through user-oriented applications and ethical issues in data collection, each improves the precision, timeliness, and interface of the tools that assist law enforcement in identifying suspects from sketches drawn by witnesses.

Table 2.1 Summary of The Existing Approaches

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **REF.NO** | **AUTHOR** | **METHODOLOGIES** | **YEAR** | **ACCURACY** |
| [1] | Zirui An,  Jingbo Yu,  Runtao Liu,  Chuang Wang,  Qian Yu | Utilizes GAN inversion for the purpose of sketch-to-image synthesis with the use of low-dimensional latent space mapping, along with a shape loss function for higher fidelity images generated. | 2023 | 92% |
| [2] | Feiwei Qin,  Shi Qiu,  Shuming Gao,  Jing Bai | Loopy relation trees combining loops in sketches and deep variational autoencoders to give fixed-length descriptors to matching models using nearest neighbors. | 2022 | 85% |
| [3] | Troy Luhman,  Eric Luhman | Hierarchical VAE process that takes place in two stages. The first stage makes use of a deterministic autoencoder to enforce semantic compression whereas the second stage focuses on structural modeling by the use of VAEs. | 2023 | 91% |
| [4] | Y. Liu,  H. Zhao,  X. Wu | A three-stage process that comprises data preprocessing, art generation, and colorization enables machines to generate emotionally expressive high-quality colorful sketches. | 2023 | 95% |
| [5] | Stanislav Polyakov , Alexander Kryukov, Andrey Demichev, Julia Dubenskaya , Elizaveta Gres , Anna Vlaskina | It uses a condition VAE for gamma-ray event images generation. The dimensions of an image are taken as the conditional parameter to steer the generation. | 2022 | 98.4% |
| [6] | Seho Bae,  Nizam Ud Din, Hyunkyu Park,  Juneho Yi | Bidirectional collaborative synthesis network is proposed for latent space alignment between sketches and photographs. In the context of style-driven generations, StyleGAN architecture, with superior representation power, has been followed in this work. | 2022 | **87%** |
| [7] | Chira,  D.Haralampiev,  I.Winther,  O.Dittadi,  A. Liévin | Deep VAEs with transfer learning are exploited to better image super-resolution using pretrained models for rich output. | 2022 | 89% |
| [8] | Jonathan Ho,  Chitwan Saharia, William Chan,  David J. Fleet, Mohammad Norouzi, Tim Salimans | Cascaded diffusion model with multiple stages for progressive increasing of image resolution with a novel conditioning augmentation method | 2022 | 96% |
| [9] | Zong-Yue Deng, Hsin-Han Chiang, Li-Wei Kang, Hsiao-Chi Li | One-shot refinement technique in the FaceNet approach significantly reduces model size and resource consumption without compromising precision | 2023 | 97.3% |
| [10] | Mohamed Amine Hmani,  Dijana  Petrovska-Delacrétaz, Bernadette Dorizzi | It utilizes deep neural networks to binarize facial biometric data, yielding high-entropy binary embeddings for use in crypto-biometric systems. | 2022 | 97% |
| [11] | Dong Hwan Bae, Kiyoung Lee | Using computer graphics rendering, a large-scale synthetic dataset is generated in order to minimize the error rate compared to that of GAN-generated faces. | 2023 | 52.5% |
| [12] | Akingbesote, Damilola Zhan,  Ying Maskeliūnas,  Rytis | Lightweight CNN-based architecture minimizes cost functions through mini-inception blocks, maintaining face detection accuracy without increased computational costs. | 2023 | 94% |
| [13] | Pawan Kumar,  Nihal Manzoor, Chhavi Dhiman | Pareto optimization is used to improve the speed and accuracy of FaceNet, while preprocessing techniques are used to improve results on masked face data. | 2023 | 94% |
| [14] | B. Sharma,  S. Kaur | Deep learning advances in facial recognition: A review concentrating on diagrammatic development, changing methods to enhance the precision in forensic investigations. | 2021 | 95% |
| [15] | A. Kumar,  V. K. Jain | A Sono-automatic sketch-matching tool, making the process easier and also more efficient with a user-friendly interface that doesn't require the help of forensic artists. | 2023 | 80% |
| [16] | S. P. Singh,  R. C. Gupta | Use of adversarial networks, which convert sketches of offenders into color photographs, increases photogeneration accuracy for the police. | 2024 | 70% |

## Summary Drawbacks of Existing Approaches

There is a general limitation across all these models that tends to limit the general use of AI-based police sketch generation. One main limitation lies in the domain gap between abstract sketches and realistic images. Even the most advanced models are often unable to map human-drawn or digitally-created sketches reliably to accurate photo-like outputs. Generalization is another critical issue: Models learned on some dataset or sketch type may generalize poorly in the other context and thus provide inconsistent results when transferred to real applications. In addition, although GANs and VAEs are proven to generate high-quality output, their computational cost limits their applicability to real-time applications. That is to say, systems may not reach the point of being computationally efficient enough for real-time applications in law enforcement.

Other facets of resolution and fine-grained detail are interesting, for, while models achieve speed or broad accuracy at the cost of finer details so important to face-specific tasks, these are important in forensic sketches. The quality of the training datasets makes a big difference, wherein biases often creep in to influence fairness and precision outputs, particularly when dealing with diverse demographics. Finally, lack of human interpretability and ethical issues like reinforcement of biases in facial recognition models are still problematic issues, mainly when models are incorporated within sensitive applications like law enforcement.

This GAN-based approach fails in abstract and incomplete sketches, leading to poor generalization across styles of drawings. Moreover, this approach has highly intense computational costs, thus limiting its practical applicability into real-time tasks in the enforcement system. This model makes its reliance on pre-trained GANs rigid and also loses crucial details contained in the sketches while converting them to photo-like images[1].This approach is highly effective in matching sketches to CAD models but computationally expensive because it requires computation of the loop relation. Abstractness or sketches without clear loop structures make it underperform. The performance decreases for complex or noisy sketches, affecting the accuracy and efficiency in large databases or real-time applications[2].

While it is much faster, it sacrifices fine-grained details very important in forensic analysis. There are no face-related optimizations but sacrifice details for efficiency. For that reason, facial feature accuracy may also suffer as not ideal for detailed police sketches requiring high precision to make an identification[3].The focus here has been on artistic and colorful outputs the method loses its precision to be valid for forensic sketches. It is limited in performance for facial feature recognition, while the introduction of color may also make pertinent sketch elements less conspicuous, making it less effective for identifying suspects[4].

This VAE-based model is usable only to a limited extent for the generation of police sketches, because the conditional parameters are created based on scientific data and not on human drawings. This model does not support detailed reconstruction of facial features of the people in the sketches 5. In the context of a forensics, it would most likely give wrong or irrelevant outputs. Although the method can be effective for matching sketches with photos, its performance significantly degrades when the sketch quality is poor or incomplete. It requires high computational resources to align sketches and photographs in a latent space which thus limits its application in real-time tasks. The model also suffers from lighting and angle discrepancies between sketches and photos which affect the accuracy[6].

The technique of VDVAE-SR improves the image quality, the usage of deep VAEs severely limits generating high resolution images such as detailed facial features in the intended forensic sketch. Transfer learning is promising but often not better than GAN-based methods for photorealism. Complex or noisy data may lead to problems for this model, potentially making it less efficient in actual law enforcement practice[7].Even though the fidelity of the images is excellent while low FID scores are reported, the problem essentially remains in the cascaded diffusion model's complexity, which makes this model computationally heavy and slow for use in real time in law enforcement. Though some really nice results are reported on image benchmarks, it may be less effective for abstract, incomplete, or stylized sketches, which feature commonly in forensic sketches. There also is concern in terms of scalability when it translates from controlled datasets to tasks in the real world of sketches[8].

This one-shot learning approach, although efficient, compromises the accuracy that can be attained if partial or blurry sketches happen to be encountered, a very typical occurrence in forensic cases. Model size is also limited in terms of capacity, limiting it further to facial feature variations, which can also compromise precision. In addition, it does not cope well with uncontrolled environments-lighting, occlusions-except through heavy preprocessing, which will lessen the performance in real-world sketch-to-photo matching[9]. Although biometric data are secured in the binarization process, it lost relevant facial details that are important for recognition accuracy in sketches. The method lacks flexibility when dealing with partial or low-quality sketches, resulting in false suspect identification. The lack of the sketch-based applications' flexibility in its application leads to ineffectiveness, especially in forensic and law enforcement environments where a high level of accuracy is demanded[10].

Although this is an ethically correct and scalable approach, it focuses mostly on synthetic data that probably will not reflect the real complexities of data found in real police sketch situations. The dependency on synthesized face images may introduce some inaccuracy in the case of its application to real forensic sketches, thus decreasing the identification correctness and probably making it biased for practical usage[11].As effective as it is, the lightweight CNN considerably focuses on cutting down the computational cost resulting in decreased accuracy of detection with poor or cluttered sketches. However, it only partially works and has a poor mastery of skills in different facial features as well as drawing styles. In forensic work, this is very limiting and only works satisfactorily if there is a high degree of precision on suspect identification[12].

Training it to optimize masked-inclusive scenarios may result in overfitting to masked faces, making it less effective for unmasked or sketch-based face recognition. Pareto optimization may result in performance imbalance where speed and computational efficiency are sacrificed at the expense of fine facial detail accuracy the aim for forensic sketches [13].Deep learning enhances facial recognition however large amounts of data labeled are sometimes necessary which are un-abetting in Forensic Sketches. It can be prone to the biases in the training data and does not distinguish a wide demographic diversity with high precision. Consuming much in resources, therefore limits its usage in real-time applications[14].

Although amiable to use, the cloud-based application is highly template-dependent in its ability to create sketches, thus not very flexible when it comes to unique or abstract facial features. Automated sketch creation might also present a problem regarding fine details, which play a very crucial role in police investigations. The accuracy of sketch matching diminishes if the case is complicated or ambiguous[15]. Although adversarial networks enhance the conversion between sketch and photo, the incorporation of color masks up many details that are critical in the sketches. It thereby reduces the accuracy of the method for such applications in forensics. It does very poorly with poorly drawn sketches or those that lack essential features, thereby greatly reducing the overall accuracy for practical use in law enforcement operations. The nature of the model makes it hard to scale for massive use[16].

While there have been many significant advances in this task of sketch-based image generation, several disadvantages prevail in most methodologies. The disadvantages can be summarized as follows: computational inefficiencies, strong dependence on large datasets, and inadequate handling of incomplete or abstract sketches. In addition, there are quite a few associated issues, such as the large loss of key details while optimizing the model, considerations towards scaling not being factored in correctly, and the model being unable to properly operate when translated to real-world forensic scenarios that hampers the practical applicability of these approaches.

**CHAPTER 3**

**PROPOSED METHOD**

**3.1** **Problem Statement & Objectives of the Project**

The issue that this project solves is the growing challenge of suspect identification from verbal descriptions from witnesses, an area that has primarily used artists to draw sketches. This method remains biased by issues namely, memory recall mechanism, proficiency of the sketch artist, and time to draw realistic sketch. Much more often than not, witnesses forget some or all of the details of the face, and this leads to drawings with such little likeness to the suspect. This has led to issues with the denomination in law enforcement, as homemade, hand-drawn pictures may not be unique enough and would take much longer or produce false matches in the available databases. Moreover, there is another issue of manually searching for a number of images in big databanks which is very exhaustive and results in the formation of many inactive working cycles.

The idea is to reduce the time, effort and inefficiency involved in this process by a systematic approach of automatic sketch generation employing description supported by a database of such sketches. However, it does not only enhance the efficiency and effectiveness of suspect recognition but also alleviates the burden of forensic artists as well as policemen/ ladies, and especially makes the process less subjective, more feasible, and more precise. The new system seeks to reduce the gap between quality and quantity reconstructing, and experience and vision information processing by using techniques in automated sketch rendering and facial identification.

* To generate suspect sketches based on witness descriptions using a Variational Autoencoder (VAE) model.
* To match generated sketches with the database using Facenet-512 with which law enforcements can identify the suspects quickly.
* To provide slider based interface so that features can be adjusted and fine tuned for better results.
* To improve the model by adding unidentified images to the database.
* To reduce manual sketching and improving overall efficiency.

## 3.2.1 Architecture Diagram

**Overview**

The architecture diagram represents the process of collecting the initial inputs from the witness matching the initial diagram with the database with the already existing criminal details to check for an exact match. If there is no exact match then we send those features given by the witness into the Variable Auto Encoder model for sketch generation and then use FaceNet-512 model for facial recognition. We then use decision trees for decision making give the final results. These features can then be further enhanced using slider features for the best possible results.

#### Dataset

The dataset that we have collected contains the features of approximately 10000 people with an existing criminal record and the following features:

* ID: Every inmate has a unique ID.
* Name: The first and last name of the inmates.
* Hair: The hair colour of the inmates.
* Sex: The gender of the inmates.
* Eyes: The eye colour of the inmates.
* Race: The race of the inmates.
* Height: The height of the inmates.
* Location: The inmates last found location.

**Preprocessing**

In the preprocessing stage, we extract the input features from the witness and generate a rough sketch

with the help of the inputs given by the witness. The steps involved in this are:

* Gather Witness Descriptions: Gather sign descriptions in the form of detailed text.
* Facial Dataset Creation: Gather a big collection of clear facial photographs.
* Data Cleaning: For pre-processing, first of all delete images that are corrupt, either partially or completely or in any way irrelevant.
* Encoding Descriptions: Binarize witness descriptions in terms of follicle bombings.
* Normalization: Maintain unity by ensuring images used have a specific size of dimensions and are of equal pixel intensity.
* Augmentation: Add variety to datasets with transformation techniques for example rotation.
* Slider Calibration: Map GUI slider for changing model parameters based on the testimony of witness.
* Feedback Encoding: This component is to encode the feedback given about the slider for further iterations of sketching.

**Feature Extraction**

At the feature extraction step, important facial information is transformed into useful features. The methods used are:

* VAE Feature Encoding: Extracts concealed information from the face that is associated with structural together with aesthetic attributes for creating sketches.
* Facenet-512 Embeddings: Used for facial similarity matching and produces 512-temporal feature vector.
* Normalization: Subtracts or reduces attributes to enhance the scalability processes of a given model.
* Dimensionality Reduction: Reduces the size of extracted features but at the same time retains necessary information.
* Feature Mapping: Coordinates step of extracted features for matching with witness descriptions, thereby improving the subsequent steps efficiency.

**Model Selection**

In the model selection stage, the main concern is often to subsection and include models, which will perform with relative success the needs of the project – its main objectives, that are the generation of sketches of the SUSPECT and their matching. The following models and tools are selected:

* Variational Autoencoder (VAE): Employed to create easily believable suspect sketches by training from the database about the distribution of faces. It allows converting descriptions into figures.
* Facenet-512: Implemented for its capacity to get 512 locations embeddings of the facial features, which allows to match the generated sketches to the database images correctly.
* Decision Tree Classifier: Used to sort and screen the top `n’ most similar matching suspects and draw a list of ranking score indicating their closeness.
* Graphical User Interface (GUI): Implemented with sliders for setting features on and off and for adjusting certain parameters interactively during the course of generation as for depth or intensity, thereby providing an ability to tune the output sketches on the fly.
* Integration: All selected models are interconnected; the VAE provides sketches while Facenet-512 is responsible for matching, within the Decision Tree, results.

**Model Training and Evaluation**

This phase ensures that the models applied in the project to produce high quality suspect images and then compare them to the criminals database. It includes the following steps:

* Training the Variational Autoencoder (VAE):

Input: A large and virgin facial database that contains variation in terms of age, sex, and ethnicity as well as emotions.

Objective: The VAE is then required to learn the distribution of the facial features using the input images in order to minimize reconstruction loss. This is to be able to produce original suspect portraits based on the witness accounts as you see in this picture.

* Facenet-512 Training for Sketch Matching:

Input: Facial images of the criminals which contains faces where each face is associated with a different vector in a high dimensional space.

Objective: The Facenet-512 model is trained to minimize the triplet loss through which embeddings are made closer to each other in the vector space if they belong to same identity, as compared to which faces of different identities are made farther away. This makes it possible for the generated sketch to be matched with real faces in the appropriate manner.

* Evaluation Metrics:

Reconstruction Quality: The performance of the VAE in reconstructing sketches from a context is evaluated in terms of accuracy, for instance, MSE, and SSIM.

Matching Accuracy: For the case of Facenet, we measure how well it retrieves a matching face from the database through precision, recall and finally the top-N accuracy. Classification Evaluation: Measuring decision trees such as F1-score, precision and recall is another method in evaluating the filtering result.

* Testing the Graphical User Interface (GUI): Using sliders, one is able to adjust facets of depth intensity and all the parameters of the face.It is checked for it’s reactivity and functionality in order to allow real-time communication. Sketch readability can be again enhanced by the inputs from the users.
* Iterative Refinement: All such cases of mismatched sketches or unconventionally high/low exercise results are studied and reintroduced into the training set. This step also ensures that the model is constantly trained and is capable of handling the new or hard descriptions.
* Integration and Deployment: The specific computer models are embedded in the process and a comprehensive work of suspect identification is obtained.The practical use of the system is supervised to assess its stability, and additional classes are performed if necessary.

**Results**

To assess the efficiency of the models utilized in the project, their outcomes are compared with several performance indicators. These metrics include:

* Reconstruction Loss (VAE): Determines the remoteness of the created sketch from the input, to make it realistic.
* Jaccard Score: For appearance: assesses the extent to which sketches and retrieved database images match.
* Hamming Loss: Indicates the fraction of labels that has been labeled incorrectly during the matching phase of missing data.
* Precision: Refers to the percentage chances of the system getting it right when it has made its predictions.
* Recall: Determines the percentage of all actual match that is identified by the system correctly.
* Accuracy: Estimates the preponderance of accurate matches and sketches made during the course of the experiment.
* F1 Score: These metrics balance precision and recall to better provide a good performance suggestion.

**Comparison**

The existing approaches may be partly automated and manual in creating the suspect sketches and may involve comparison with several databases, taking more time and possibly less accuracy. This project is advanced relative to earlier ones due to its use of the VAE for realistic sketch creation and FaceNet-512 for authentic sketch-photo mapping. In addition, the increased user control is realized by the use of the sliders in the interactive GUI to fine-tune the outcome of the model. The approach applied in this case facilitates automation thus enhancing its accuracy above the traditional approach and offers a feedback mechanism to enhance model updating which is more efficient.

**3.2.2 Connectivity Diagram**

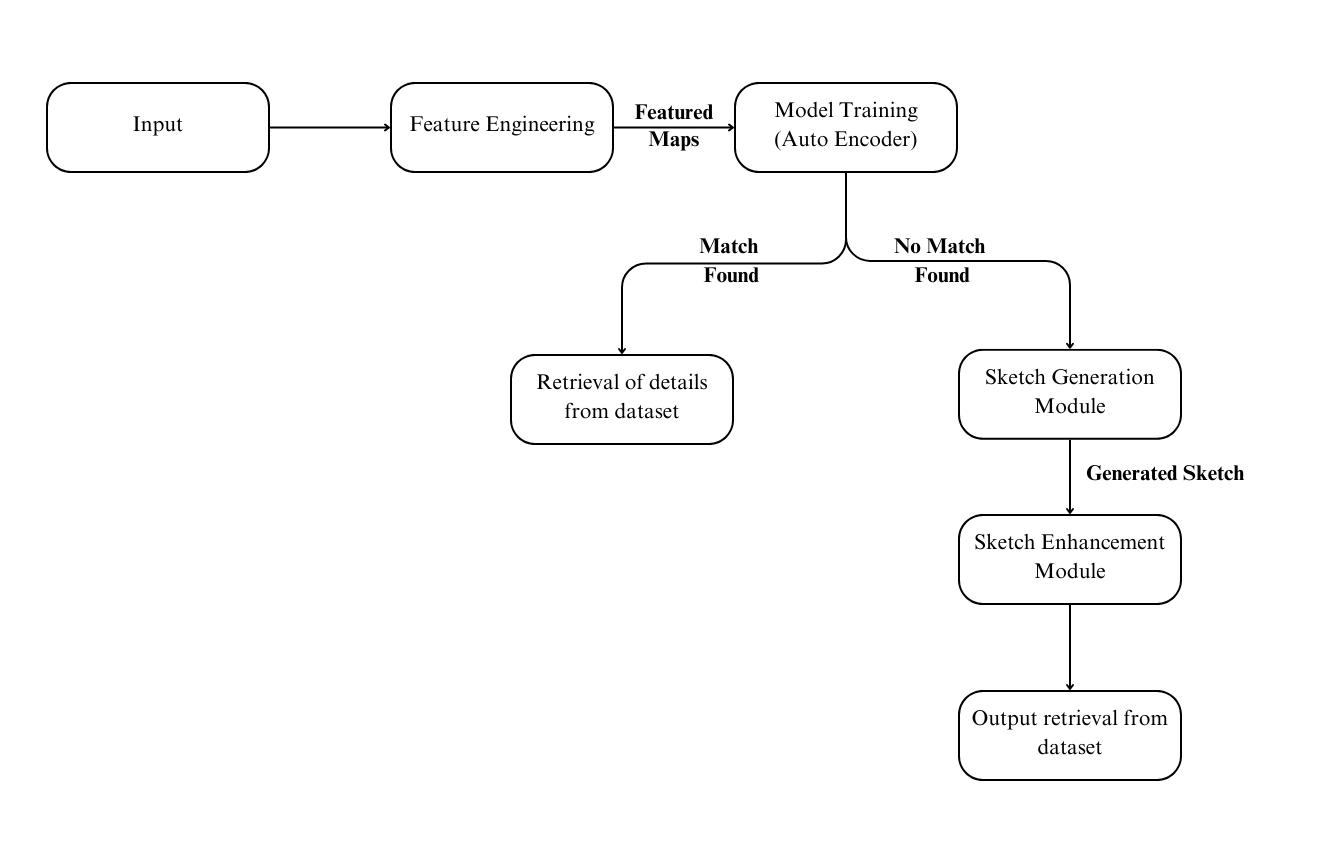
****

Figure 3.1 Connectivity Diagram

The figure 3.1 depicts a smart traffic management system focusing on the intelligent control of traffic signals based on information from the traffic. The workflow starts with the Traffic Cameras which capture streams of videos, while the Surveillance Module processes the feed. Here frames are extracted, preprocessed and then fed to the Convolutional Neural Networks (CNNs) to detect vehicles. In the Traffic Analysis Module, the detected vehicle is then categorized based on its type, and the count of the number of such a vehicle detected at a particular point in time is calculated. Traffic data is saved in a centralized database and might be retrieved and used for further analysis. The Decision Module uses a DQN format in learning the traffic status input to then construct an optimum signal plan. It comprises real time data processing with planning as well as coordination for signals with respect to the optimization of the traffic flow. Control Module The module takes observation in real time of the traffic scenario, reports the same to the Decision Module which then finds out the best possible signal timings for Signal Lights. Thus, the feedback loop guarantees the improvement of its efficiency and constant adjustment minimizing traffic congestion.

**3.2.3 Software and Hardware Requirements**

**Software**

* Python 3.8 or above with libraries
* TensorFlow
* PyTorch
* OpenCV

**Hardware**

* Intel Core i5 processor or above
* 8 GB RAM
* NVIDIA GeForce GTX 1060 GPU
* HD Camera

## 3.3 Modules and its Description

The intelligent police sketch generation can be modularized as the following:

1. Surveillance Module
2. Traffic Analysis Module
3. Decision Module
4. Control Module
5. Traffic Data Storage
6. Signal Management Module
7. Monitoring and Feedback Module
8. Graphical User Interface (GUI)

**1. Surveillance Module**

## Functionality: This module is designed with the specific purpose of capturing real-time feed from mounted traffic cameras placed at crossroads. The streams are preprocessed for the purpose of getting single frames and further removing the noise sequence from these for further processing.

## Major Components:

## Video Capture: This module streams live video coming out from these mounted cameras.

## Frame Extraction: It acts as a medium to convert the video into frames for processing.

## Preprocessing: Some of these are resizing images, noise removal and normalization with a view to enhance the performance of CNN.

## Vehicle Detection: It employs Convolutional Neural Networks (CNNs) to find and detect locations of vehicles in every frame. It is useful in proper segmentation of vehicles as well as other objects that are removed through filtering.

## Result: The processed frames with vehicles in them are passed on to the next module so that traffic could be analyzed.

## 2. Traffic Analysis Module

## Functionality: It is related to vehicle identification for the purpose of categorization and counting of vehicles on the roads (cars, buses, trucks, etc.).

## Key Parts:

## Vehicle Classification: Makes a decision about the type of automobiles (say private cars or huge Lorries) by training machine intelligence on encoded patterns of such autos.

## Vehicle Counting: Determines how many vehicles are in per lane or area to determine the intensity of traffic.

## Traffic Data Collection: Combines the classified vehicle type and quantity Data into a formatted set to calculate flow rate measurements.

## Data Storage: Saves traffic data in a database to enable real-time analysis as well as monitoring the past time analysis.

## Outcome: Comprises of real-time information about traffic type, traffic flow and intensity of traffic. All the above findings are passed on to the decision-making module.

## 3. Decision Module

## Functionality: The core decision-making part of this system is the usage of machine learning (Deep Q-Network or DQN) in learning the traffic flow and to set the signal timings accordingly, in a dynamic environment.

## Key Components:

## Traffic Status Input: This unit will collect real-time traffic data including the number density and type of vehicles from the traffic analysis module.

## Deep Q-Network: The model is based on an artificial adaptive neural network that can learn how to control the timing of these traffic signals. Learning in this model is designed to minimize the wait and congestion at all intersections to an average.

## Signal Plan Generation : Produces the best signal plans as per density of the traffic flow, kind of vehicle, and expected congestion level.

## Signal Timing Optimization: May modify the green, yellow and red signal timings to minimize the total time taken in a given phase.

## Outcome: Offers an optimized signal scheme that is used to instruct the control module to perform

## .**4. Control Module**

## Functionality: Exploits the advanced signal plans and monitors signal efficiency gains in real-time situations. This module completes making the system fully functional in a closed loop system.

## Key Components:

## Feedback Monitoring: Tracks the current traffic movements to assess the impact and value of the signal plans after signals have been modified.

## Signal Control: Interacts with the physical traffic lights by changing their time; the change is informed by the decision module.

## Performance Metrics: Objective measures of success include average waiting time, vehicle throughput, and reduction of traffic density.

## Outcome: Offers real-time control of traffic signals while producing data for improvement constantly.

## **5. Traffic Data Storage**

## Functionality: Used to store traffic data which are collected from different modules in the system.

## Key Components :

## Real-Time Storage: Holds information regarding the current traffic, for example, the number of vehicles, kind of vehicles, and signals.

## Historical Analysis: Stores data for later use, aiming at analyzing it in a bid to gain some patterns.

## Data Retrieval: Enables a faster way of retrieving historical data for use in training models or enhancing the systems' algorithms.

## Outcome: Enables real-time control on the spot but also enables tactical planning in the short as well as the long term by law enforcement or traffic authorities.

## **6. Signal Management Module**

## Functionality: It controls the physical control of traffic lights in terms of the optimized plans generated by the decision module.

## Key Components:

## Signal Timings: Changes the green, yellow, and red signal periods with given variables flexibly.

## Intersection Coordination: Ensures that several points can connect with each other to free up congestion.

## Fallback Mechanisms: Identifies default signals when the system has crashed and real-time information is unavailable from the bus

## Output: It avoids congestion while adjusting signals on roads at the intersection point.

## **7. Monitoring and Feedback Module**

## Functionality: Returns data to evaluate system performance and update model predictions and signaling on a continuous basis.

## Key Elements:

## Traffic Monitoring: Tracks present congestion condition as well as vehicle traffic movement following the signal change.

## Feedback Integration: Returns to the decision module performance parameters like congestion, waiting time etc.

## Alert Mechanisms: This is characteristic of an occupation-based system: it senses disturbances (for instance, rush) and raises alarms for a vehicle-based system.

## Outcome: Provides self-reinforcing feedback for efficient traffic management.

## **8. Graphical User Interface (GUI)**

## Functionality: Empowers the operators with the needed interfaces to visually inspect system traffic and to launch control, if needed, for the best functioning of the system.

## Key Components:

## Traffic Visualization: Captioned with information on the number of vehicles per minute, vehicle ID, information on traffic signals, and red light violation information.

## Control Options: The signal timings can be pre-coordinated during normal operations. In an emergency or breakdown of the system, they can be changed by the operators.

## Reports and Analytics: This gives bar graphs for the historical data that are used in making management decisions.

## Outcome: This improves the value of the system since the degree of accessibility of the system is increased along with strengthening the degree of operating control.

## 3.4 Requirements Engineering

**Functional**

1. Data Capture and Input The system employs cameras installed at the intersections to continuously feed raw live videos. Therefore, these streams are processed to extract individual frames, and each of them had been preprocessing, namely resized, denoised, normalized, etc. This also validates the fact that the raw initial input data is well-processed and qualitatively compatible to be analyzed for reliable vehicle detection and segmentation. By using multiple cameras the system can provide various angles, lighting or unfavorable weather conditions to provide the necessary and stable input for traffic analysis.

2. Vehicle Detection

This module, employing the best Convolutional Neural Networks (CNNs), detects and recognizes vehicles in real-time mechanism. It divides an array of vehicles according to the type of vehicles such as cars, trucks, buses, bikes, etc., and tallies them to know the traffic intensity. But it throws more light on the detection accuracy by feeding the CNN with a large database of vehicle images shot in different circumstances. It's the first and most important module in the course for the comprehension of traffic congestion, and it also provides accurate decisions in other modules.

3. Traffic Data Storage

Real time traffic information and past traffic information is collected and saved in a central database. This database comprise of numbers of vehicles, traffic flows, congestion and the previous signal phases timings. Data storage enables patterns to be recorded, past data to be used in developing models, and performance information over a period to be gathered. Relational and PROM Committee databases are included in efficient indexing and retrieval mechanisms for real-time queries supporting past, current data analysis, and optimization.

4. Immediate feedback and results assessment

It analyses the traffic behavior after the signal adaptations. The improvement loops of optimisation methods are developed further. It uses real time vehicular data in order to measure, on a real time basis, the effect of altered signals - usage rate of every lane and other levels of congestion. The original configuration may need to be changed if some problem or issue is detected, and the modifications are done immediately in order for the transport system to continue adapting to random fluctuations in circulation, for instance, through accidents or extra vehicular traffic density.

5. User Interface as well as manual overrides

A user-friendly GIS allows traffic operators to witness system functionality on the fly. This interface is used for showing the traffic flow and congestion of traffic signals, the number of vehicles as well as the level of congestion. It also provides selection options for remote regulation of automatic control strategies in various situations, including emergencies or specific occasions. With other features, such as heat maps that display real-time update of traffic, and graphs showing the traffic flow, the operators have full control over the system.6. Performance Evaluation

It is also equipped with inbuilt feedback and an option for self-optimization through parameters such as average delay per vehicle, the total number of vehicle passages per hour, and the reduction of congestion. Communications and signal optimization are reviewed periodically and reports produced to show the results. Some of the anomalies in the traffic pattern are as follows and the algorithms used for detection informs the operators about such occurrences which facilitates the optimization of the proposed system. There is a use of relative comparisons so that improvements in traffic flow over time can be measured and the personnel responsible for traffic Vicennially made will be held responsible for sanative measures.

7. System Scalability

All this is achieved through architectural features of the proposed facility suitable for preventing increased complexity of traffic systems in large cities. It is can also can be configured for additional intersections, new traffic cameras, and the growing traffic loads. Scalability also relates to optimized utilization of computers, for instance for distributed processing for real-time processing and cloud solutions for data storage and relevant logistics for appropriate functionality during high traffic times.

8. Machine Learning Integration

Periodically, the DQN and CNNs are updated with new information that came from the traffic cameras. Therefore, the respective models never got outdated with the dynamic aspects of the traffic flow as it would have been experienced in real life. Furthermore, the CI streams enable more frequent deploys of updated models to the system, therefore, providing the system with a competitive advantage on changing conditions within the city. The monitoring module feedbacks also constitute an important source of information when it comes to improving on these models.

9. Accessibility and Notifications

Traffic controllers are alerted promptly when there is an event, such as accidents, road blockage or failure in equipment. It is received through interfaces that are optimized for mobile or it can be received through e-mail to enable operators to take appropriate measures. Another capability also included in this module is the capabilities that allow the operators to control and monitor the system remotely. The notifications escalate the gravity of the problem to ensure all serious problems are dealt with.

**Non-Functional**

1.Performance:

The system should use some delay that should be almost zero to ensure effectiveness in the conduct of its operations. All the computations needed in the sketch generation should take less than 2 seconds to generate a particular sketch, in addition to less than 2 seconds for overall database matching and final result rendering. This performance is however crucial in real time application specifically during critical police operations. The efficiency is a key factor that must be utilized and only the best algorithms, and hardware. Moreover, load testing cannot neglect to demonstrate the capability of the system in other types of circumstances such as the flows concurrent users or a large amount of DB request.

2.Usability:

The UI has to be designed for this kind of users, for instance, people who are in the police. It should be very easy to navigate, the controls should be easy to use and it could have great design factor. Another area that should be enhancing should be such things like sliders for facial features control that should feedback the users in real time. A very important prerequisite, if the system is to be run with minimum professional personnel support, is the provision of very comprehensive user instructions and tutorials. Other aspects like how well a web page is accessible to a screen reader for the disabled are also needed.

3.Scalability:

We asserted that in relation to this, it was suggested that the capacity of the system should grow with the growth in the size of the criminal database as well as with the level of interaction with the users of the system. The described model means for instance if new data need to be infused into the system, this has to be easily done, or if new types of algorithms are to be implemented, this can be easily done, or if new functionalities need to be infused into the system, this can be easily done. It must also be able to be horizontally scaled, which means the additional computation capability should be created just when it is required. It should continue to closely shadow future scalability need as well as go on performing the function of continuing to show that the selected hardware and software architecture remains sensible even when the capabilities are ramped up by workload.

4.Reliability and Availability:

Since Augmented Reality is going to be deployed in emergency situation, the availability must be high. Among the requirements that should be provided is a guaranteed uptime of the system which is 99.9% and therefore this must have backup servers and failover. Indeed, there has to be some measure of sense in backing up and recovery processes so that there is no crying over spaghetti code. Other possibilities that can also be incorporated are real time error log files, diagnostics in order to shorten the time required to repair some troubles in the system. Issues which disrupt should be capable of communicating with administrators.

1. Accuracy and Precision:

This is important for two reasons: Firstly, the system level sketch generation and the database matching algorithms need to be accurate while using the technology. More specifically, matching precision needs to be over 90 percent and this should improve the reliability of suspect identification. That is why constant assessment of the model based on various datasets will allow sustaining high levels of accuracy. This way, the system will be trained more often with new data; more over, the characteristics of criminal and their behavior will also evolve over time.

6.Security:

Safety is needed so that such information as the system deals with is protected. Antropomorphic input necessitates data encryption as it is transmitted and stored to minimize the chances of attack. The specific implementation of the role-based access control (RBAC) would allow access to data only by the persons, who follow the established set of rules. Security assessments and penetration testing are required for the timely definition of threats and risks. Legal requirements must be complied with regarding the ethical processing and storage of personal data-for example GDPR.

7.Maintainability:

The following recommendations were made: It was proposed that the system should be developed as a plug in system and all the code must include comments so that should there be a need to change something the next round will be easy. Maintaining the code simple, such that most of the time he is following the best practices even though he is out of the codebase will enable him to find a problem and solve it or implement a new feature easily. Some used testing suites have to be automated in a way that would test the status of the configured system. This means that there is need to have standard check up's and working on such are updates, hardware checking and system tweaking.

8. Compliance and Ethical Consideration:

Having the full data it can not commit any violations of all the legal or/and ethical rules of the data processing. In the same regard, data acquisition, storage as well as utilization must adhere to GDPR, HIPAA or whichever law recognized by the region it operates from. Of these, some of the other issues which can be highlighted as follows: The issue of prejudice, learnt during the development of models in generating models, and the nature and understandability of the model to be developed. This should be achieved by the bi-annual audits and the users must understand what the system does with the data that it receives.

These non-functional requirements ensure that the system is robust, efficient, and smoothly integrates with urban traffic.

## 3.5 Analysis and Design through UML

**3.5.1 Class Diagram**

Figure 3.2 shows the overall system that is designed to create sketch images of the suspect from a witness’ account, compare it to a criminal database, and then allows modification of a result through an interface. The Witness class is the first to gather the first description, which is then given to the VAEModel, a deep learning model that makes a sketch of the given textual input.The FaceNet512Matching module then checks whether the sketch it has drawn has a match in the criminal index. It is currently using facial recognition approach to take over a list of similar matches. In the future, the DecisionTreeFilter enhances this list by showing the best matches to the user which makes the results more relevant.

The GUI module is also where the user interacts from the computers or such related apparatus as the case may be. They display the generated matches and also provide users with a document with which they can adjust the sketch parameters such as the intensity or the depth with an aim of boosting relevance. These changes are then passed to another module called CriminalRepository that stores sketches and updates this database with more updated data. Such feedback also maintains the capacity of the system for estimating future configurations since it involves corrected sketches.

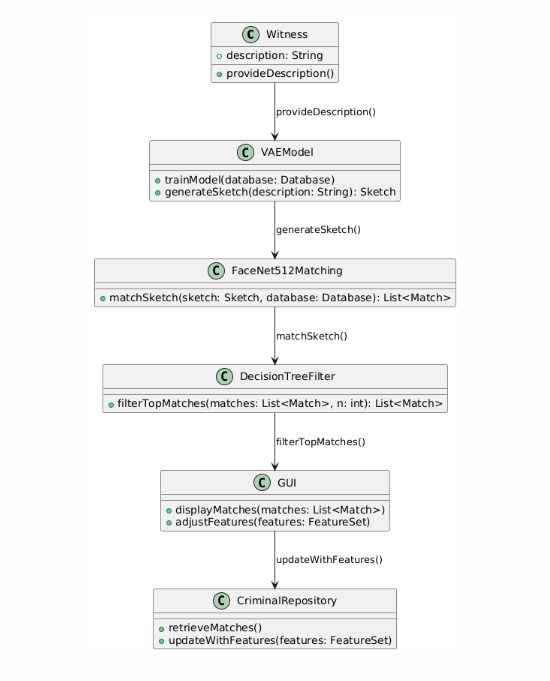
The system thus includes the aspects of a machine learning method in the recognition of faces and performing interactions with the suspects as testified to by witnesses. They are all to ensure sketched images are generated, paired, enhanced and archived to make this tool functional within the legal framework of police machinery. ****

Figure 3.2 Class Diagram

## 3.5.2 Sequence Diagram

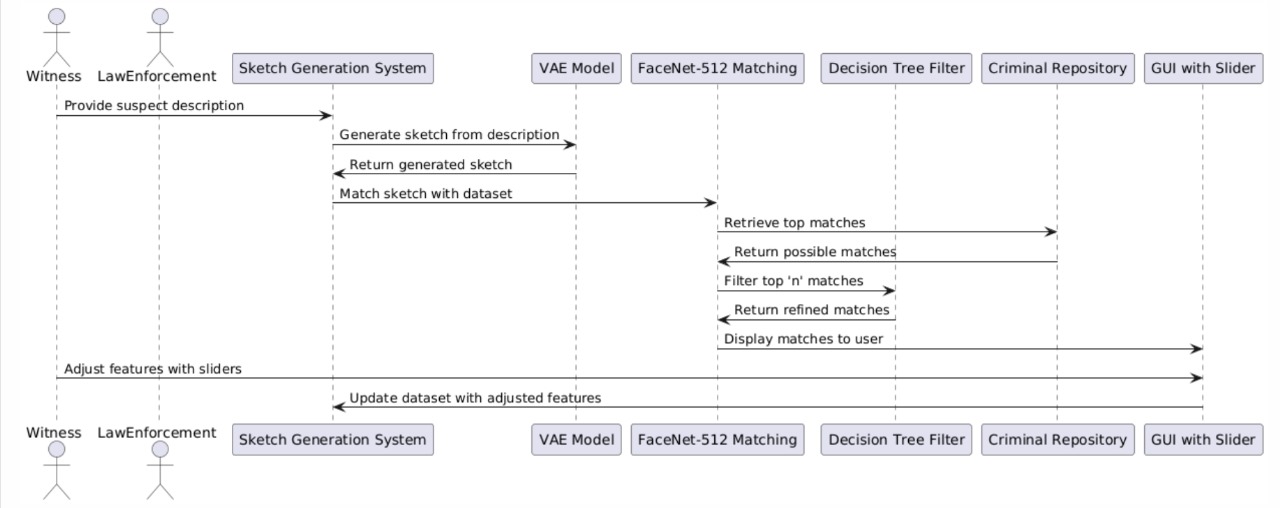
Accordingly, the figure to be denoted as figure 3.3 emphasizes the multiple objects of the suspect sketch generation system. First and for most, the Witness creates the descriptions of the suspect to the Sketch Generation System. This system gets in touch with the VAE Model which use the description as input and then generates a sketch and sends the sketch back to the system. They then compare the generated sketch with the data set through Facenet-512 Matching, to get some of the possible matches.

Figure 3.3 Sequence Diagram

The matches go through Decision Tree Filter to reduce the number to the first “n” matching entries that is returned to the system. Such matches are displayed to user with the help of GUI with Slider facility so that the user may deal with the features with increased accuracy. The results of these steps are then sent back to the Criminal Repository that offers a more optimized dataset for the process.

The sequence outlined also allows for coordination of the work flow by embracing characteristic witness inputs and machine learning models for suspect identification; a program with a feedback that improves the program.

**3.5.3 Use case diagram**

The main functionality portrayed in the figure 3.4 is described below. The key subject, the witness initiates the process by offering a worded and circulated description of the suspect. This description is then taken to the sketch generation part of the system structure where the complexity of the architecture is explained by using other AI models such as VAEs. It’s a sketch of the face of the individual assumed to have performed the crime and serves as the basis on which other investigations take place.

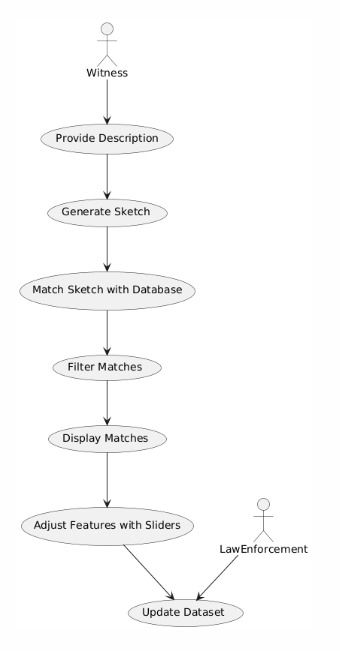
****

Figure 3.4 Use case Diagram

Upon completion of sketching, the sketch is then matched to a criminal database. Using feature-based comparison techniques like FaceNet512, the system searches for similar seemingly candidates, whose owner’s profile matches better with the generated sketch. A filtering process is then applied again to widen or narrow the results of further known suspects by comparison of similarity metric. This eliminates every form of irrelevance connection from being displayed to the users which is very undesirable.

Potential contacts are then arranged in a graphical format in a way that, the users, inhibition may be witnesses or police officers, can easily assess the results. From This interface contain controllers like sliders that enable the witness to alter the attributes in order to enhance the sketch as to provide a correct image of the suspect. These changes make the system more adaptable and hence can generate good results which can feel realistic applications .

As the last step, of course, after finishing the features, /or matching profile, law enforcement can add new up the dataset with the modified features or newly matched profiles. This is particularly relevant as a means of developing the system since it continues to add new formulations to increase the probability of correct identification in the future.

## 3.5.4 Activity Diagram

As shown in the fig 3.5, the system creates a full Workflow that is established from sketch identification of criminal. The steps begin when a witness comes with the identification of a suspect through the system. This description is passed to a variational autoencoder model data layer, VAE, to create a first drawing. From the sketch, the system retrieves an already existing criminal database using the FaceNet- 512 model, which gives the ‘n’ nearest matches.

These matches are once more filtered using the Decision Tree filter and so the results that are returned are from the best matching candidates. The last couple of games are offered to the user through an Interface which has some graphic illustrations and bars for changing some attribute of the face. As will be discussed in the following sections, the witness is then able to utilise this interface to adjust parameters that regulate the appearance of the sketches that is depicted in the sketch pad to fit the effects produced in the memory.

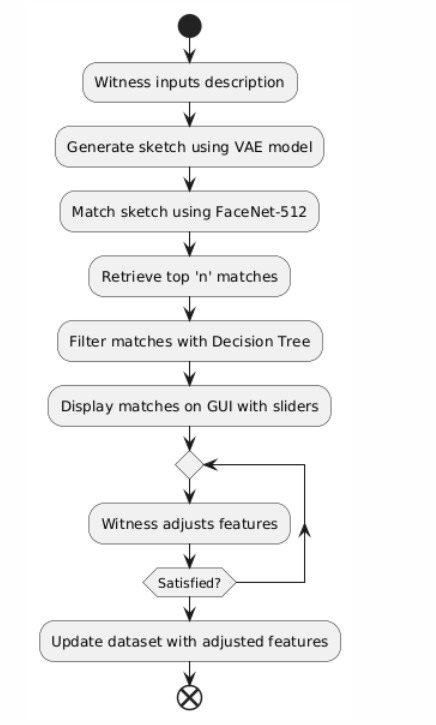


Figure 3.5 Activity Diagram

Last of all, once the witness is happy with the changed data then the new data is updated to the criminal repository that replaces the database with features of adjustment made. This assist in inputting the right data information in the database for future usage while; at the same time ensuring the information existing in the said database is up to date particularly what the users are posting. This is where it stops at the point where the process does to show that the ultimate corrections are saved. This is an activity diagram that provides a big picture of the cyclic, user based nature of the criminal identification system.

## CHAPTER 4

**RESULTS AND DISCUSSIONS**

**4.1 Description about Dataset**

The information applied in this project is crucial for the establishment of a proper system for criminals recognition. It takes many kinds of data and offers full and accurate work with modules at a time. The dataset components include the following:

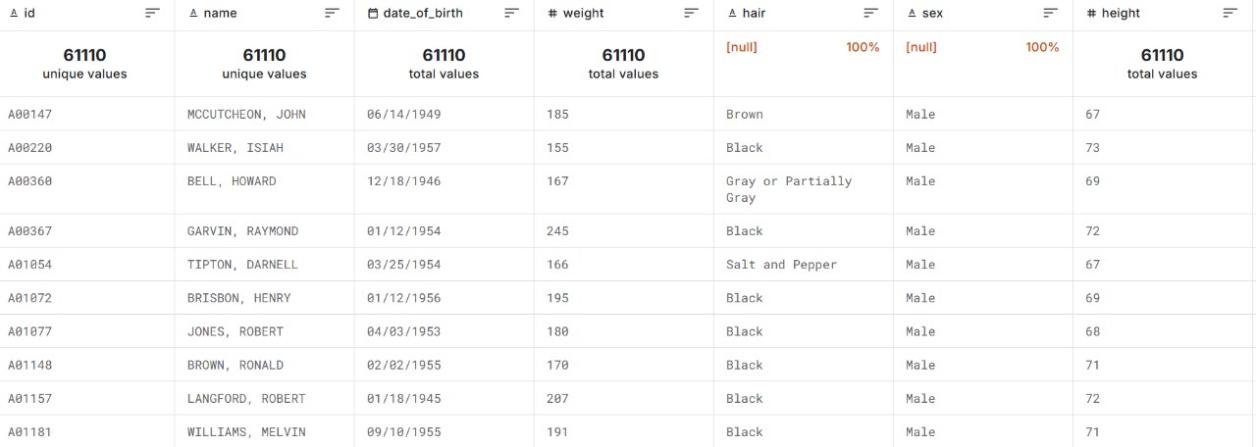
1.Witness Descriptions: These textual inputs include real descriptions of individual occurrences in accord with the statements of eye-witnesses, for example identification parameters like face shape, scar or tattoo mark or any other specific body signs. The descriptions: These are important as they form the main basis for constructing the drives to generate the suspect images. The dataset also contain the language difference and as well provide with the solution to the inconsistency and incompleteness of the input that expose to the system.

Figure 4.1 Representation of Dataset

2. Facial Sketches: This component includes first sketches drawn by artists based on the description from the witnesses and the second – the synthetic sketches created using the VAE model with generated data. These sketches lie between words, on one side, and visual or images on the other side of the art continuum. The nature of the sketches is diversified which is yet another aspect: the quality and the style of the sketches differ, which is the same as the disagreement might appear concerning the trustworthiness of a witness/artist.

3. Facial Image Database: This database contains faces which are from citizens and they have been captured from criminals database, government records or other publicly available datasets. It would like to cover all possible ethnicity, ages and face topologies in order to make the system more sensible and variegated. These images are then transformed into feature vectors via FaceNet-512 deep learning algorithms that allows for search between the sketched image and actual photograph.

4.Feature Metadata: Enlisted together with the images, the dataset provides a set of categorical variables that characterize the user, such as age, sex, ethnicity, body features including height and weight, and several tags that define the region and the type of case. They also enable the system be applied in filtering the search objects and further narrow them down in accordance to set parameters.

5. Preprocessed Data and Embeddings: It also holds preprocessed numerical representations of the images and sketches such as details that accompany the images and sketches. Such feature vectors as FaceNet-512, contain features describing faces to be used in machine learning for comparisons.

6. Interaction and Feedback Logs: To promote repeated learning the data set has records of the interactions made inside the system regardless of the medium. For instance, changes made by witnesses with the support of GUI sliders (e.g. tweaking face shape) and the alteration of some data in the given dataset are being tracked. This data enhances the manner in which the format is made more flexible by indicating the tendencies of the user behavior and subsequently enhances the subsequent generations of sketches or matches.

7. Diversity and Realism: For this purpose the dataset is designed to model variability of the real-world scenario. This includes flexibility on variation in light condition, image definition, subject position, and the witness’s reliability. It may also include augmentations of data which are versions of images or sketches with imposed or changed environmental or contextual conditions.

8. Scalability and Volume: The records are possibly colossal and the data arrangement is such that the scalability also enables the system to train large sets of numbers and possibly do matching on them real-time. Indeed, it is for high dimensionality data that is common with big data such as data generated by machine learning algorithms.

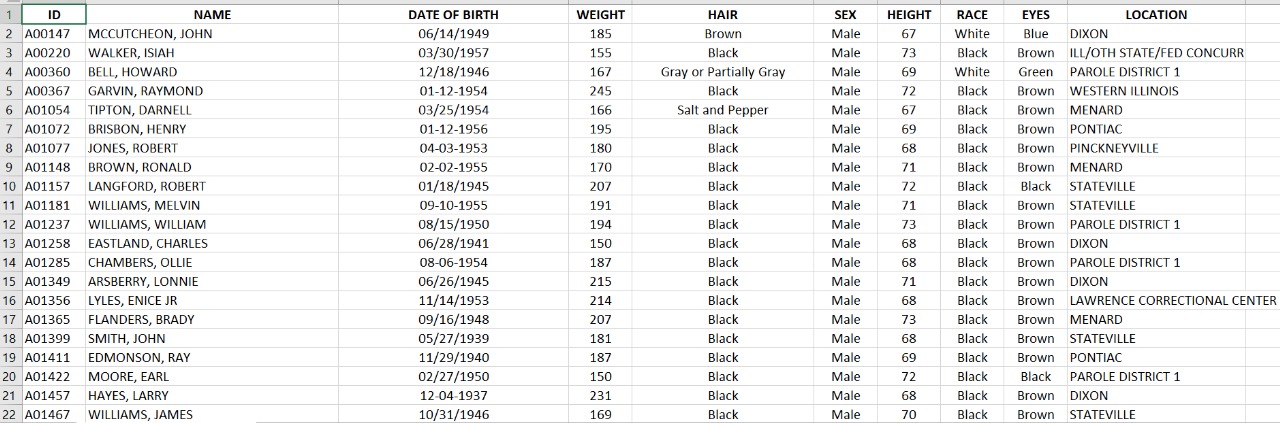
Alto converted this dataset into a multi fundamental approach, multi-perspectives dataset in which the sophisticated, narrow aspects of both textual, visual and contextual data intermingled. Its design also embraces how versatile the system is in order for it to work effectively in asserted and possibly unanticipated scenarios, the system is a useful tool for a dependable and diversified criminal identification system.

Figure 4.2 Representation of Data

**4.2 Experimental Results**

The VAE is used to learn the mapping of facial images into a lower dimensional space and reconstruct images from it. The training process is evaluated using two loss functions: reconstruction loss which preferably calculated with Mean Squared Error and KL divergence loss to balance the regulator of the latent space. We assume that the total loss and therefore the reconstruction capability as well as the quality of the feature encoding in more than 20 epochs decreases for the VAE. For additional improvement in performance, intermediate checkpoints are stored and the learning rate is controlled with a scheduler. These results show the VAE ability to generalize over different facial features, which are necessary to the downstream tasks.

The gotten features such as race, eye size and nose width are quantized and normalized so that a similarity calculation can be done on the user-defined entities. The system calculates weighted Euclidean distances regarding features that are defined by the user and entries of the dataset. Matches are ranked based on similarity scores, with higher scores indicating a closer match. Experimental results highlight the importance of feature weighting since it directly affects the quality of matches. For instance, giving more weight to critical features such as race or eye shape improves accuracy by accentuating those traits that are visually significant.

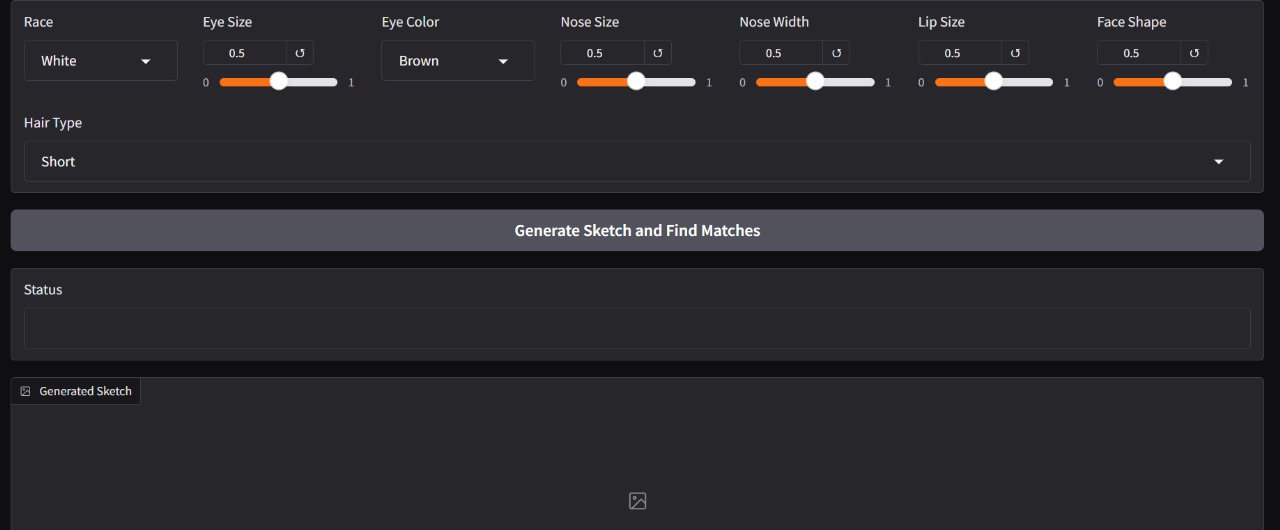


Figure 4.3 User Interface of the Sketch Generation System

A sketch is generated from a particular placement of features in the embedding space of VAE decoder. In assessing these sketches, they are analyzed with regard to user expectations about the level of realism and accuracy. The latent feature vector is passed into the VAE’s decoder and produces an image which has the attributes specified. A simple analysis where the success is determined by comparison of the sketch with the input parameters by eye. The sketches generated demonstrate an ability to map high-dimensional latent vectors and exhibit

the models’s ability to generate visually appropriate output.

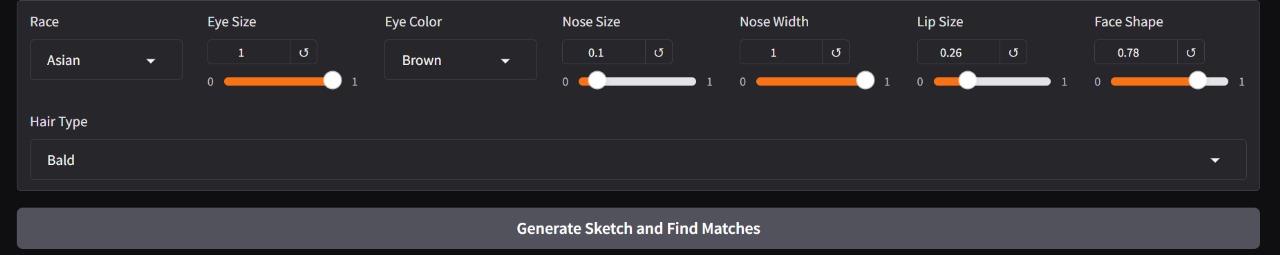


Figure 4.4 Representation of the Slider

The final outcomes are shown in the form of "n” comparable firms, with match similarity and other pertinent data, such as name and position, shown. Users also generate retrieved images that enable them to visually confirm matches. Thus, the effectiveness of the matching process can be put under the accuracy term and define it as the correspondence of the obtained matches to the input features. User feedback then peels off the coarse edges and refines the system by means of tweaking weights or thresholds.

The global results prove that the system is capable of drawing reasonable sketches and finding exact matches depending on the characteristics of the input data. The fragmentation loss gathered to train the VAE confirms the quality of the sketches generated, whereas the measures to match confirm that the faces visually similar to the probe are distinctive in the database.

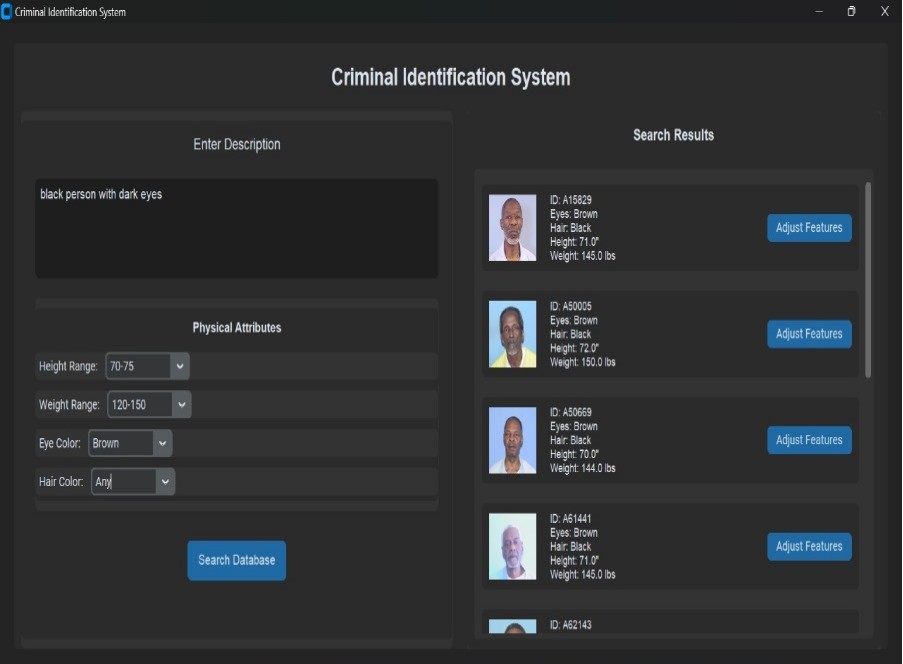


Figure 4.5 Results of given inputs

**4.3 Significance of the Proposed Method**

The proposed system goes towards a unique combination of current deep learning methodologies with features approach for the synthesis of face sketching and matching. This approach harnesses the capabilities of VAEs that learn an embedding of high dimensional face space while offering a feature space for comparison. This is a blend methodology that hopes to fulfill some of the crucial wants of applications that need generative and analytical abilities like facial reconstruction, forensic, and identification verifications.

**Advantages**

1. Feature-Driven Versatility:It means that due to the reliance of the current system on the feature-based input, there is almost incalculable flexibility. The users themselves can set certain parameters, for instance, hair and face type or colors of the eyes, while choosing different combinations when drawing sketches in order to meet certain requirements. Due to that, the model can be tailored and utilized almost to various user requirements, unlike other strictly imaginal models.

2. Hybridization of Generative and Analytical Models:The proposed system has been designed to have a twofold functionality of synthesizing realistic sketches of faces and or matching faces that are closest to each other by hybridizing the generative model of a variational autoencoder with facial feature similarity search. This makes it, in all ways, an all-in-one tool for doing creative and analytical work.

3. Sketch Generation of Excellent Quality:The proposed utilization of VAEs, therefore, means that produced sketches will be pretty realistic and detailed. It also means that while the VAE learns high levels of abstraction of the data, it is able to capture subtle [s]hifting of distances between facial features, which allows for both realism and accuracy in the sketches made by the system according to the user-specified attributes.

4. Precision in Face Matching:Due to the use of a weighted Euclidean similarity measure, alignments of most similar facial landmarks are made with high accuracy. This way the system learns priority given to specific attributes, such as race, nose size, or lip size, and then achieves precise similarity scores with some variations that make the system strong and reliable.

5. Scalability for Large Datasets:The model was made, with a view of accommodating large volumes of data into the system and fast comparison of features. This makes it possible to employ for various practical purposes associated with thousands of face records like, identity databases or criminal records.

6. Error Tolerance and Robustness:Individual features of the system are considered, and based on the architecture, the feature of the system that takes into account error tolerance is critical. The absence of input data of certain or uncertain kinds is solved in the system through normalized features and default mappings. This helps to provide further evaluation even in case of admitted missing data or erroneous inputs from users.

7. User-Friendly Interface:The Gradio integration to the system, therefore, comes in as a means to present the results in an easy-to-understand, fully visually interactive way – virtually bringing the system to eyes of user who is not necessarily a technical person at all. Such guidance as face attributes sliders and dropdown lists will make the process of defining inputs very fluent in order to communicate with the model.

8. Potential for Forensic and Law Enforcement Applications: The sketch-generation capability according to the describing inputs is very useful in forensics especially when reconstructing faces from actual descriptions given by witnesses. Compared to face-matching capability, the system has promoted the identification of culprits or even the tracing of lost individuals.

There is literature that offers solutions of the challenging problems of mapping structured data and visual representations and this is what this approach solves. How it can generate, compare, and match facial sketches at the request of a user in feature space is very promising for as diverse industries like police, cinema, and many others. Furthermore, the modularity of the envisioned system guaran teeths that otherel generative models, larger datasets and much more sophisticated feature mappings can be introduced to expand the basic framework, with mesurable effects in numerous fields.

\

## CHAPTER 5

**CONCLUSION AND FUTURE ENHANCEMENTS**

**5.1 Conclusion**

In conclusion, that we presented the face-sketch generation and matching system as one of the major steps forward in the convergence of generative modeling and feature-based analysis. It has a dual-functionality, based on Variational Autoencoders, for generating photo-realistic sketches, and similarity of features for matching. It addresses some of fundamental issues in these domains: reconstructions and identification of facial features as accurate as possible in forensics, law enforcement, and creative businesses. The feature-driven approach ensures an excellent balance between customization and precision: It is equally applicable practically in different issues arising practically.

Among the strengths of carrying out this project is that feature inputs are structured in a way that can easily be translated into visuals. This capability enables the user to input special characteristics like the size of the eyes, the shape of the nose or the race and to obtain correct facial sketches which indicate that the system can be easily adjusted toward real-world requirements. Weighted similarity metrics also hence enhance certainty that facial matching can be versatile and accurate and facilitate comparisons of faint differences on big data sets. This generative and analytical combination of capabilities designed for the system makes it possible to find faces in diverse scenarios, including criminal pursuits and searching for a missing person.

The above concept of developing an automated system is made user friendly through incorporation of the graphical interface from Gradio. The introduced interactive platform allows the fact that an extensive amount of training is not required; the system is usable by most of the stakeholders. It can be applied infinitely, due to both an obvious functionality and flexibility from small police departments urgently requiring face matching up to illustrators and designers creating individual characters. The system is equally well suited for noisy or imperfect inputs to guarantee good performance even in situations where the information is limited or noisy-an all too familiar story for forensic applications.

Last but not the least, the presented system is scalable and can be made vastly more complex in the future. Such design allows the general generative models to be updated, extends to a broader input set, and improve the feature extraction methods. All of these would further increase the flexibility of the system and extend its applications for use in the virtual reality, gaming and biosecurity. Most importantly it is completely compatible with the existing workflow and database and this means it will not have a drastic big disruption effect on industries for it to be adopted.

From the social angle, this project carries lots of potentials. For practical applications of the technology to enhance effectiveness and reliability of criminal identification for practical solutions it provides the possibility of reconstructing faces from such accounts. It democratises the faces sketch technology by reducing reliance on those conventional, hire sketch artists that mainly work for well-endowed institutions.

In other words, face-sketch generation and matching remain a fitting example of how the generative AI can be connected with the feature analysis. Not only it will be able to address current problems related to facial picturing and identification, but it will also form the basis for future advancements of HCI and AI. The system is prepared to bring the change in numerous fields since it has the flexibility to cover various concerns and further enhances its functions.

**5.2 Future Enhancements**

While the proposed face-sketch generation and matching very effective in terms of robustness, there still exists significant potential for increase in the complexity of the system so as to make it more flexible for use and faster in terms of performance. For instance, one of the major prospects of development to be implemented in the future is the incorporation of the contact with precise generative models, including GANs and diffusion models that might enhance the degree of reality and the detail of thesketches drawn by the program. It may also enhance the capacity of such models in recreating basic forms of facial structure common in complex datasets, and also make the system more responsive to its users in terms of the facility’s illustrations.

The second potential augmentation is large and complex data components. The system can be introduced to global applications through inserting Big Data processing pipeline and convex solution of cloud storage. This would be very useful if the, criminal databases that a law enforcement agency needs to process are of big size. Optimizing the model for better performance to stream on such datasets would ensure that the interaction with the user was tightly controlled both for response and space specifically during high usages.

It can be further expanded by incorporating other technologies such as the three dimensional modeling of faces and place and virtual reality play to areas of play like gaming, flick making and even envelope immersion. These will allow the system to draw more than simple 2D layouts, but a full and complete finished 3D model thus creating another dimension for usage and creation.

**APPENDICES**

In this chapter we can see the sample code for the Intelligent Police Sketch System, one can gain a better understanding of how the task is implemented in a more practical manner. The modules are implemented using Python programming language

import os

import numpy as np

import pandas as pd

import torch

import torch.nn as nn

import torch.nn.functional as F

from torch.utils.data import Dataset, DataLoader

from torchvision import transforms

from PIL import Image

import gradio as gr

import warnings

warnings.filterwarnings('ignore')

# Constants

DEVICE = torch.device('cuda' if torch.cuda.is\_available() else 'cpu')

IMAGE\_SIZE = 128

MAX\_IMAGES = 100

BATCH\_SIZE = 16

LEARNING\_RATE = 0.001

NUM\_EPOCHS = 20

LATENT\_DIM = 128

# Paths

DATA\_PATH = r"C:\Users\HP\OneDrive\Desktop\Major Project\Copy of person\_edited(1).xlsx"

IMAGE\_DIR = r"C:\Users\HP\OneDrive\Desktop\Major Project\front"

VAE\_WEIGHTS\_PATH = 'vae\_weights.pth'

class FaceDataset(Dataset):

def \_init\_(self, image\_dir, transform=None):

self.image\_dir = image\_dir

self.transform = transform

self.image\_files = [f for f in os.listdir(image\_dir) if f.endswith(('.jpg', '.jpeg', '.png'))][:MAX\_IMAGES]

print(f"Loading {len(self.image\_files)} images")

def \_len\_(self):

return len(self.image\_files)

def \_getitem\_(self, idx):

try:

img\_name = self.image\_files[idx]

img\_path = os.path.join(self.image\_dir, img\_name)

if not os.path.exists(img\_path):

print(f"Image not found: {img\_path}")

return None # Return None for missing images

image = Image.open(img\_path).convert('RGB')

if self.transform:

image = self.transform(image)

return image

except Exception as e:

print(f"Error loading image {img\_name}: {e}")

return None # Return None for errors

class VAE(nn.Module):

def \_init\_(self):

super(VAE, self).\_init\_()

self.latent\_dim = LATENT\_DIM

self.hidden\_dims = [32, 64, 128, 256]

# Encoder

modules = []

in\_channels = 3

for h\_dim in self.hidden\_dims:

modules.append(

nn.Sequential(

nn.Conv2d(in\_channels, h\_dim, kernel\_size=3, stride=2, padding=1),

nn.BatchNorm2d(h\_dim),

nn.LeakyReLU(),

nn.Dropout(0.1)

)

)

in\_channels = h\_dim

self.encoder = nn.Sequential(\*modules)

self.flatten\_size = self.hidden\_dims[-1] \* (IMAGE\_SIZE // 16) \* (IMAGE\_SIZE // 16)

# Latent space

self.fc\_mu = nn.Linear(self.flatten\_size, self.latent\_dim)

self.fc\_var = nn.Linear(self.flatten\_size, self.latent\_dim)

# Decoder

self.decoder\_input = nn.Linear(self.latent\_dim, self.flatten\_size)

modules = []

hidden\_dims = self.hidden\_dims[::-1]

for i in range(len(hidden\_dims) - 1):

modules.append(

nn.Sequential(

nn.ConvTranspose2d(hidden\_dims[i], hidden\_dims[i + 1],

kernel\_size=3, stride=2, padding=1, output\_padding=1),

nn.BatchNorm2d(hidden\_dims[i + 1]),

nn.LeakyReLU(),

nn.Dropout(0.1)

)

)

# Final layer

modules.append(

nn.Sequential(

nn.ConvTranspose2d(hidden\_dims[-1], 3, kernel\_size=3, stride=2, padding=1, output\_padding=1),

nn.Tanh()

)

)

self.decoder = nn.Sequential(\*modules)

def encode(self, x):

x = self.encoder(x)

x = torch.flatten(x, start\_dim=1)

mu = self.fc\_mu(x)

log\_var = self.fc\_var(x)

return mu, log\_var

def decode(self, z):

x = self.decoder\_input(z)

x = x.view(-1, self.hidden\_dims[-1], IMAGE\_SIZE // 16, IMAGE\_SIZE // 16)

x = self.decoder(x)

return x

def reparameterize(self, mu, log\_var):

std = torch.exp(0.5 \* log\_var)

eps = torch.randn\_like(std)

return mu + eps \* std

def forward(self, x):

mu, log\_var = self.encode(x)

z = self.reparameterize(mu, log\_var)

return self.decode(z), mu, log\_var

class FaceMatchingSystem:

def \_init\_(self):

# Initialize VAE

self.vae = VAE().to(DEVICE)

if os.path.exists(VAE\_WEIGHTS\_PATH):

try:

self.vae.load\_state\_dict(torch.load(VAE\_WEIGHTS\_PATH, map\_location=DEVICE))

except Exception as e:

print(f"Error loading VAE weights: {e}")

if os.path.exists(VAE\_WEIGHTS\_PATH):

os.remove(VAE\_WEIGHTS\_PATH)

self.vae.eval()

# Load dataset

self.dataset = pd.read\_excel(DATA\_PATH)

self.dataset = self.dataset.head(MAX\_IMAGES)

# Image preprocessing

self.transform = transforms.Compose([

transforms.Resize((IMAGE\_SIZE, IMAGE\_SIZE)),

transforms.ToTensor(),

])

def map\_features(self, person):

"""Map string features to numerical values, excluding age."""

race\_mapping = {'White': 0, 'Black': 1, 'Asian': 2, 'Hispanic': 3}

eye\_color\_mapping = {'Brown': 0, 'Blue': 1, 'Green': 2, 'Black': 3}

try:

# Map features and handle potential errors

race\_val = race\_mapping.get(person['race'], 0) / 3 # Normalize race

eye\_size\_val = float(person['eye\_size'])

eye\_color\_val = eye\_color\_mapping.get(person['eye\_color'], 0) / 3 # Normalize eye color

nose\_size\_val = float(person['nose\_size'])

nose\_width\_val = float(person['nose\_width'])

lip\_size\_val = float(person['lip\_size'])

face\_shape\_val = float(person['face\_shape'])

hair\_type\_val = float(person['hair\_type']) / 3 # Normalize hair type

return np.array([

race\_val,

eye\_size\_val,

eye\_color\_val,

nose\_size\_val,

nose\_width\_val,

lip\_size\_val,

face\_shape\_val,

hair\_type\_val

])

except Exception as e:

print(f"Error processing features for person {person.get('id', 'unknown')}: {e}")

return np.zeros(8) # Return a zero array if there's an error

def find\_matches(self, features, top\_n=5):

"""Find matches based on feature similarity, excluding age."""

try:

matches = []

# Convert feature values to numeric, excluding age

query\_features = np.array([

features['race\_val'],

features['eye\_size'],

features['eye\_color\_val'],

features['nose\_size'],

**REFERENCES**

[1]Zirui An, Jingbo Yu, Runtao Liu, Chuang Wang, Qian Yu, “SketchInverter: Multi-Class Sketch-Based Image Generation via GAN Inversion”, IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, pp.4308-4318, 2023, doi:10.1109/WACV56688.2023.00430

[2]Feiwei Qin, Shi Qiu, Shuming Gao, Jing Bai, “3D CAD model retrieval based on sketch and unsupervised variational autoencoder” Advanced Engineering Informatics, China, pp. 101427 , 2023, doi: 10.1016/j.aei.2021.101427

[3]Troy Luhman, Eric Luhman, “High Fidelity Image Synthesis With Deep VAEs In Latent Space”, published at arXiv:2303.13714 ,2023 doi: 10.48550/arXiv.2303.13714

[4]Y. Liu, H. Zhao, X. Wu, “cGAN-based Sketched Image Art Generator Using Deep Learning” IEEE Transactions on Computational Imaging, 2023, doi: 10.1109/ACCESS.10170433.

[5]Stanislav Polyakov , Alexander Kryukov, Andrey Demichev, Julia Dubenskaya , Elizaveta Gres , Anna Vlaskina , “ Using conditional variational autoencoders to generate images from atmospheric Cherenkov telescopes”, pushlished at arXiv:2211.12553 ,2023, doi: 10.48550/arXiv.2211.12553

[6]Seho Bae, Nizam Ud Din, Hyunkyu Park, Juneho Yi, “Face Photo-Sketch Recognition Using Bidirectional Collaborative Synthesis Network” 16th International Conference on Ubiquitous Information Management and Communication (IMCOM) ,  Seoul, Korea, Republic of ,2022, pp. 1-8 ,doi: 10.1109/IMCOM53663.2022.9721719

[7]Chira, D., Haralampiev, I., Winther, O., Dittadi, A., Liévin, V. (2023), “Image Super-Resolution with Deep Variational Autoencoders.” In: Karlinsky, L., Michaeli, T., Nishino, K. (eds) Computer Vision – ECCV 2022 Workshops. ECCV 2022. Lecture Notes in Computer Science, vol 13802. Springer, Cham, doi:10.1007/978-3-031-25063-7\_24

[8]Jonathan Ho, Chitwan Saharia, William Chan, David J. Fleet, Mohammad Norouzi, Tim Salimans, “Cascaded Diffusion Models for High Fidelity Image Generation”, The Journal of Machine Learning Research, Volume 23, Issue 1,2022, doi:10.5555/3586589.3586636

[9]Zong-Yue Deng, Hsin-Han Chiang, Li-Wei Kang, Hsiao-Chi Li, “A lightweight deep learning model for real-time face recognition” , published at sensors ,2020,doi: 10.3390/s20216114

[10]Mohamed Amine Hmani, Dijana Petrovska-Delacrétaz, Bernadette Dorizzi, “Locality preserving binary face representations using auto-encoders” Publication: IET Biometrics Volume 11, Issue 5,pp 445-458, 2022, doi: 10.1049/bme2.12096

[11]B. Bae, J. Chong, and M. Cho, “DigiFace-1M: 1 Million Digital Face Images for Face Recognition” Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2023, doi: 10.1109/WACV56688.2023.00129.

[12]Akingbesote, Damilola ; Zhan, Ying ; Maskeliūnas, Rytis ; Damaševičius, Robertas, “A novel deep facenet framework for real-time face detection based on deep learning model” published at Algorithms.. Basel : MDPI. 2023, vol. 16, iss. 6, art. no. 292, p. 1-24, 2023, doi: 10.3390/a16060292

[13]Pawan Kumar, Nihal Manzoor, Chhavi Dhiman, “Masked-face recognition using deep metric learning and FaceMaskNet-21”, 8th International Conference on Signal Processing and Integrated Networks

(SPIN), Noida, India, pp. 569 – 573, 2021, doi: 10.1109/SPIN52536.2021.9566002 .

[14]B. Sharma, S. Kaur, "Enhancing Suspect Identification: Automated Composite Sketch Generation and Recognition" Applied Intelligence, vol. 51, no. 10, pp. 7112–7123, 2021, doi: 10.1007/s10489-021-03150-3.

[15]A. Kumar, V. K. Jain, "Face Sketch Recognition" Lecture Notes in Networks and Systems, Springer, 2023, doi: 10.1007/978-981-97-6036-7\_9.

[16]S. P. Singh, R. C. Gupta, "Performance enhancement of generative adversarial network for photograph–sketch identification" International Journal for Research Trends and Innovation, vol. 8, no. 1, pp. 25-30, 2024, doi: 10.52792/IJRTI2401045.