A

Major Project Report on

AI-Driven Forensic Face Sketch Construction and Recognition

submitted in partial fulfillment of the requirements for
the award of the degree of

Bachelor of Engineering

in

Artificial Intelligence and Data Science

By

Pasunuri Kathyayini (1601-21-771-088)

Cheera Rohan (1601-21-771-099)

Gillala Vaishak Reddy (1601-21-771-102)

Under the esteemed guidance of

Ms. Talla Sai Sree

Assistant Professor

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY

HYDERABAD – 500075

MAY 2025

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY

HYDERABAD – 500075

INSTITUTE VISION

"To be the center of excellence in technical education and research".

INSTITUTE MISSION

"To address the emerging needs through quality technical education and advanced research".

DEPARTMENT VISION

"To be a globally recognized center of excellence in the field of Artificial Intelligence and Data Science that produces innovative pioneers and research experts capable of addressing complex real-world challenges and contributing to the socio-economic development of the nation."

DEPARTMENT MISSION

1. To provide cutting-edge education in the field of Artificial Intelligence and Data Science that is rooted in ethical and moral values.

2. To establish strong partnerships with industries and research organizations in the field of Artificial Intelligence and Data Science, and to excel in the emerging areas of research by creating innovative solutions.

3. To cultivate a strong sense of social responsibility among students, fostering their inclination to utilize their knowledge and skills for the betterment of society.

4. To motivate and mentor students to become trailblazers in Artificial Intelligence and Data Science, and develop an entrepreneurial mindset that nurtures innovation and creativity.

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY

HYDERABAD – 500075

DECLARATION CERTIFICATE

We hereby declare that the project titled **AI-Driven Forensic Face Sketch Construction and Recognition** submitted by us to the department of **Artificial Intelligence and Data Science**, **CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY, HYDERABAD** in partial fulfillment of the requirements for the award of **Bachelor of Engineering** is a bona-fide record of the work carried out by us under the supervision of **Ms. Talla Sai Sree** . We further declare that the work reported in this project, has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma of this institute or of any other institute or University.

Project Associates

Pasunuri Kathyayini (1601-21-771-088)

Cheera Rohan (1601-21-771-099)

Gillala Vaishak Reddy (1601-21-771-102)

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY

HYDERABAD – 500075

BONAFIDE CERTIFICATE

This is to certify that the project titled **AI-Driven Forensic Face Sketch Construction and Recognition** is a bonafide record of the work done by

Pasunuri Kathyayini (1601-21-771-088)

Cheera Rohan (1601-21-771-099)

Gillala Vaishak Reddy (1601-21-771-102)

in partial fulfillment of the requirements for the award of the degree of **Bachelor of Engineering** in **Artificial Intelligence and Data Science** to the **CHAITANYA BHARATHI INSTITUTE OF TECHNOLOGY, HYDERABAD** carried out under my guidance and supervision during the year 2024-25. The results presented in this project report have not been submitted to any other university or Institute for the award of any degree.

Ms. Talla Sai Sree — Guide

Prof. K. Radhika — Head of the Department

Submitted for Final Year Major-Project viva-voce examination held on ______________

Examiner-1

Examiner-2

ABSTRACT

Facial sketching and identification are valuable resources in contemporary criminal investigations and are critical components of suspect identification when photographic evidence is not available. Conventional forensic sketching depends upon the artistic skill and interpretation of the artist to render eyewitness descriptions into facial composites. Although this technique has proven to be beneficial, it is subjective, time-consuming, and significantly relies upon the experience of the artist and the memory of the witness. Consequently, the outlines tend to be inaccurate and inconsistent, thereby diminishing their usefulness in high-stakes investigations. While recent progress in automated face recognition with Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs) has been encouraging, it remains plagued by important limitations. CNN-based approaches tend to miss detailed, identity-related features, resulting in low feature correlation and weak recognition performance. GANs, in contrast, improve the realism of images but tend to blur or modify important identity features, which can critically undermine the system's reliability and trustworthiness.

This project introduces a new AI-based forensic face sketching and identification framework to address the limitations of both traditional and current automated methods. In the initial phase, investigators can generate high-fidelity composite sketches by choosing from a library of modular facial features—eyes, eyebrows, noses, lips, hairstyles, and facial hair. Every feature is selected to build the overall identity match in a positive way. For guaranteeing proportionality and realism, an AI-driven recommendation system comes to the aid by providing compatible feature combinations, thus reducing human bias and inconsistency.

After the sketch is finished, the system renders it as a photorealistic digital image through a sketch-to-photo conversion module, maintaining significant identity features while improving visual quality. The image is then input into a recognition engine that compares it to a criminal database, providing the top five most similar faces and a breakdown of individual feature similarity scores. By combining modular facial synthesis, AI suggestions, photorealistic conversion, and smart matching, this system minimizes manual labor considerably, improves precision, and streamlines the suspect identification process. Overall, it offers law enforcement agencies a robust, scalable, and data-driven method of solving crimes more efficiently and effectively.

Keywords: Forensic sketch, Face Recognition, Deep Learning, Convolution Neural Network, Generative Adversarial Network, Criminal investigations

ACKNOWLEDGEMENTS

We would like to express our deepest gratitude to the following people for guiding us through this course and without whom this project and the results achieved from it would not have reached completion.

Ms. Talla Sai Sree, Assistant Professor, Department of Artificial Intelligence and Data Science, for helping us and guiding us in the course of this project. Without her guidance, we would not have been able to successfully complete this project. Her patience and genial attitude is and always will be a source of inspiration to us.

Prof. K. Radhika, the Head of the Department, Department of Artificial Intelligence and Data Science, for allowing us to avail the facilities at the department.

We show gratitude to our honorable Principal **Prof. C.V. Narsimhulu** for providing all facilities and support.

We are also thankful to the faculty and staff members of the Department of Artificial Intelligence and Data Science, our parents and our friends for their constant support and help.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 Introduction to Forensic Face Sketch Construction and Recognition Systems

Criminal investigations usually rely on different types of evidence to determine suspects, but in the absence of photographic or video evidence, forensic face sketching is an essential tool. The capability to reconstruct a suspect's features from eyewitness accounts is essential in crime solving. Police forces have long used talented forensic artists to try to convey verbal data in visual forms with sketches, yet the process is highly subjective. It relies upon interpretations by the artist and also on the recollection of the witness, and in the midst of the turmoil, this fails. Forensic skills, as well as personal biases, can also contribute to impacting accuracy and leading to inaccuracies that hamper investigations. Furthermore, manual sketches consume a lot of time, preventing the progression of the case and minimizing the chances of rapid identification of suspects. Forensic identification is also a task of equal magnitude since freehand sketches are far from actual photographs.

Conventional identification techniques extract essential facial attributes and try to match them against criminal files, but the lack of color, texture, and elaborate shading makes direct comparison difficult. Most conventional image-processing techniques are edge detection and structural similarity, which ignore the artistic simplification of drawings. Such a disconnection between sketches and real images typically results in a high false positive or negative rate, rendering traditional recognition ineffective.

Advances in technology in AI and machine learning are revolutionizing forensic sketching and recognition, addressing most of the flaws. AI-enabled tools can produce and refine sketches with greater accuracy and speed, such that critical facial features closely match real identities. These technologies narrow the gap between photographs and sketches in databases, increasing the rates of recognition and making it possible for more accurate identification of suspects. AI eliminates dependence on the quality of the artist, mini mizes human error, and streamlines the whole process of investigation by automating and optimizing the process of sketching. Other than accuracy, AI-driven forensic technology offers enormous value in terms of speed, scalability, andusability. Investigations are under pressure where every second is precious. AI systems can generate drawings in minutes, greatly reducing times of investigation and increasing the likelihood of making arrests within a timely fashion. Small police forces with no access to forensic artists can apply AI for standard, high-quality suspect identification on a regular basis. Above all, these improvements enhance justice and public safety through reduced wrongful arrests, strengthening testimonies of witnesses, and assuring law enforcers are able to function with more certainty and accuracy when solving crimes.

1.2 Problem Definition

Correct suspect identification is a pillar of contemporary criminal investigations. In many real-life cases, particularly when there is no photographic or surveillance evidence, the main source of visual information regarding a suspect is from eyewitness descriptions. Historically, these verbal descriptions are converted into visual sketches by forensic artists—a process that, although useful, is riddled with several key limitations. The process of sketching is very subjective and inconsistent, greatly depending on the eyewitness's capacity to remember under traumatic or stressful conditions. Even if memory is good, the sketch produced relies on the skill and interpretation of the artist, which may not be similar from one artist to another. Therefore, most sketches do not identify distinctive facial characteristics of suspects with accuracy, and identification rates become low. Furthermore, the manual process is time-consuming, making it ill-suited for time-sensitive investigations or large-scale operations.

Beyond sketch construction, the second major challenge lies in recognizing and matching these sketches with existing digital databases of criminal records. Traditional face recognition systems are designed for high-quality photographic images and often perform poorly when dealing with hand-drawn or stylized sketches. The intrinsic modality gap—the disparity in appearance between drawings and actual images—hinders these systems from deriving meaningful and comparable features. Therefore, most current systems do not yield accurate or dependable matches, reducing their applicability to active investigations.

Combined, these limitations establish a critical gap in the criminal justice process:

1. There is no dependable method to create realistic, identity-maintaining sketches from eyewitness feedback in a uniform and scalable manner.
2. There is no robust and effective means of matching sketches with real-world

images with good accuracy, particularly for varied or unstructured sets of data.

This two-fold difficulty severely prolongs investigations, diminishes the chances of making prompt suspect identification, and adds excessive burdens to witnesses as well as forensic experts.

To this end, the project aims to solve these problem by introducing an AI-Driven Forensic Face Sketch Construction and Recognition system that combines smart composite sketch creation with deep learning-based recognition. The objective is to design a streamlined, objective, and efficient process that reduces human bias, maintains essential identity features, and facilitates correct suspect recognition from sketches. By closing the gap between human description and digital identification, the system equips law enforcement agencies with a more data-driven, scalable, and reliable investigative asset.

1.3 Objectives and Outcomes

The main aim of this project is to automate and rationalize the conventional forensic sketching and suspect identification procedure through artificial intelligence. By combining intelligent composite sketch design with deep learning-based face recognition strategies, the system seeks to remove human bias, enhance efficiency, and enhance accuracy in suspect identification. The following aims were set to spearhead the creation of this solution:

- To create an interactive system for constructing composite forensic face sketches on the basis of individual facial elements like eyes, nose, lips, eyebrows, and hair to enable investigators to construct a clear and organized facial profile.
- To introduce an AI-based feature recommendation mechanism that offers compatible facial elements while constructing the sketch to make the proportions more accurate and lower inconsistencies.
- To take composite sketches to digitally enhanced, colored versions that maintain identity-specific features along with enhanced visual clarity and recognition, thus optimizing them for easier comparison.
- To engineer and train a facial recognition system using deep learning that will compare the digitally enhanced sketches favourably with the database of active criminal records.

- To get back the top five most visually and structurally comparable matches from the database, along with percentage-wise analyses of similarity in individual facial features.

- To have a scalable, efficient, and user-friendly workflow minimizing manual effort and decreasing the turnaround time for suspect identification.

- To assist law enforcement agencies with a data-assisted and dependable tool for investigation, especially in cases without photographic evidence.

1.4 Motivation

The impetus for this project comes from the urgent necessity to update one of the most fundamental—but too often neglected—resources in criminal investigation: identification of suspects via facial sketches. As forensic science has progressed exponentially in areas such as DNA and electronic surveillance, the method of sketch-based identification has remained stuck in a bygone era for decades, continuing to rely on artistic license and witness memory. This gap between technological progress and practical implementation was a primary stimulus for starting this project.

In our research and exploration, we noticed a critical gap—not just in accuracy, but in equity and accessibility. Smaller law enforcement agencies or those in resource-constrained settings may not have access to trained forensic artists at all times, leaving them at a disadvantage during crucial early hours of an investigation. Meanwhile, reliance on one or two interpretations of memory under duress can skew outcomes and unintentionally influence case direction. Our goal was not merely to digitalize sketching, but to democratize it—making sure that any investigative team, wherever they were located and whatever resources they had, could create a believable and coherent suspect profile.

We were also motivated by a larger question: what if the sketch isn't merely a lead, but a match? The notion that a well-crafted, AI-enhanced sketch might not only indicate a suspect but assist in verifying their identity from a database in real time, gave the project a significant investigative advantage. Beyond convenience, this is a transformation of intent—from suggestion to identification. In addition, while we reflected on actual tragedies brought about by wrongful arrests or slow identifications, we understood that a more standardized and data-driven system would be able to restore people's trust in criminal investigations. In addition, as we thought about actual tragedies that resulted from mistaken arrests or late identifications, we knew that a more data-driven, standardized, and transparent system is required. Technology, used ethically and responsibly, is able to decrease the margin for human error without eliminating human supervision.

This initiative was conceived in that crossroads—of empathy and engineering. Combining AI with forensic necessity, our drive is to develop a system that honors both the time sensitivity and dignity of justice. It's not just about crafting a solution, but making sure it gets investigators closer to the truth—faster, more fairly, and with more assurance.

1.5 Thesis Organization

The thesis has been structured into various key parts to clearly lay out the research on construction and recognition of forensic face sketches. The Introduction serves to overview the research, which is started off by introducing the topic, the problem definition, the objectives and anticipated results of the research, as well as the reason why the study has been conducted. The section wraps up with the preview of how the thesis has been structured. The Literature Review provides an in-depth review of five influential research papers in the area, discussing the different approaches and technologies employed in forensic face sketch recognition. This section also provides a summary table for convenient comparison and determines the research gaps that currently exist, which are addressed in the present work. In the Proposed System section, the system methodology is presented, including the architecture of the system and a thorough explanation of how the system would be implemented. The results and discussions are also included here, showing the success of the proposed method. Finally, the Conclusion and Future Scope section provides a brief overview of the research findings' main conclusions, presenting a final evaluation of the contributions of the proposed system to the field, and specifying the possible avenues for future work and enhancements in forensic face sketch recognition.

1.6 Existing System

The forensic process of face identification has traditionally had its basis in manual and semi-manual methodologies. Although they provided the bedrock for tracking suspects from eyewitness descriptions, current criminal investigations based on time-sensitive, accurate, and consistent parameters expose serious shortfalls.

1. **Forensic Sketch Construction**

   Traditionally, the process of creating a face sketch of a suspect starts with an interview between a forensic artist who has been trained and an eyewitness.

The artist will listen as the witness details the suspect's features—eye shape, nose size, jawline, or hairstyle—and then try to translate those verbal descriptions into a hand-drawn portrait. The success of this process relies significantly upon two human variables: the reliability of the eyewitness's memory and the artistic ability of the artist. Both are unfortunately prone to variation. Trauma, fear, and time can all affect memory in a detrimental way. In the same vein, even the best artist can inadvertently misread or amplify features.

To reduce the reliance on artistic skill, some semi-automated systems were introduced, which allowed investigators to build a face using predefined sets of eyes, noses, mouths, and other features from a database. Tools like these sped up the process and made it more accessible to non-artists. However, they often produced stiff, unrealistic composites. Witnesses had a hard time selecting perfect matches from a narrow library of components, and the resulting images often fell short in emotional expressiveness or natural appearance. Furthermore, those systems also did not have dynamic guidance or feature suggestions, which made it difficult to build consistent and balanced facial representations.

2. **Face Recognition from Sketches**

   Once a sketch is created, matching a suspect based on it is the second key step. In conventional processes, this was done by hand, comparing the sketch to mugshots or applying rudimentary software utilities that try to match the sketch against a database of recognizable individuals. The process is laborious and hardly accurate, especially if the sketch doesn't look much like photographs of real faces. Differences in style, shading, and absence of photographic elements make sketches increasingly dissimilar to real faces.

   Where some police databases incorporate sketch-based searching, these frequently use old-style visual matching technologies and only yield true matches when the sketch closely approximates actual appearance. With this, countless leads are overlooked, and investigation momentum is delayed. In situations involving high risks, this causes serious repercussions—failed identifications, wrongful arrestments, or even untracked suspects.

These challenges mirror the need for a modernized, intelligent system that enhances the whole sketching and identification process to become more accurate, quicker, and available—particularly in situations where there is no photo evidence. A system that assists investigators with drawing realistic, proportionate sketches and then identifying them correctly to known offenders can go a long way in facilitating efficient criminal investigations and delivering justice without delay or inaccuracies.

1.7 Challenges in Current Approaches

Although they have been in use for ages in forensic investigations, traditional systems of face sketching and identification have some fundamental problems that keep them from being more effective in actual criminal investigations. These problems not only undermine the accuracy of suspect identification but also more importantly delay the investigation process, usually at the expense of justice.

1. **Heavy Reliance on Eyewitness Memory and Artist Perception**

   One of the essential challenges is the subjective nature of sketching. Eyewitnesses, particularly under stress or after considerable time lapse, might provide fuzzy, imprecise, or contradictory descriptions. Artists have to decipher these descriptions and translate them into visual sketches, a procedure that is inherently subject to error, artistic prejudice, and non-standardization. The resulting sketch may be widely different depending upon how the artist perceives the same input, impacting its utility for identification.

2. **Lack of Realism and Proportionality in Constructed Sketches**

   Even with half-automatic composite sketch software, realism is still a big problem. Such systems provide predetermined sets of facial features that can be mismatched with the individual being remembered. The combinations might look unnatural or cartoon-like, without emotional nuance and facial balance. Without the ability to modify proportions or see slight variations, the sketches produced tend to fail to represent a recognizable face.

3. **Time-Consuming and Skill-Intensive Process**

   Conventional sketching is a labor-intensive process involving one-on-one meetings between the artist and the witness. This decelerates investigations, particularly where timely identification is important. Furthermore, dependence on trained forensic artists constrains scalability—such specialists might not be present in all law enforcement agencies, particularly in smaller or low-resource jurisdictions.

4. **Inability to Match Sketches to Photographs**

   After a sketch is finished, the next obstacle is recognition—comparing it to images in current criminal records. Most systems have difficulty comparing hand-drawn or computer-created sketches with actual photographs because they differ in texture, shading, and detail. This visual difference renders traditional facial recognition systems unsuitable, since they tend to be tuned to compare real-world images, not drawings.

5. **Low Accuracy and High False Positives**

   Due to the creative nature of sketches and inconsistency in witness recall, matching attempts frequently yield erroneous results or a complete lack of match. Systems can flag innocent individuals or overlook actual suspects, resulting in wasted investigative effort, wrongful suspicion, or worse—opportunities to prevent subsequent crimes being missed.

6. **Absence of Intelligent Assistance or Automation**

   Existing systems provide little or no feedback throughout the sketch construction process. There is no smart feedback to assist investigators in selecting compatible or realistic facial features, and no adaptive support to rectify disproportional components. This puts more faith in human judgment and can lead to inconsistent or impossible outputs.

These challenges emphasize the imperative of a more sophisticated, intuitive, and data-based process for forensic sketching and identification—one that reduces human bias, increases realism, speeds up matching, and facilitates sound suspect identification.

CHAPTER 2
LITERATURE SURVEY

2.1 Paper-1: Domain Alignment Embedding Network for Sketch Face Recognition

The paper[1] Domain Alignment Embedding Network for Sketch Face Recognition-introduces a new method called Domain Alignment Embedding Network (DAEN) for the task of Sketch Face Recognition (SFR)—a domain where forensic sketches need to be matched to real-life face photographs for identity verification, often in law enforcement and criminal investigations. The major challenges in this domain include (1) the modality gap between hand-drawn or software-generated sketches and high-quality digital photos, and (2) the limited availability of paired sketch-photo data, which makes deep learning-based training prone to overfitting. The authors propose a solution that combines deep metric learning with a domain alignment strategy inspired by few-shot learning.

DAEN employs a training episode strategy where a series of few-shot learning-like episodes are formed. In each episode, randomly selected classes form domain-aware query and support sets—composed of both sketches and photos. The model learns to embed both modalities into a shared embedding space using a novel Domain Alignment Embedding Loss, which has two components: the Sketch Domain Loss (LSDL) and the Photo Domain Loss (LPDL). These loss functions are computed using softmax functions over Euclidean distances between features of sketch-photo pairs, encouraging the network to align identities across domains while maintaining inter-class separability. The backbone used for feature extraction is ResNet-18, chosen for its balance of accuracy and computational efficiency.

To evaluate the effectiveness of DAEN, the authors conducted experiments on two benchmark datasets:

- UoM-SGFSv2, which contains 600 subjects, each with two sketches (EFIT-V generated and manually refined versions), forming two subsets—Set A and Set B.
- PRIP-VSGC, which includes 123 subjects with one photo and one sketch each, created using the Identi-Kit software by an artist.

The experiments follow standard protocols with multiple train/test splits. The DAEN model was compared against various state-of-the-art approaches, including traditional intra-modality methods like EP(+PCA), and inter-modality deep learning methods like SP-Net, Identity-aware CycleGAN, and DEEPS. On the UoM-SGFSv2 Set A, which is more challenging due to larger modality gaps, DAEN outperformed the next best method by nearly 37% at rank-1 accuracy. On Set B, it led by 22%. In the PRIP-VSGC dataset, which poses a small-sample challenge, DAEN surpassed other methods by around 10% at rank-10 accuracy.

In addition to performance, the method showed advantages in efficiency and scalability. Due to the lightweight ResNet-18 backbone, DAEN had fewer parameters and shorter inference time compared to heavier models using VGG-16. The paper concludes that DAEN not only bridges the sketch-photo modality gap more effectively but also enables reliable recognition with limited data—making it a robust tool for real-world forensic applications.

2.2 Paper-2: Feature Encoder Guided Generative Adversarial Network for Face Photo-Sketch Synthesis

Feature Encoder Guided Generative Adversarial Network for Face Photo-Sketch Synthesis[2] by Jieying Zheng, Wanru Song, Yahong Wu, Ran Xu, and Feng Liu suggests a better GAN-based methodology called EGGAN for the reciprocal synthesis of face sketches and photos. The need arises due to shortcomings of existing approaches that resulted in low image quality, texture loss, facial distortion, and inability to preserve identity-specific details such as beards or spectacles. In response to these difficulties, the authors propose a new feature encoder to steer generator training in a cycle-consistent GAN framework.

EGGAN is an extension of the CycleGAN model with two generators and two discriminators that allow both directions of transformation between the sketch and photo domains. A distinguishing aspect of EGGAN is the use of a shared feature encoder that discovers a common latent space representation of both domains. In contrast to classical models in which encoders explicitly control image synthesis, this encoder passively controls the training process by using two new loss functions: feature loss and feature consistency loss. These losses reduce the difference between real and synthesized images in the feature space while preserving detailed and identity-specific information. The overall loss function combines adversarial, cycle-consistency, identity, feature-based, and total variation losses to produce smooth, artifact-free images.

It is trained and evaluated on two datasets released to the public: CUFS (comprising CUHK, AR, and XM2VTS subsets) and more challenging CUFSF, containing variation in light conditions and face expressions. For training, resizing and normalization are done for the images, and Adam optimizer optimizes the model. EGGAN not only demonstrates better visual quality in sketch-to-photo and photo-to-sketch synthesis but also surpasses current methods like FCN, pix2pix, CycleGAN, DualGAN, and CSGAN in SSIM (Structural Similarity Index), LPIPS (Learned Perceptual Image Patch Similarity), and face recognition accuracy. Importantly, EGGAN has a perfect 100% face recognition rate on the CUFSF dataset, which indicates its real-world effectiveness for identity-sensitive tasks.

Along with enhanced synthesis performance, EGGAN preserves computational efficiency with only a slight rise in model complexity. The encoder adds merely around 2.2% to the overall parameter size, and training convergence occurs within 200 epochs. Ablation studies also confirm the significance of each component, especially the contribution of feature-guided losses to improving perceptual quality. In summary, EGGAN provides a strong, effective, and scalable solution for face photo-sketch synthesis and shows promise for applications in image-to-image translation tasks of style or modality conversion.

2.3 Paper-3: An Identity-Preserved Model for Face Sketch-Photo Synthesis

This work titled An Identity-Preserved Model for Face Sketch-Photo Synthesis [3] by Yunfan Lin, Shengping Ling, Kun Fu, and Peiyi Cheng proposes a deep learning-based model known as *Identity-Preserved Adversarial Model (IPAM)* for enhancing the face sketch-to-photo synthesis task. This task is vital in forensic applications where sketches of suspects must be compared with photos in available criminal databases. Although most current methods produce realistic images, they do not preserve identity-conveying facial features, which results in inferior recognition performance. The given IPAM solves this problem by integrating an adapted U-Net generator, dual discriminators, and a new identity-preserving constraint to preserve essential facial features during the photo-to-sketch conversion.
The IPAM model is constructed on top of a Generative Adversarial Network (GAN) architecture. The generator uses an extended U-Net structure with additional skip connections to strengthen the connection between the input sketch and the generated photo. Two discriminators are utilized: one for differentiating between sketch-photo pairs (cross-domain), and another for making sure the output is within the distribu-

tion of natural face photos. Also, an identity-preserving loss is proposed with cosine similarity, which makes the generated image preserve identity-related features by minimizing feature distance with the original sketch in a deep feature space with a pre-trained face recognition model.

For experimentation, the authors employ two benchmark datasets: the CUHK Face Sketch Database (CUFS) and the CUHK Face Sketch FERET Database (CUFSF). CUFS consists of controlled, artist-created sketches, whereas CUFSF exhibits more realistic challenges including variation in lighting and exaggeration. The model is trained on 88 and 100 image pairs of CUFS and CUFSF respectively, and tested on recognition measures like Rank-1 identification rate and verification accuracy at a false acceptance rate of 0.1%.

Results show that IPAM consistently achieves better performance over state-of-the-art approaches such as Pix2Pix, CycleGAN, PS2-MAN, and MDAL regarding image quality as well as recognition performance. Ablation studies also confirm the contribution of each module of IPAM—demonstrating that the dual-discriminator framework, identity constraint, and extended generator all together contribute to the realism and identity preservation of the fake images. The model thus turns out to be extremely efficient in closing the gap between face sketch inputs and face recognition systems in actual forensic usage.

2.4 Paper-4: Dual-View Normalization for Face Recognition

This paper[4] by Guo-Shiang Hsu and Chih-Hung Tang, introduces a new framework named *Dual-View Normalization (DVN)* for improving face recognition using pose normalization. Previous normalization methods commonly transform non-frontal faces to a single frontal pose for enhanced recognition accuracy. But this effort goes beyond that by producing not one but two normalized views—a frontal view and a yaw-45°side view—from any face pose. This dual-view strategy is motivated by forensic best practices that tend to use both frontal and side profiles for identification verification. The DVN model seeks to enhance identity preservation, recognition accuracy, and visual realism at the same time.

The DVN structure includes a fixed ArcFace-based face encoder, two layers of dual-view generators, and two sets of discriminators. The Layer-1 generators learn to map a source image to its frontal and side-view normalized versions. Layer-2 generators further hone these outputs in order to have better identity preservation by reducing the cross-pose identity loss. Two discriminators (global and local) work to guarantee photo-realistic outputs of the final images. Several loss functions optimally designed as identity loss, reconstruction loss, symmetry loss, and Wasserstein GAN-based adversarial loss are employed in the model so that it would have an adequate balance between fidelity and consistency in identity.

Authors test their approach on constrained datasets as well as unconstrained ones. For constrained environments, experiments are performed on the Multi-PIE dataset, whereas for unconstrained conditions, the IJB-A and IJB-C datasets are utilized. Extra training data is taken from CASIA-WebFace. In their experimental configuration, DVN performed better than most state-of-the-art normalization methods, particularly for large pose variations. An ablation study proves the worth of every architectural element, ensuring that the dual-view configuration and two-layer generation scheme both play important roles in performance improvement.

In summary, the DVN approach obtained a Rank-1 identification rate of 97% on Multi-PIE and robust verification performance on IJB-A and IJB-C. Employing mean-fused identity vectors of both normalized views offered improved recognition compared to single-view or concatenated feature approaches. The work places DVN as a robust and identity-preserving face recognition method for face recognition in diverse poses, making it especially beneficial for real-world forensic and surveillance applications.

2.5 Paper-5: Instance-Level Heterogeneous Domain Adaptation for Limited-Labeled Sketch-to-Photo Retrieval

This work[5] introduces a new paradigm called *Instance-Level Heterogeneous Domain Adaptation (IHDA)* to deal with the issue of sketch-to-photo retrieval where labeled target data is scarce. Conventional domain adaptation methods typically do not perform well in such scenarios because of domain discrepancies and disjoint identity spaces between the source and target collections. The proposed IHDA framework aims to effectively transfer instance-level knowledge from a richly annotated photo domain (source) to a sketch-photo domain (target) by leveraging shared attributes and domain adaptation strategies.

The IHDA model consists of dual encoders for the heterogeneous domains (photos and sketches), identity and attribute classifiers, and a domain discriminator integrated through a Gradient Reversal Layer for adversarial training. Identity learning is addressed through classification and triplet losses, whereas attribute learning is directed by common attribute spaces formed from non-color attributes in the source domain. Entropy minimization and attribute consistency losses are employed for target domain unsupervised learning to allow the model to successfully bridge the modality gap. The framework jointly minimizes marginal and conditional domain distribution differences for stronger feature alignment.

The authors compared IHDA with three pairs of datasets: UT-Zap50K $\rightarrow$ QMUL-Shoes, CelebFaces $\rightarrow$ IIIT-D Viewed Sketch, and Market-1501 $\rightarrow$ PKU-Sketch. In all situations, IHDA performed better than a variety of state-of-the-art alternatives. On QMUL-Shoes, it produced a rank-1 accuracy of 68.7%, an increase by 7%. On the difficult IIIT-D Viewed Sketch, it obtained 95.7%, and on the demanding PKU-Sketch, it produced an impressive 85.6% rank-1 accuracy—well above baselines by more than 50. Large-scale ablation experiments confirmed the role of each part, demonstrating that both adversarial domain adaptation and attribute-guided learning are essential.

In addition to its strong empirical performance, the IHDA framework holds practical value for real-world forensic applications, particularly in scenarios where sketch-based data is scarce and expensive to annotate. Its ability to adapt knowledge from richly labeled datasets into limited-label or unlabeled domains addresses a key bottleneck in sketch-photo retrieval systems. The modular design of the IHDA pipeline also allows for easy extension into other cross-modal tasks such as sketch-to-video retrieval or surveillance-based attribute tagging, making it a robust and scalable solution for future developments in heterogeneous face recognition.

The paper shows that IHDA facilitates efficient knowledge transfer for cross-modal retrieval without labeled data in the target domain. Its architecture makes it applicable in forensic and surveillance scenarios where identification using sketches is necessary but data labeling is prohibitive or costly.

2.6 Summary of Literature Survey

Forensic facial sketch identification is now a critical law enforcement tool where photographic evidence cannot be provided. Conventional methods used manual sketching and rule based comparison methods, but these could not adequately address the artistic style inconsistencies, memory changes in witnesses, and the inherent gap between sketches and actual photographs. Early machine learning algorithms, including Support Vector Machines (SVM) and hand-designed feature extraction methods, tried to fill this gap but were not flexible enough to cope with the variety and complexity of real-world situations. With the advent of deep learning, forensic face recognition has made tremendous progress.

Generative Adversarial Networks (GANs) and Convolutional Neural Networks (CNNs) have been especially influential, allowing for more precise sketch-to-photo conversions. GAN-based architectures, like cyclic GANs, have played crucial roles in maintaining identity attributes while bringing the sketches closer to realism. In addition, Domain Alignment Embedding Networks (DAENs) have enhanced recognition accuracy by projecting sketches and photos into the same feature space. Nevertheless, even though these models are highly successful, they need large, high-quality datasets to generalize across demographics and sketching styles, hence deployment in real-world scenarios is challenging. To tackle these challenges, researchers have developed identity-preserving methods, including feature decoupling learning and identity-aware super-resolution models, that maintain essential facial features throughout the transformation process. Other methods, such as dual-view normalization techniques, have assisted in normalizing facial representations across different sketching styles to enhance consistency in recognition.

However, issues remain—many current models lack computational efficiency and do not generalize well to low-resolution or partial sketches, both of which occur frequently in forensic use. There has also been recent work with hybrid models fusing CNN-based feature extraction with domain adaptation frameworks in order to enhance forensic recognition in practical environments. One such new direction is Instance Level Heterogeneous Domain Adaptation (IHDA), enabling more robust matching across datasets with a small amount of labeled data. There has also been an advance with multifeature fusion methods with the combination of Histogram of Oriented Gradients (HOG) and Gabor Wavelet Transform(GWT) in coping better with occlusion and varying lights, rendering forensic sketch recognition more versatile. Despite these advancements, significant challenges remain. Many deep learning models require substantial computational power, making them difficult to deploy in real-time forensic investigations. Future research needs to focus on developing lightweight architectures that maintain accuracy while being efficient enough for real-world use.

Additionally, ensuring cross-demographic generalization is critical—models must be trained on diverse datasets to avoid bias and ensure fairness across different age groups, ethnicities, and artistic styles. Looking ahead, the combination of self-supervised learning and transformer-based visual models could unlock new avenues to enhancing forensic face recognition, which will make such systems more trustworthy and universally usable in law enforcement.

Table 2.1: Summary of Literature Review on Forensic Face Sketch Recognition

S.No	Paper Title	Keywords	Problem Domain	Methods Used	Limitations	Future Work	Findings
[1]	Domain Alignment Embedding Network for Sketch Face Recognition	Sketch recognition, deep metric learning	Face Sketch-Photo Matching	Domain Alignment Embedding Network (DAEN)	Struggles with diverse sketch styles	Improve style robustness, real-time deployment	High accuracy across datasets
[2]	Feature Encoder Guided GAN for Photo-Sketch Synthesis	GANs, image translation	Photo-Sketch Synthesis	Cyclic GANs with two generators/ discriminators	High computational cost	Adaptive loss, diverse datasets	Improved photorealism, recognition
[3]	Identity-Preserved Model for Face Sketch-Photo Synthesis	Identity preservation, translation	Face Synthesis	U-Net, dual discriminators, ResNet-50, Light CNN-29v2	Demographic limitations	Better sketch-style adaptation	85%+ recognition with identity retention
[4]	Dual-View Normalization for Face Recognition	Face normalization, CNN, ArcFace	Face Recognition	Light CNN with Dual-View Normalization	Sketch quality sensitivity	Self-supervised learning	Better sketch-photo consistency

[6]	Deep Learning-Based Architecture for Photo-Sketch Recognition	CNN, Morphable Model	Face Recognition	3D Morphable Model, DCNN, transfer learning	Needs large training data	Improve sketch realism, data expansion	Better facial variation handling
[5]	IHDA for Limited-Labeled Sketch-to-Photo Retrieval	Domain adaptation, cross-modal	Sketch-to-Photo Retrieval	IHDA framework	Labeled data dependency	Explore category-level retrieval	80%+ accuracy in forensic settings
[7]	SP-Net: Framework to Identify Composite Sketch	Composite sketch, CNN, contrastive loss	Composite Sketch ID	SP-Net, Siamese net, VGG-Face	Struggles with hand-drawn sketches	Adapt to Identikit-based images	28.3% rank-1, 80% rank-10 (E-PRIP)
[8]	Realistic Photo–Sketch via Composition-Aided GANs	GAN, face parsing, translation	Face-Sketch Synthesis	CA-GAN, SCA-GAN	FID metric issues	Improve realism, FID reliability	99.7% NLDA accuracy
[9]	Multi-Task Explainable Quality Networks for Forensics	Image quality, explainable AI	Forensic Recognition	XQNet-ConvNet, EfficientNet	Bias in low-quality images	Edge device optimization	Enhanced interpretability, bias reduction

[10]	Domain Balancing on Long-Tailed Domains	Domain balancing, class imbalance	Face Recognition	Residual balancing mapping	Poor real-time adaptation	Dynamic domain adaptation	96.56% VA@FAR= 10^{-6}
[11]	CCA Feature Fusion with Patch of Interest	CCA, patch-based local matching	Sketch Image Retrieval	CCA with dynamic patches	Computational cost	Improve scalability	Better local feature matching

2.7 Comparative Analysis

This section provides a detailed comparative study of the models analyzed in the literature review, classifying them according to the main challenges they tackle in forensic face sketch recognition. These challenges are sketch-to-photo synthesis, identity preservation, domain adaptation, style transfer, and generalization across demographics. The comparison in the next table provides insightful information on the strengths, weaknesses, and performance features of each model, which shows trade-offs between recognition accuracy, computational complexity, training data needs, and their applicability in real-world forensic usage.

Summary of Comparative Analysis

- GAN-style model, which are very good at domain gap reduction (accuracy up to 85%) but need a lot of computation.
- Identity-preserving methods, for example, feature decoupling with LightCNN, achieve very high accuracy (90 per cent) by emphasising the preservation of identity characteristics, but at the computational cost.
- Normalisation methods such as dual-discriminator U-Net manage the variation in sketch quality well, but suffer a loss in performance when there are extreme stylistic departures.
- Super-resolution and graph-regularized models provide flexibility in the low-resolution setting, and identity-aware networks reach accuracy of 85 percent or more on such inputs.

- Hybrid models and IHDA work well in real-world situations but need huge amounts of data to handle a variety of situations.

This comparative study provides guidance on which models to use in which forensic situations. Both methods have their own strengths and weaknesses, and will enable us to create better forensic sketch recognition systems.

Table 2.2: Comparative Study of Techniques Used for Forensic Face Recognition

Challenge	**Model**	**Technique**	**Strengths**	**Limitations**	**Accuracy**
Domain Gap Reduction	Cyclic GAN Framework	Two generators, two discriminators	Maintains identity, effective at sketch-to-photo synthesis	Requires large dataset, high computational cost	80–85%
	Domain Alignment Embedding Network	Embedding feature alignment	Robust feature harmonization across sketch and photo domains	Moderate performance in highly stylized sketches	78–82%
	Dual-View Normalization with LightCNN	Dual-view normalization with ArcFace	Consistency across artist styles	Sensitive to extreme variations in sketch quality	$\sim$80%
Identity Preservation	Identity-Preserving GANs (IPAM)	Identity attributes with GANs	Retains core identity traits during transformation	High training cost, limited adaptability to diverse databases	$\sim$85%
	Feature Decoupling Learning + LightCNN	Separation of identity and non-identity features	Effective separation for better identification	Reduced performance with lower-resolution sketches	90%+

Challenge	Model	Technique	Strengths	Limitations	Accuracy
Sketch Quality & Style Variability	U-Net with Dual Discriminators	Leaky ReLU, identity-verifying models	Good adaptation to variable sketch quality	Moderate recognition drop with extreme styles	75–80%
	Dual-View Normalization with ArcFace	Normalization across views	Adapts to stylistic & quality variations	Reduced effectiveness on low-quality sketches	∼80%
Low-Resolution /Blurred Sketches	Identity-Aware Super-Resolution Network	Super-resolution with identity preservation	Improves low-resolution sketch quality	Limited applicability in highly blurred sketches	∼85%
	Graph-Regularized Dictionary Learning	Graph regularization and residual learning	Effective with low-detail, low-res sketches	Performance declines on high-variance datasets	70–80%
Adaptability to Real-World Conditions	HOG + Gabor Wavelet Fusion with CCA	Multi-feature fusion	Handles occlusion, lighting variability	Limited performance in low-resolution cases	75–80%
	Instance-Level Heterogeneous Domain Adaptation (IHDA)	Instance-level domain adaptation	Effective cross-domain matching in complex environments	Requires substantial training data for adaptation	80%+

2.8 Analysis of Existing Research Gaps

1. **No Unified System:** Current solutions lack an integrated approach combining face sketch construction and recognition, leading to disjointed workflows.
2. **Missing Recommendations:** No system offers AI-driven facial feature recommendations during sketch creation, limiting consistency and realism.

3. **GAN-Based Models:** GANs excel in reducing domain gaps (up to 85% accuracy) but demand high computational resources, hindering real-time use.
4. **Identity Preservation:** Methods like feature decoupling with LightCNN achieve 90% accuracy by preserving identity, yet they're computationally expensive.
5. **Normalization Challenges:** Dual-discriminator U-Net handles sketch quality variations well but struggles with extreme stylistic differences, reducing reliability.
6. **Low-Resolution Handling:** Super-resolution and graph-regularized models, along with identity-aware networks, offer flexibility for low-resolution inputs (85%+ accuracy).
7. **Hybrid Models & IHDA:** These perform well in real-world scenarios but require large, diverse datasets to generalize effectively across cases.

CHAPTER 3
PROPOSED SYSTEM

3.1 Introduction

The proposed system strategically adopts a three-phase architecture to address the inherent complexity and multi-faceted nature of forensic sketch-based identification. Traditional methods that attempt to directly map sketches to photographic images often suffer from low accuracy, limited interpretability, and a lack of user control. By breaking down the task into face-based sketch creation, sketch enhancement, and face recall, the system recognizes that successful identification is not a linear transformation but a layered, successive approximation of human facial identity. This conceptual framework allows each stage to address a specific cognitive and technical issue, making the resulting pipeline more reflective of how human recognition and memory construction naturally work.

The first step, feature-based sketch production, is based on a realistic model of human memory, especially during stressful or traumatic situations typically related to observing a crime. Psychological studies have repeatedly demonstrated that people hardly ever remember a whole face as a unified entity; instead, they recall individual features—like the shape of the eyes, the shape of the lips, the shape of the jawline, or the cut of the hair. These piecemeal memories can differ in vividness and accuracy, so it is unrealistic and cognitively taxing to ask a witness to construct a complete face from memory all at once. Acknowledging this limitation, the system brings in a modular solution that enables forensic professionals and witnesses to construct a composite sketch feature by feature, based on a list of AI-suggested options. This approach greatly minimizes the intellectual load on the witness by limiting the field of decision-making at each stage, facilitating a more precise and confident recall of personal features. By allowing the face to be regarded as an assembly of independent but interplaying elements, the system simulates natural intellectual processes, and the sketch creation stage becomes more intuitive and less vulnerable to error. Furthermore, the modular approach to design guarantees that the resulting sketch is logically structured, allowing it to be easier for the next refinement and retrieval phases to interpret and process the input in a useful manner. This stage thereby not only maximizes the system's usability and efficiency but also guarantees that the forensic process continues to be compliant with the realities of human memory and investigative practice, providing a solid foundation for the next phases.

The second stage, sketch refinement, bridges the gap between rough composite sketches and the visual diversity needed for dependable face retrieval. Rather than leaping from feature-based drawings directly to gallery search, this middle improvement stage enables the system to enhance coherence, detail, and recognizability of the face while maintaining primary identity features. Conceptually, this helps to overcome the modality gap between drawn or feature-based sketches and photographic images, a long-standing hindrance to forensic face matching. It guarantees that retrieval models are run on inputs more proximate in distribution to actual images, and thus both robustness and fairness of the retrieval process are improved.

The third and last phase, identity-preserving face retrieval, rounds off the pipeline by converting the cleaned-up sketch into actionable outcomes. In place of absolute visual matching, the retrieval stage prioritizes feature-wise similarity comparison, providing forensic practitioners with more insight into how and why a match was proposed. Not only does the stage boost the credibility of the system in actual legal applications, but it also enables users to make well-informed choices instead of relying blindly on automated results. By designing retrieval as a separate phase, the system sidesteps the black-box opacity that usually characterizes end-to-end approaches and promotes higher accountability and interpretability.

All three phases are integrated smoothly into a full, user-friendly web platform built using the MERN (MongoDB, Express.js, React.js, Node.js) stack. This platform guarantees that the advanced backend models are hidden from the user, with an intuitive, interactive interface that can be easily used by forensic experts without technical knowledge. From choosing AI-suggested features to enhance sketches and trigger searches, all interactions are optimized to enable collaboration between human input and machine intelligence. By putting the technical process in a comprehensible, usable application, the platform takes the arcane machine learning operations and converts them into useful forensic tools waiting to be deployed in the real world.

In all, the modular, step-wise approach paired with an easy-to-use web interface enables the system to fulfill its primary goals: enhancing identification precision, promoting transparency in the system, aiding forensic processes, and providing long-term scalability. Each step is tailored to overcome a significant challenge in forensic face matching, and by compartmentalizing these issues into bite-sized units, the system fosters a strong and flexible basis for future innovation. Ultimately, the platform reimagines forensic face sketch identification as a collaborative, technology-supported practice that more effectively serves investigative purposes.

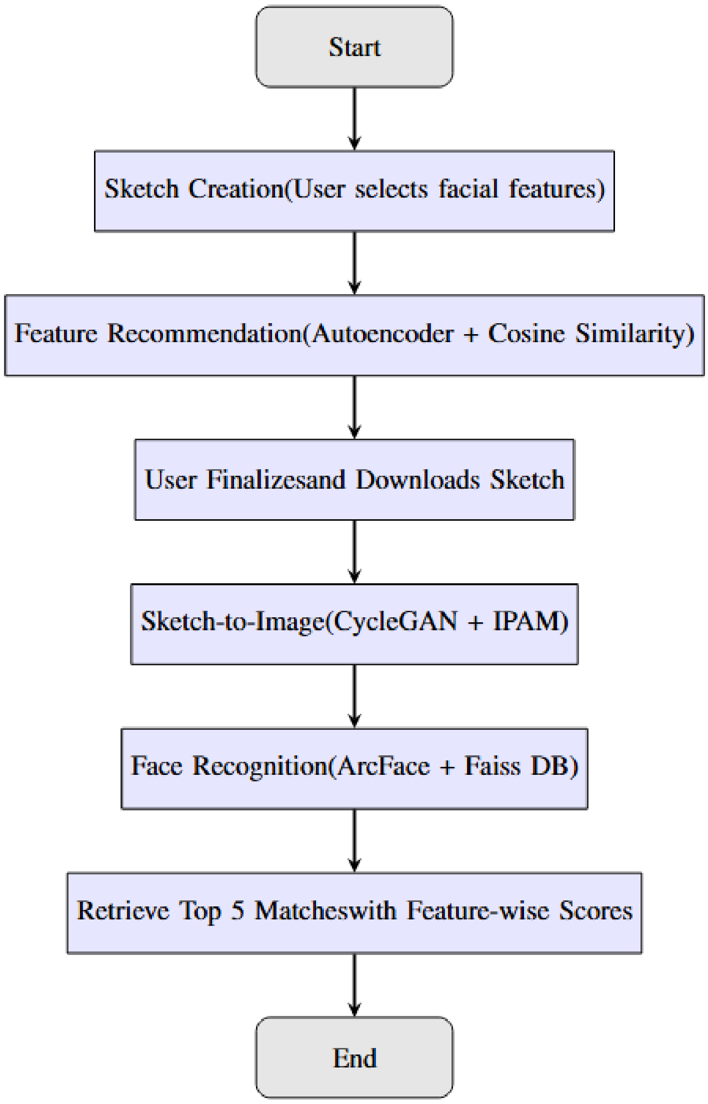

Figure 3.1: Flowchart for Forensic Face Sketch Construction & Recognition System

3.2 Methodology

3.2.1 Facial Feature Recommendation Model

The initial process of the system is aimed at enabling high-quality composite sketching through an AI-based feature suggestion module. The users are provided with the provision to manually select fundamental facial structures like eyes, nose, lips, eyebrows, hair, and beard from a pre-curated dataset. However, the choice is supplemented by smart suggestions so that anatomical consistency and reasonable composition are assured.

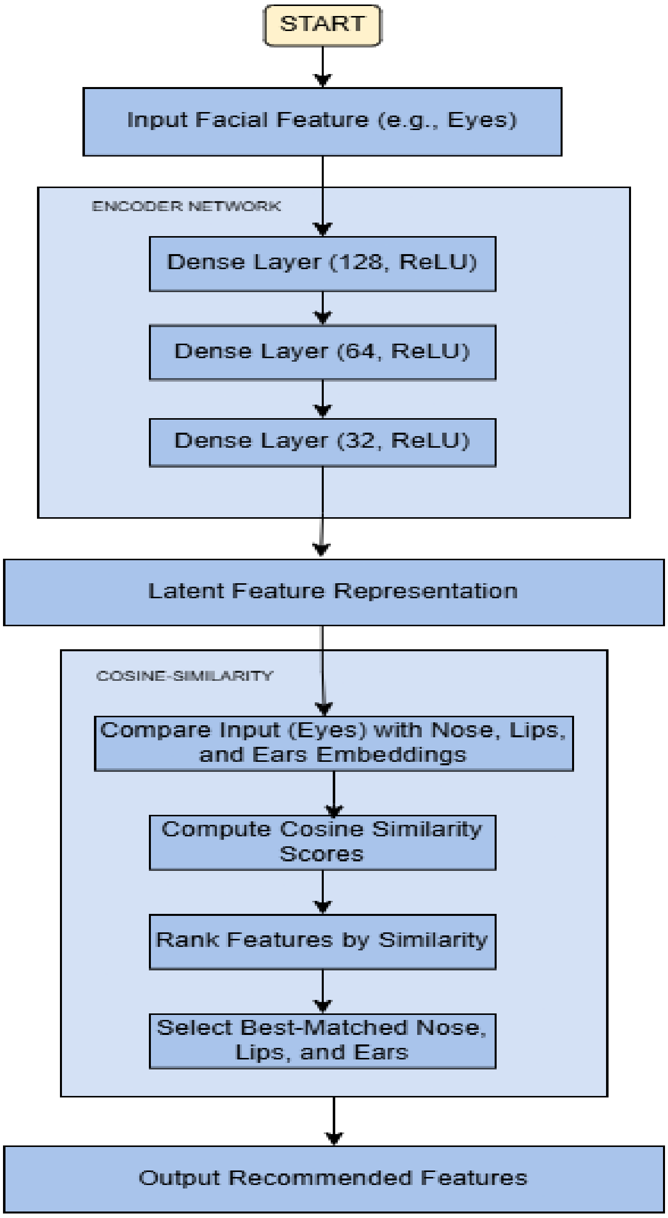

Figure 3.2: Flowchart for Facial Features Recommendation Model

Such a recommendation system is developed over a deep autoencoder model trained on the collected facial feature datasets. The encoder part encodes the visual data of every feature into a compact latent embedding that well represents fundamental morphological features. Employing cosine similarity metrics calculated over such embeddings, the system finds and recommends features that are most compatible with the user-chosen component. This methodology reduces unrealistic amalgams and aids users, particularly those with no advanced forensic artistry skills, in creating internally consistent and realistic sketches.

The autoencoder is separately trained for each type of facial region, and embeddings capture fine-grained features. This is possible to the extent that, for example, the system can suggest a nose shape suitable for a chosen pair of eyes to provide demographic consistency in age, ethnicity, and gender features. The recommendation process takes place in real time, giving users instant feedback and leading them to produce sketches that are better suited as inputs for the later refinement and retrieval phases.

3.2.2 Sketch-to-Refined Image Synthesis Model

After completing the composite sketch, the next phase targets improving the sketch into a higher-resolution, more detailed, more refined, and identity-conserving image. In contrast with the general sketch-to-photo translation methods that try to produce photorealistic results, this model prioritizes maintaining the distinctive identity features in the initial sketch while detailing the image with more detailed facial features and enhancing structural consistency.

This model uses a tailored version of the CycleGAN architecture with Identity-Preserving Attention Modules (IPAM) that are specifically designed to guide the attention of the network to key identity-defining areas of the face. The model does not try to hallucinate excessive photorealistic details; instead, it improves the sketch by sharpening contours, shading, texture gradients, and feature definitions. The modules of attention are designed to ensure that important characteristics like eye shape, lip shape, and nose shape are accentuated and retained, with a very high level of fidelity to the original intent of the sketch.

Model training uses a compound loss function combining adversarial loss (for overall realism), cycle-consistency loss (to impose bidirectional reconstruction), perceptual loss (to maintain structural attributes at a high abstraction level), and identity-preserving loss (with emphasis on retaining distinct facial features). The training set includes unpaired sketches and face photographs, reinforced by augmentation techniques like rotations, scalings, and injection of noise to improve robustness against sketch change.

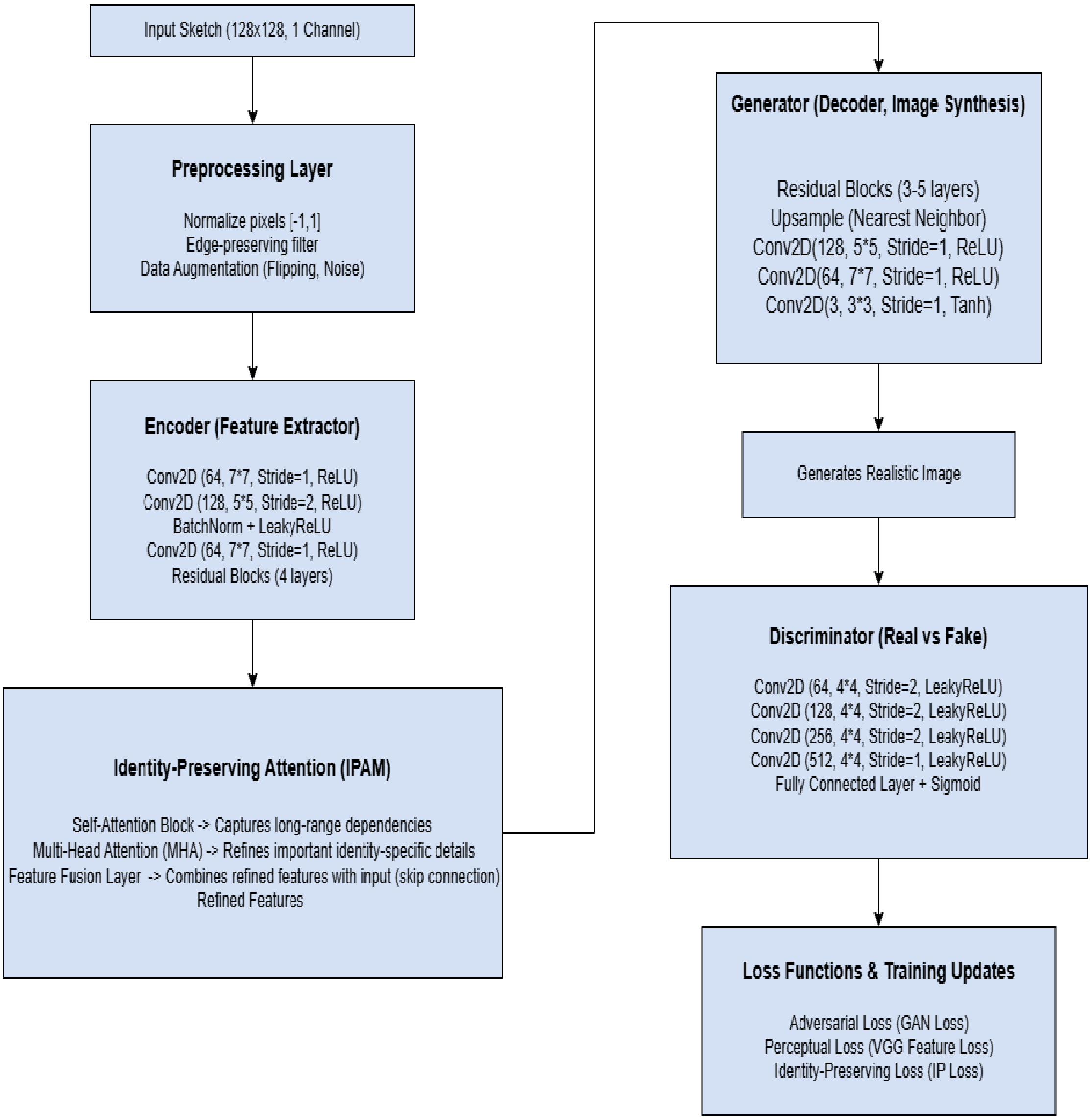

Figure 3.3: Flowchart for Sketch to Digital Image Conversion Model

3.2.3 Face Recognition and Feature-Wise Matching Model

The third and last stage of the pipeline is devoted to suspect retrieval via a two-level matching process: global face matching and localized feature-wise matching. Such a dual strategy improves the interpretability and explainability of retrieval outcomes, enabling forensic analysts to evaluate not only the general quality of the match but also the contribution of certain facial areas.

The enhanced image generated in the preceding step is passed through a face recognition network developed based on the ArcFace architecture and having a ResNet-50 backbone. The network transforms the face into a 512-dimensional embedding optimized for angular margin-based separation, such that embeddings of the same identities group tightly while those of dissimilar identities are far apart.

Region-specific embeddings are also extracted for the eyes, nose, and lips simultaneously. These feature-wise embeddings allow the system to calculate localized similarities, giving a feature-wise breakdown of which features best match retrieved candidates. Retrieval is done via FAISS (Facebook AI Similarity Search), a library for optimized large-scale nearest-neighbor search. Global embeddings and feature-wise embeddings are both indexed separately so that retrievals can be made based on holistic similarity as well as localized feature matches.

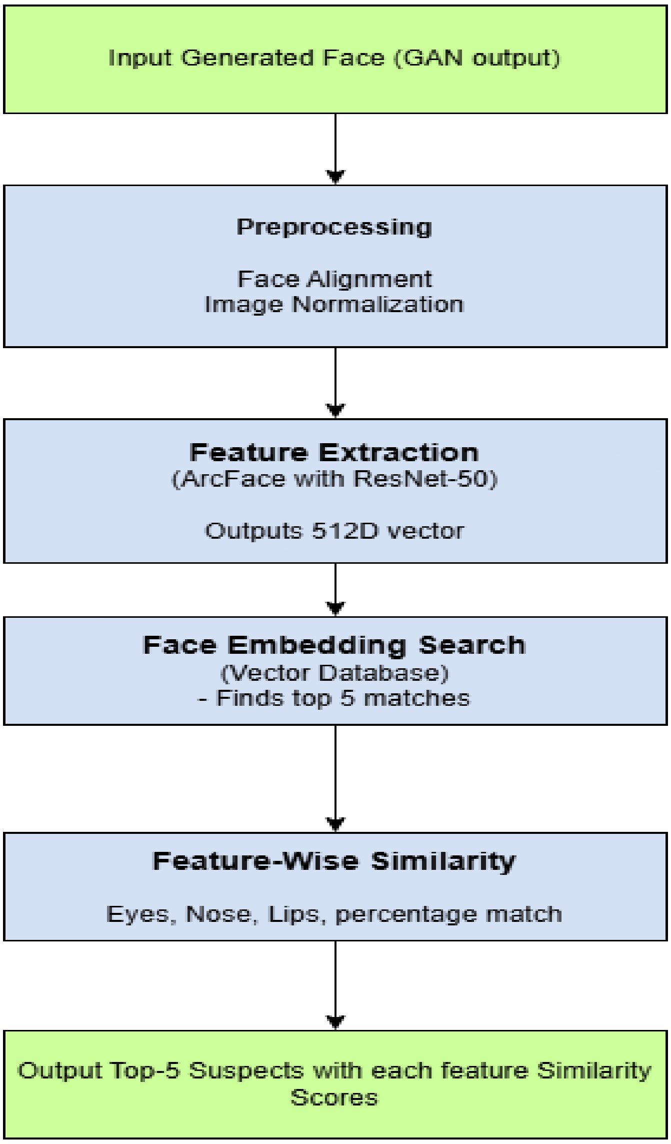

Figure 3.4: Flowchart for Face Recognition Model

By offering the users not only a list of nearest matches but also a report of feature-based similarity, the system equips forensic professionals to make more logical choices. In this way, the explainability of the retrieval process is markedly improved, and this is crucial in forensic deployments where interpretability and accountability hold the key to success.

3.2.4 Data Processing Pipeline

The performance of the suggested models relies on the quality and integrity of the foundational data. The importance of powerful data preprocessing workflows is thus imperative. For datasets of facial features, images are manually curated and annotated to be of high quality across different demographics. Features are resized to the same resolution, normalized in terms of pixel intensity, and transformed randomly to promote diversity. While training the feature recommendation model, balanced presentation of feature types is ensured not to introduce any bias.

The training sketches for the refinement model are preprocessed to improve definitions of edges and eliminate noise. Histogram equalization and bilateral filtering are used in order to equalize sketch quality. Face images employed in face recognition are processed through landmark detection and alignment processes to maintain uniform orientation and scale, essential for the success of deep metric learning. Precomputed embeddings for every suspect database record are indexed in optimized FAISS indices, which facilitate quick query response times during retrieval processes. The data pipeline is scalable, making it simple to expand the database as new suspect records are incorporated.

3.2.5 Platform Development: MERN Stack

In order to present the system as an easy-to-use and production-grade tool, a web-based platform is implemented based on the MERN stack, which features modularity, scalability, and great responsiveness. The frontend, implemented in React.js, offers a dynamic user interface supporting feature selection, composite sketch composition, real-time display of recommended features, finalizing the sketch, and visualization of retrieved results. The interface is kept simple, reducing the learning process for forensic experts while providing great depth in functions.

The system's backend infrastructure, implemented in Node.js and Express.js, serves as the critical communication conduit between the user-confronted frontend and the specialized deep learning microservices. It is tasked with processing and responding to API requests for feature-based sketch generation, uploading and managing user-inputted sketches, requesting the refinement of sketches into higher-resolution images, and carrying out identity-preserving face retrieval queries. In addition to these fundamental functions, the backend also incorporates basic forensic functionality in the form of secure authentication of users, workflow integrity maintenance using session management, and stable activity logging facilities, where all interactions remain traceable, secure, and aligned with forensic requirements.

The database layer, which is based on MongoDB, holds user-created sketches, chosen features, processed images, suspect database records, and retrieval logs. The schema design is optimized for efficient querying and retrieval, which is essential for real-time responsiveness in investigative processes. The whole system is containerized with Docker, and each model service can be deployed as a microservice. This approach to architecture makes it simpler to update, maintain, and scale, allowing the system to easily accommodate changing needs in forensic applications.

3.3 System Design and Architecture

3.3.1 System Architecture Overview

The proposed system employs a layered modular architecture designed to separate functionalities across different tiers, maximizing maintainability, scalability, and fault tolerance. The architecture enforces a clean separation of concerns, such that changes or failures within one layer do not negatively impact the others. Every tier is optimized for certain purposes, and they communicate via secure, authenticated RESTful APIs to preserve data integrity and improve system resilience.

The major architectural components are as follows:

1. User Interface Tier
The User Interface (UI) Tier constitutes all those client-side elements on which the end-user has a direct interface. This tier is accountable for the satisfactory, intuitive, and responsive experience. It comprises:

- **Feature Selection Tools:** The interactive panels that enable users to pick out individual facial features like eyes, nose, mouth, etc., while constructing the sketch.

- **Sketch Visualization:** In-the-wild rendering of the face sketch as the user chooses or edits features to provide instant feedback.

- **Recommendation Panels:** Context-based suggestions for facial features (e.g., styles of eyes or noses) based on dynamically created suggestions derived from previous user input or description selections.

- **Retrieval Result Displays:** A graphical gallery showing search results when the user tries to retrieve similar faces from the database based on similarity scores.

2. Service Tier

The Service Tier is the system's backbone, managing all backend operations and coordinating communication among the UI, inference models, and database. Its role is:

- **Routing Requests:** Handling API requests from the UI and routing them to the correct microservices or database endpoints.
- **Input Validation:** Preventing incoming data — e.g., chosen features or sketches — from violating expected formats and constraints, reducing the risk of security attacks.
- **Model Interface Management:** Managing the invocation of model inference services, bundling inputs, and unbundling outputs in a frontend-friendly format.
- **Database Interaction:** Coordination of read and write operations to and from the MongoDB database, including data caching mechanisms for quicker retrieval.

3. Model Inference Tier

Model Inference Tier is a key section where the intelligence of the system lies. It consists of several deep learning models running as individual microservices. Each model has a specific function:

- **Feature Recommendation Model:** Proposes potential facial features that enhance the user's existing sketch, according to acquired aesthetic and forensic rules.
- **Sketch Refinement Model:** Performs post-processing on user-drawn sketches to enhance their realism or align with forensic requirements.
- **Face Retrieval Model:** Compares the finished sketches to available records within the database using embeddings and provides the best matches.

The microservices architecture enables each model to scale independently in accordance with usage patterns. For instance, as face retrieval requests intensify in a particular session, the retrieval service will be scaled without impacting the other services. Containerization with Docker or another such technology ensures environment consistency and effortless deployment across different levels.

4. Data Storage Tier

The Data Storage Tier maintains a centralized MongoDB database that securely stores both structured and unstructured data, supporting efficient querying, scalability, and fault tolerance. The primary data components managed in this tier are:

- **User Actions:** Logs of user activities, including feature selections, sketch modifications, retrieval attempts, and system feedback, to facilitate analytics and continuous system improvements.
- **Sketches:** Raw user-generated sketches along with all intermediate versions created during the sketch refinement and feature recommendation processes.
- **Refined Images:** Enhanced or realistically reconstructed versions of sketches produced through model-based post-processing, ensuring better matchability during retrieval.
- **Embeddings:** High-dimensional vector representations of sketches and images, generated via deep learning models, that enable efficient and accurate similarity-based search and retrieval.
- **Criminal Database:** A specially curated collection containing:
  - Mugshots and photographs of known or suspected individuals.
  - Personal Metadata, such as name, age, gender, height, physical descriptors (scars, tattoos, etc.), and criminal history.
  - Associations with Cases, such as linked crimes, case statuses, and investigation notes, which are stored to assist forensic investigators during retrieval analysis.

The criminal database is integrated seamlessly with the retrieval model, ensuring that search operations are fast, secure, and compliant with legal and ethical data handling standards. Indexes, sharding strategies, and data backup policies are employed to maintain high performance, minimize latency, and prevent data loss.

3.3.2 Component-Wise Architecture

Frontend (React.js):

- Facial Feature Selection Module
- Sketch Creation and Finalization Interface
- Real-time Recommendation Display

- Sketch Refinement and Identity-Preservation Module
- Facial Recognition Interface
- Retrieval Results Visualization

Backend (Node.js + Express.js):

- Feature Recommendation Service
- Sketch Management Service
- Refinement Inference Service
- Retrieval Query Service

Deep Learning Model Services:

- Autoencoder with Cosine Similarity for feature recommendation
- GAN-IPAM for sketch refinement
- ArcFace-ResNet50 + FAISS for retrieval

Database (MongoDB):

- Users Collection
- Facial Features Collection
- Sketches Collection
- Refined Images Collection
- Criminal Details Collection
- Retrieval Logs

CHAPTER 4
IMPLEMENTATION

The real-world implementation of the system proposed includes the meticulous co-ordination of various components, from preparing datasets to training deep learning models, backend API creation, and frontend integration. Every module of the system is developed with the primary objective of obtaining accuracy, usability, scalability, and modularity so that the end product can be easily deployed in actual forensic settings.

Implementation into steps was methodically planned out, starting from a sound dataset specific to the forensic arena. Next came specialized deep-learning architecture being tuned or created to be used at various stages in the pipeline while guaranteeing every model was intended to meet an explicit requirement — from sketch upgrade to identity-relevant feature retrieval. Particular emphasis was on the deployment design with technologies picked with the primary focus of leveraging maximum responsiveness, reliability, and ease of maintainability of the system.

By breaking the project into distinct phases, we were able to ensure that every phase — data preparation, model building, database design, retrieval logic, and platform deployment — was optimized step by step and validated stringently before integration into the entire system. This modular approach also offers flexibility for future improvements or extensions, with individual modules being upgradable without impacting the overall process.

The subsequent parts present an itemized, step-by-step description of the system's implementation, starting with the building blocks: data preprocessing and collection.

4.1 Data Collection and Preprocessing

4.1.1 Image Collection

A strong and heterogeneous dataset is the foundation of our system. We assembled a vast corpus of criminal face images from various sources in order to preserve real-world variability. These faces included a vast range of conditions, such as variations in pose, lighting, age, ethnicity, facial expression, and background environments. This deliberate variability was essential to make sure that models trained on these data could generalize well to new, unseen forensic sketches. Care was taken to obtain images that mimic the difficulties most commonly encountered during actual criminal investigations, including partial occlusions and non-standard illumination.

4.1.2 Grayscale Conversion

To mimic the nature of forensic sketches, grayscale conversion was applied to all images gathered using OpenCV. Nonetheless, in contrast to traditional grayscale processing, every image was stored in a three-channel format with the same grayscale values in every channel. This was used to ensure compatibility with deep learning models that had prior pre-training on RGB images to prevent re-engineering of input pipelines. This strategy allowed the models to be modified rapidly while preserving structural likeness to the source pre-trained networks and thereby optimizing transfer learning.

4.1.3 Feature Extraction

A critical preprocessing step involved the extraction of nine key facial features that play a major role in forensic identification. These features were:

- Eyes (Left and Right)
- Nose
- Mouth
- Eyebrows
- Ears
- Chin
- Cheeks
- Forehead
- Jawline

This set of features was selected from forensic science literature with a focus on the facial regions most consistently remembered by witnesses. Features were manually and semi-automatically extracted from images using facial landmark detection models. The features were then systematically stored to be utilized later during the sketch construction process, providing a modular and controlled input system for users.

4.2 Forensic Sketch Construction Interface

4.2.1 React.js Frontend Development

A simple, intuitive web-based interface was constructed utilizing React.js for maximum forensic analyst, investigator, and even non-technical user accessibility. The frontend is designed with high focus on modularity, responsiveness, and visualization. The use of React component architecture enabled dynamic facial feature loading, real-time handling, as well as delivering real-time user feedback throughout sketch development.

4.2.2 Drag-and-Drop Feature Implementation

The central interaction paradigm was drag-and-drop. Users would be able to pick from a pre-extracted facial feature gallery and drop them onto an empty canvas. Positioned there, the features could be:

- Resized
- Rotated
- Translated (moved around)

Real-time transformations allowed users to adjust the position of each feature until a composite sketch closely resembling the witness's recollection was produced. This process closely resembles traditional forensic artist workflows, providing practitioners with familiarity while improving speed and reproducibility.

4.2.3 Real-time Preview and Validation

Users get instant previews of their changing sketches, allowing them to refine iteratively. By providing ongoing feedback, the site minimized user errors and cognitive overload, making the process of sketch construction efficient and user-friendly.

4.3 Facial Features Recommendation Model

The facial feature suggestion model maps an input feature (e.g., eyes) to a low-dimensional latent space by utilizing an encoder neural network with sequential dense layers and ReLU activation (128 $\rightarrow$ 64 $\rightarrow$ 32 units). The latent representation retains important feature-specific information. Then, the model calculates cosine similarity between the input feature embedding and the embeddings of other facial parts (nose, lips, ears) in the system. Through the assessment of similarity scores, the model ranks the candidate features and chooses the most harmonious matches. This allows coherent and visually pleasing facial composites to be generated, which is essential in applications such as digital face synthesis and forensic sketching.

Algorithm 1 Facial Feature Recommendation Model

Input: Facial feature (e.g., Eyes)
Output: Recommended Nose, Lips, and Ears
Encode input feature using Encoder Network:
  Dense Layer (128 units, ReLU activation)
  Dense Layer (64 units, ReLU activation)
  Dense Layer (32 units, ReLU activation)
Obtain Latent Feature Representation
for each stored feature (Nose, Lips, Ears) **do**
  Encode stored feature
  Compute Cosine Similarity with input feature
end for
Rank stored features based on similarity scores
Select the best-matched Nose, Lips, and Ears
Return Recommended features

4.4 Sketch Refinement and Identity Preservation

4.4.1 Overview

Though composite sketches are useful, they are not rich enough in photorealistic detail to support useful deep learning-based matching. So, an intermediate translation step was added to detail and improve the sketches to prepare them for identity-preserving retrieval.

4.4.2 Model Integration

To enable successful translation from sketches to images, two models were combined:

- **IPAM (Identity Preserving Adversarial Model):** Designed for converting sketch features into detailed facial textures while maintaining identity characteristics.
- **GAN (Generative Adversarial Model):** Used to ensure cycle-consistency between the sketch and processed image domains, such that identity details were not discarded during translation.

The two-model configuration improved robustness and visual quality and guaranteed that the output images were consistent with the input sketches.

Algorithm 2 Refining Sketch using GAN with IPAM

Require: Sketch image S (128×128, 1 channel), real image I_{real}, hyperparameters $\lambda_1, \lambda_2, \lambda_3$

Ensure: Generated image I_{gen}

Preprocessing:
  Normalize S to $[-1, 1]$
  Apply data augmentation (random flipping, noise)
Feature Extraction (Encoder):
  $F_1 \leftarrow \text{Conv2D}(S, 64, 7 \times 7, \text{stride} = 1, \text{ReLU})$
  $F_2 \leftarrow \text{Conv2D}(F_1, 128, 5 \times 5, \text{stride} = 2, \text{ReLU})$
  $F_3 \leftarrow \text{Conv2D}(F_2, 256, 5 \times 5, \text{stride} = 2, \text{ReLU})$
  $F_4 \leftarrow \text{Conv2D}(F_3, 512, 5 \times 5, \text{stride} = 2, \text{ReLU})$
  Apply BatchNorm and Residual Blocks (4 layers) to F_4
Feature Refinement (IPAM):
  $F_{\text{ref}} \leftarrow \text{MultiHeadSelfAttention}(F_4)$
  $F_{\text{fused}} \leftarrow \text{FeatureFusion}(F_{\text{ref}}, S)$ ▷ Skip connection
Image Synthesis (Generator):
  $U_1 \leftarrow \text{Upsample}(F_{\text{fused}}, \text{Nearest Neighbor})$
  $U_2 \leftarrow \text{Conv2D}(U_1, 128, 5 \times 5, \text{stride} = 1, \text{ReLU})$
  $U_3 \leftarrow \text{Conv2D}(U_2, 64, 7 \times 7, \text{stride} = 1, \text{ReLU})$
  $I_{\text{gen}} \leftarrow \text{Conv2D}(U_3, 3, 3 \times 3, \text{stride} = 1, \text{Tanh})$
Discrimination:
  $D_{\text{real}} \leftarrow \text{Discriminator}(I_{\text{real}})$
  $D_{\text{gen}} \leftarrow \text{Discriminator}(I_{\text{gen}})$
Loss Calculation:
  $L_{\text{GAN}} \leftarrow \mathbb{E}[\log D(I_{\text{real}})] + \mathbb{E}[\log(1 - D(I_{\text{gen}}))]$
  $L_{\text{VGG}} \leftarrow ||\phi(I_{\text{real}}) - \phi(I_{\text{gen}})||_2$
  $L_{\text{IP}} \leftarrow ||F(I_{\text{real}}) - F(I_{\text{gen}})||_2$
  $L \leftarrow \lambda_1 L_{\text{GAN}} + \lambda_2 L_{\text{VGG}} + \lambda_3 L_{\text{IP}}$
Training Updates:
  Update Discriminator to maximize L_{GAN}
  Update Generator to minimize L
Output: I_{gen}

4.4.3 FastAPI Backend for Model Serving

The deep learning models were wrapped in a FastAPI backend. The frontend interacted with this backend by:

- Passing user-created sketches as API requests.

- Getting the matching fine-grained, high-resolution facial images.

This design facilitated low latency, stateless communication, and straightforward scaling of the translation service, which is applicable to real-world deployments.

4.5 Facial Matching

This model follows a face recognition pipeline beginning with generated face (GAN output). It preprocesses the image by applying face alignment and normalization to input standardization. Feature extraction utilizes ArcFace along with ResNet-50 and generates a 512D vector capturing facial traits. Face embedding search then cross-checks the vector with a database, searching for the top 5 matches using cosine similarity. Feature-similarity comparison assesses individual facial features (eyes, nose, lips) and calculates a percentage match. The result is the top-5 suspects along with similarity scores that allow for identification. The integration takes advantage of deep learning frameworks such as PyTorch for ArcFace and optimized vector comparison for matching.

Algorithm 3 Face Recognition Pipeline for GAN-Generated Faces

Require: Generated face image I_{gen} (GAN output), face embedding database D
Ensure: Top-5 suspects with similarity scores

Preprocessing:
  Perform face alignment on I_{gen} to standardize orientation
  Normalize I_{gen} pixel values to $[0, 1]$
Feature Extraction:
  Load ArcFace model with ResNet-50 backbone
  $V_{\text{gen}} \leftarrow \text{ArcFace}(I_{\text{gen}})$ ▷ Outputs 512D vector
Face Embedding Search:
  For each embedding V_i in database D:
    Compute cosine similarity $S_i \leftarrow \frac{V_{\text{gen}} \cdot V_i}{||V_{\text{gen}}|| \cdot ||V_i||}$
  Sort similarities S_i in descending order
  Select top-5 embeddings $\{V_{i_1}, V_{i_2}, \ldots, V_{i_5}\}$ with scores $\{S_{i_1}, S_{i_2}, \ldots, S_{i_5}\}$
Feature-Wise Similarity:
  For each top-5 embedding V_{i_k} (where $k \in \{1, 2, \ldots, 5\}$):
    Extract features for eyes, nose, lips from I_{gen} and corresponding database image I_{i_k}
    Compute percentage match P_{i_k} for each feature
    Aggregate P_{i_k} as weighted average across features
Output:
  Return top-5 suspects $\{I_{i_1}, I_{i_2}, \ldots, I_{i_5}\}$ with similarity scores $\{S_{i_1}, S_{i_2}, \ldots, S_{i_5}\}$ and feature match percentages $\{P_{i_1}, P_{i_2}, \ldots, P_{i_5}\}$

4.6 Database Setup

4.6.1 Database Schema Design

A relational database schema normalized to handle criminal data effectively was developed. The schema consisted of two primary tables:

- **Image Table:** Has image IDs and their respective file URLs.
- **Details Table:** Holds text data like name, age, body description, and known criminal history.

Decoupling images from text metadata enabled faster data retrieval, modular updating, and improved query optimization during matching.

4.7 Image Embedding and Indexing

4.7.1 Embedding Extraction and Storage

Each image from the criminal database was fed through the fine-tuned ResNet50 model to obtain its 512-dimensional embedding. These embeddings were saved together with their image IDs for convenience during search.

4.7.2 FAISS Indexing

To perform fast ANN search, FAISS library was utilized. The following settings were used:

- Index Type: Flat Inner Product (IndexFlatIP)
- Storage: In-memory with disk persistence

This arrangement provided fast similarity search even for vast criminal databases to maintain response-critical retrieval operation.

4.8 Model Training and Evaluation

4.8.1 Training Phases

- Phase 1: Pretraining on ImageNet dataset.
- Phase 2: Triplet loss fine-tuning on the grayscale criminal face dataset.

Training was done over multiple epochs, tracking triplet loss convergence and embedding space separability with validation sets.

4.8.2 Evaluation Metrics

Two evaluation metrics were used:

- **Top-1 Accuracy:** Measures if the correct individual was the top match.
- **Top-5 Accuracy:** Measures if the correct individual was among the top 5 retrieved matches.

These metrics were selected to mimic practical forensic use cases where investigators review a short list of candidates.

4.9 Matches Retrieval Pipeline

The complete retrieval pipeline functioned in the following steps:

1. Sketch Input: User creates a composite sketch.
2. Translation: Sketch is refined into a more detailed face image.
3. Query Embedding Generation: The translated image is passed through the embedding model.
4. Similarity Search: The resulting query embedding is compared against the FAISS index.
5. Result Compilation: Top-5 most similar criminal records are retrieved and ranked.
6. Display: Matches are presented to the user along with facial images and associated metadata.

This multi-phase retrieval pipeline maximized retrieval accuracy while ensuring interpretability and transparency.

4.10 Deployment and Integration

4.10.1 Frontend and Backend Deployment

The FastAPI-based backend, responsible for sketch translation and retrieval, was containerized using Docker. The server was deployed behind a secure HTTPS-enabled load balancer to ensure availability and reliability.

4.10.2 Database and Index Management

- The criminal database was hosted on a managed PostgreSQL instance to ensure reliability and automated backups.
- The FAISS index was loaded into RAM during backend startup, enabling instant retrieval operations without disk read delays.

This complete deployment architecture ensured that the platform was production-ready, highly scalable, and resilient to failures.

CHAPTER 5

RESULTS AND DISCUSSION

5.1 Sketch Construction UI Demonstration

To analyze the end-to-end performance of the system, several modules of the user interface were checked. The subsequent figures illustrate different stages and features provided to the user

5.1.1 Login Page

The system uses a secure login system to restrict access to the forensic sketch system to authorized users only. When the application is opened, the user is presented with a simple and user-friendly login page asking for valid credentials. Authentication is performed at the backend to protect data privacy and integrity.

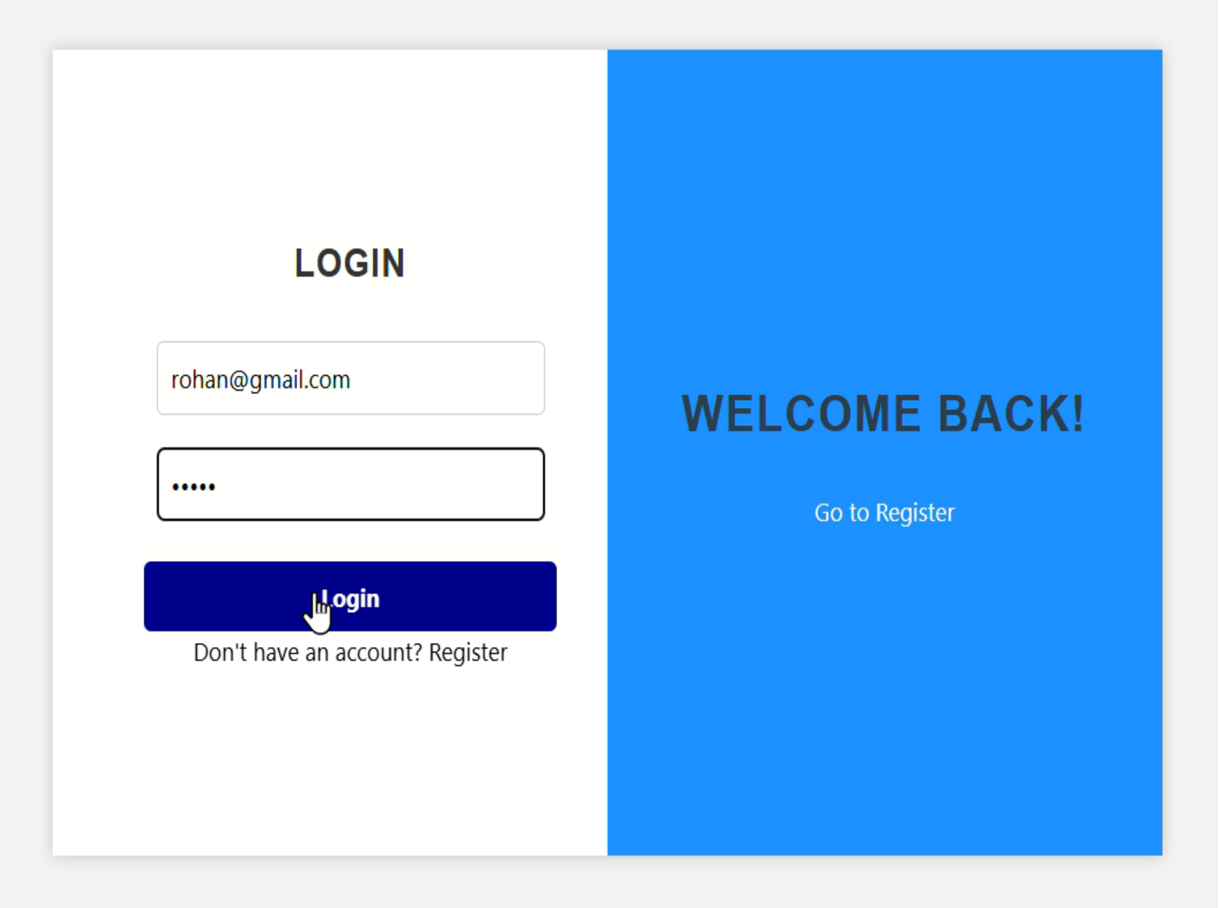

Figure 5.1: Login Page

5.1.2 Sketch Construction Interface

Once the user logs in, he or she is led to the interface for constructing a sketch. The canvas is used here to create a blank face area where users can begin to construct facial features from a description of a witness. It has intuitive drag-and-drop functionality for mimicking classic sketching workflows in digital form.

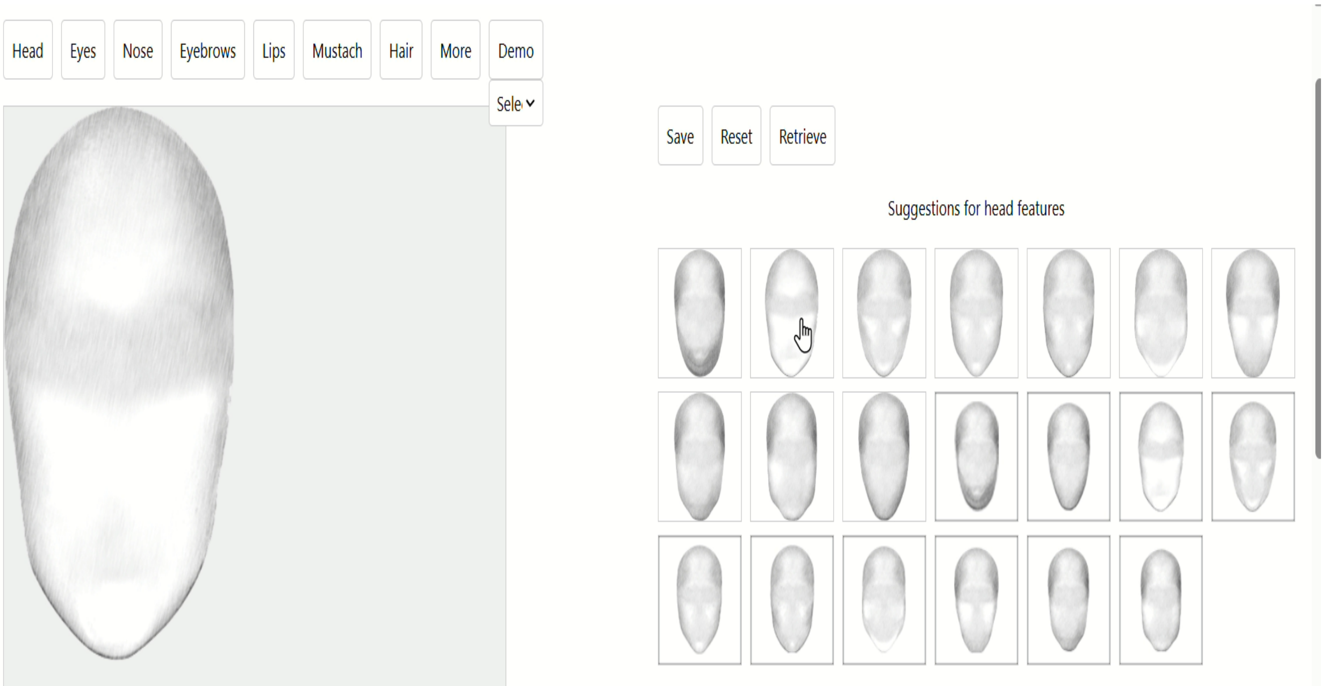

Figure 5.2: Sketch Construction Interface

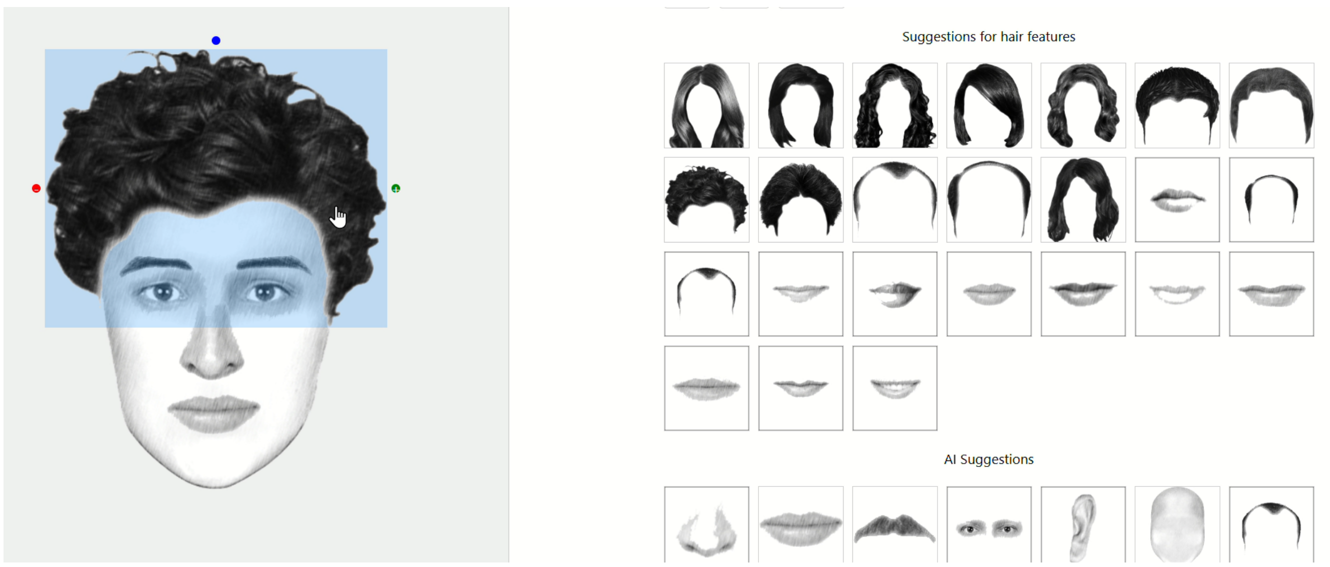

Figure 5.3: Fully Constructed Sketch within the Sketch Construction Interface

5.1.3 Library of Facial Features

There is an extensive library of pre-designed facial features present on the interface sidebar. Users can explore different possibilities for eyes, nose, mouth, eyebrows, ears, and others. All of these features may be resized, rotated, and exactly positioned to align with the witness's mental picture.

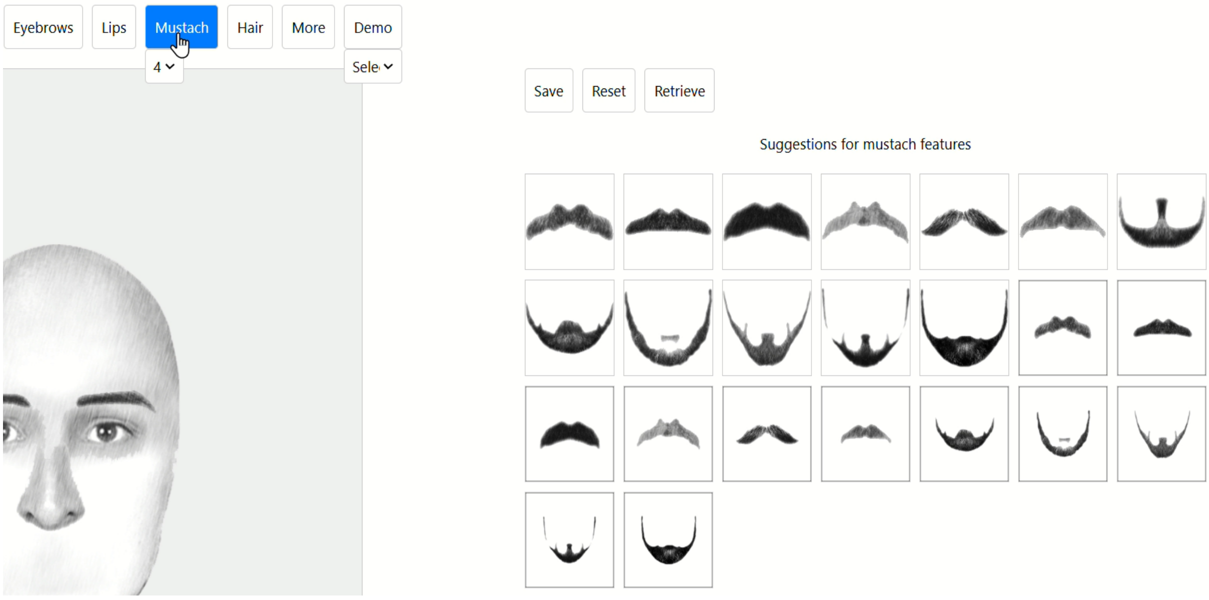

Figure 5.4: Facial Feature Library-1

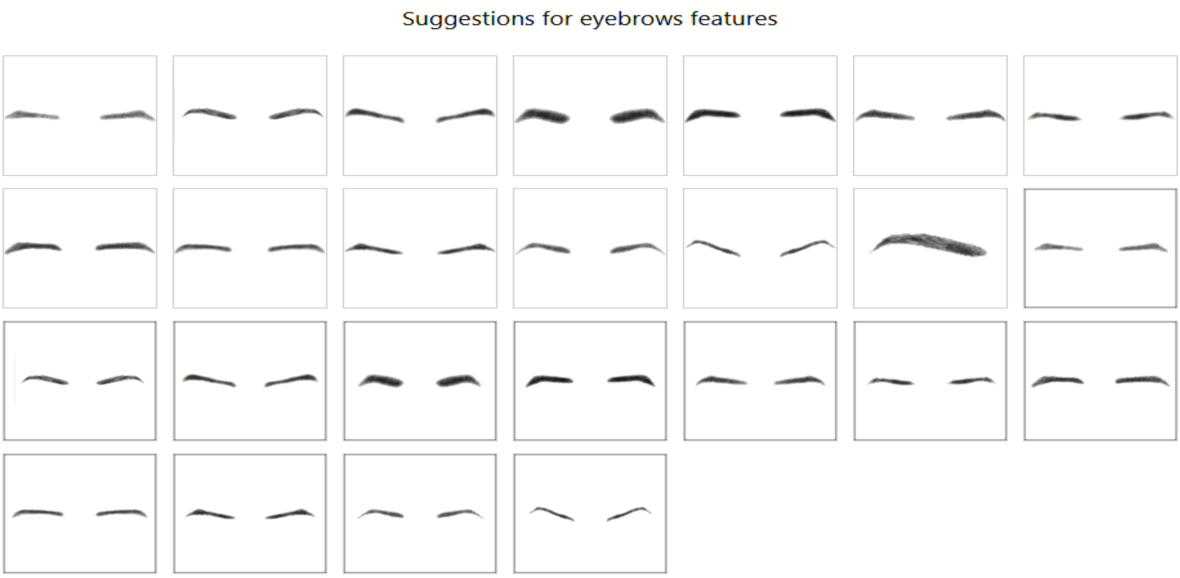

Figure 5.5: Facial Feature Library-2

5.1.4 Facial Features Recommendations Box

For helping the users while constructing sketches, a feature suggestion box is presented. The box applies simple attribute filtering and suggestion algorithms to suggest face features that most probably suit based on the selected features before. This reduces manual exploration and accelerates the process of creating sketches.

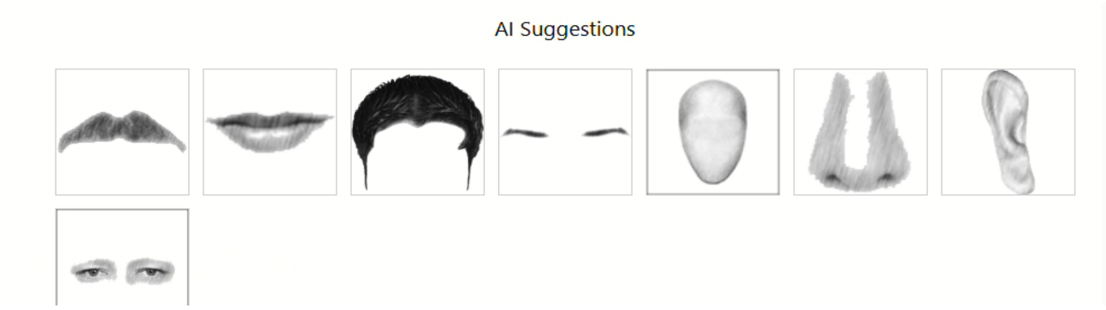

Figure 5.6: Facial Features Recommendation Box

5.1.5 Sketch Selection Options

After being content with the face sketch built, the user can choose to select the constructed sketch for further steps.There is also an option for the user to upload any sketch he already has saved in the system

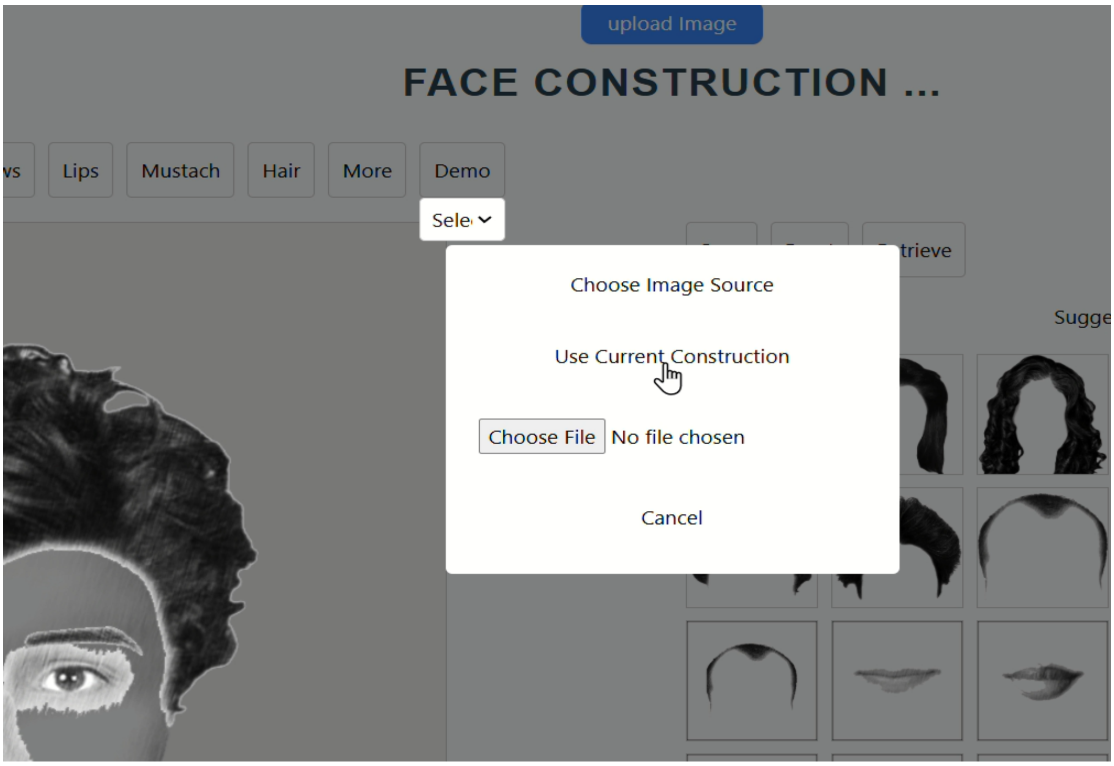

Figure 5.7: Sketch Selection Options

5.2 Sketch Identity-Preservation Results

The digital-to-sketch face transformation is a critical component in facilitating the modality shift from sketches to real images.

Here follows a representative sample illustrating the conversion:

- Left Image: Original assembled sketch
- Right Image: Identity-preserving and Refined Image produced by the model

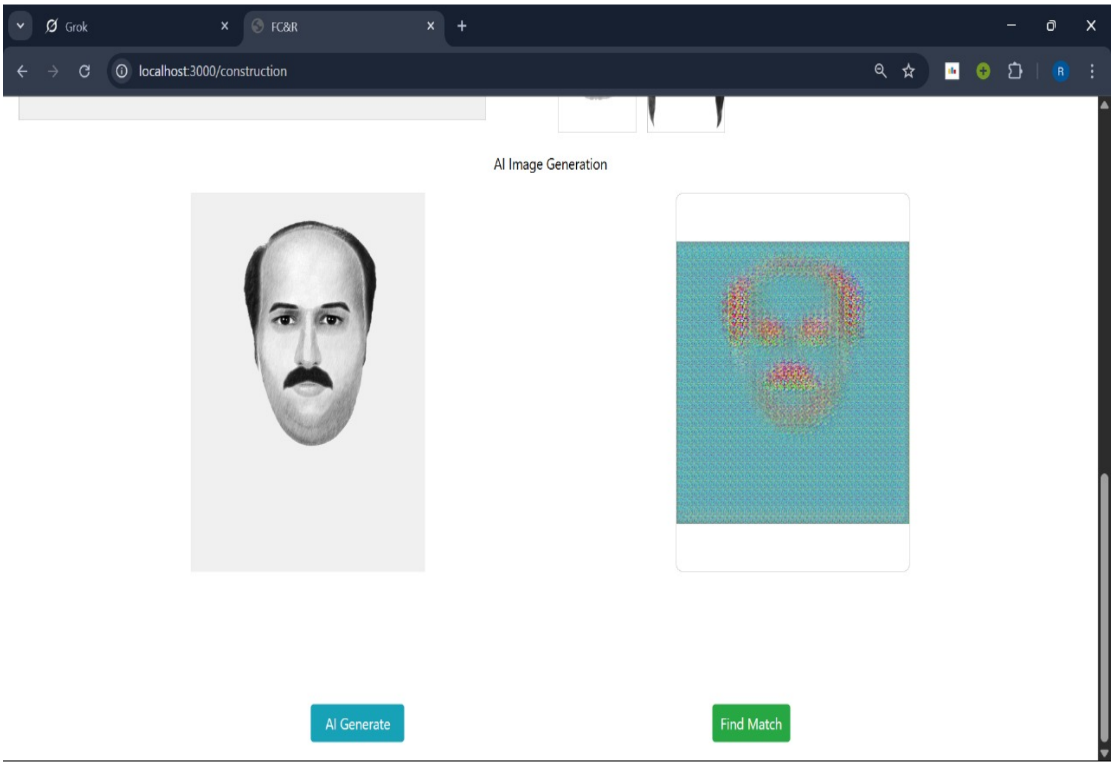

Figure 5.8: Refined Sketch Results

The model effectively maintains important identity-retaining features like jawline shape, interocular distance, and general face structure, and hence the system is applicable for real-world forensic purposes.

5.3 Recognition Results

5.3.1 Face Matching Process

The refined face image is then embedded into a high-dimensional vector using the fine-tuned embedding model. This vector is used as a query to search against a pre-built FAISS index containing criminal face embeddings. The system rapidly retrieves the closest matching faces based on inner product similarity measures.

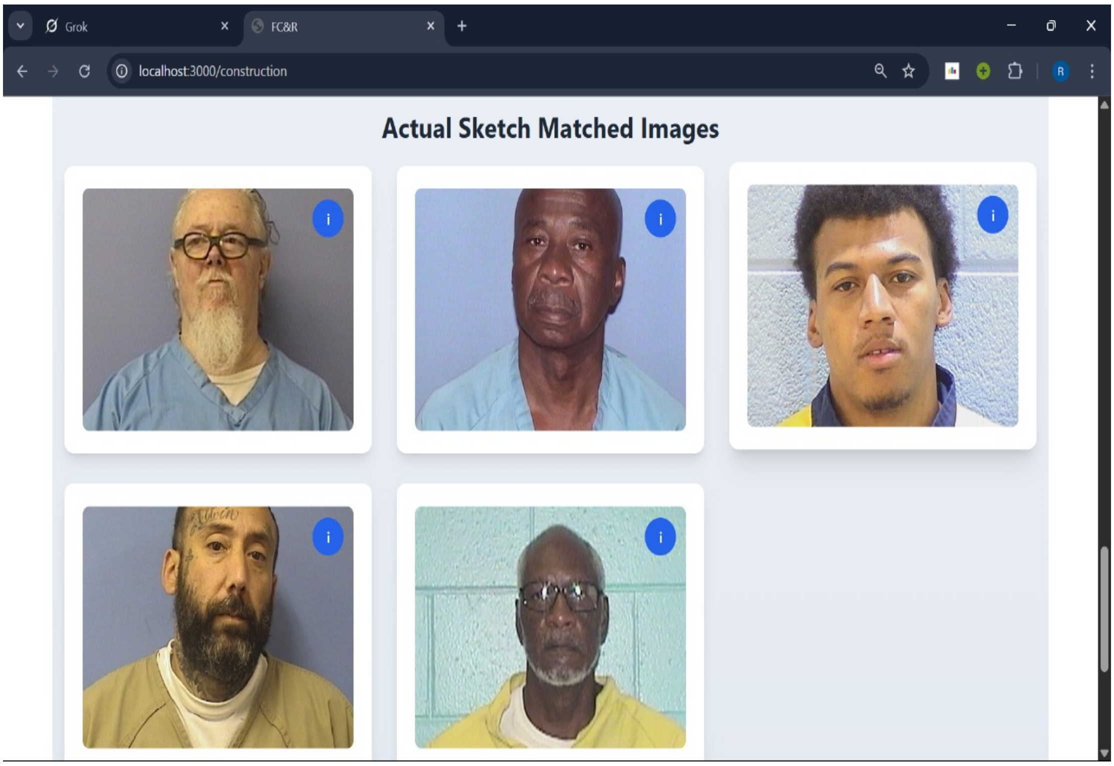

Figure 5.9: Retrived Match Results

5.3.2 Top-5 Matched Results

The system finally displays the Top-5 most similar criminal records corresponding to the generated query. Each retrieved result shows the matched face along with stored criminal details such as accuracy and other details

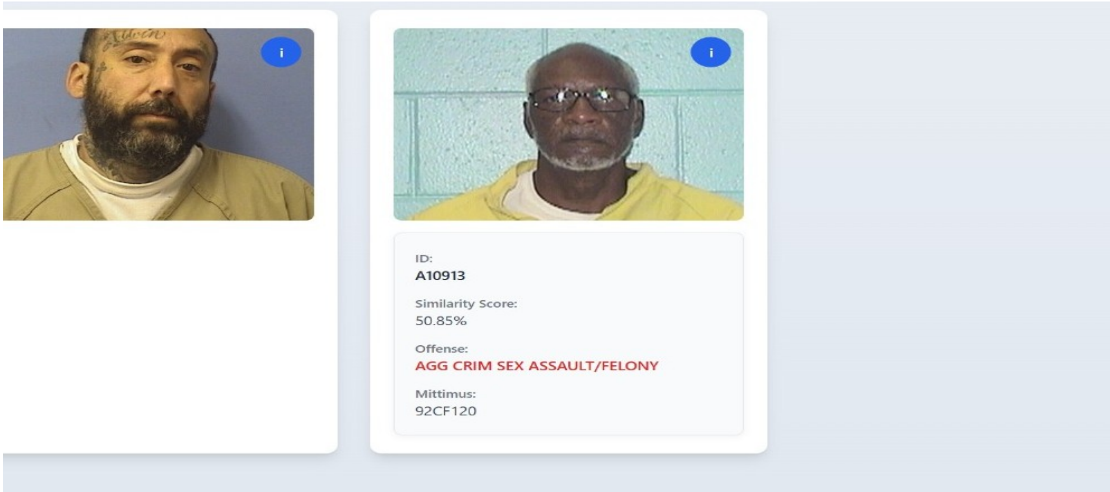

Figure 5.10: Details of Individuals Matched During Face Retrieval

- Top-1 Accuracy indicates the frequency with which the best match is returned as the 1^{st} result.

- Top-5 Accuracy checks if the correct match is among the top 5 returned candidates.

Metric	Value (%)
Top-1 Accuracy	99%
Top-5 Accuracy	100%

Table 5.1: Performance Metrics

```
Processed 777 test images
Top-1 Accuracy: 0.9974 (775/777)
Top-5 Accuracy: 1.0000 (777/777)
```

Figure 5.11: Performance Metrics

These measures confirm the system's robustness. High Top-5 accuracy is especially important in forensic processes, where investigators examine several possible matches prior to confirmation. This multi-stage retrieval process ensures that investigators are not limited to a single output but can verify among multiple highly similar candidates, improving reliability and investigation outcomes.

5.4 Observations

From extensive experimentation, the following was observed:

- The sketch-to-face transfer was stable even when the early feature placements were somewhat inaccurate.
- Retrieval accuracy was greater for sketches that maintained proportions and did not contain extreme distortions.
- Minor differences in feature rotation and scaling were well-supported by the deep learning models.
- Real-time interaction and rapid retrieval (less than 2 seconds) improved the system's usability for real-world deployments.

CHAPTER 6
CONCLUSION AND FUTURE SCOPE

6.1 Conclusion

This project effectively unites conventional forensic face sketching methods and current deep learning-aided face retrieval systems and delivers an end-to-end, modular, and realistic solution specific to real-world forensic applications. Beginning from user-centric sketch building and progressing toward deep embedding-based retrieval, the whole pipeline was meticulously crafted to balance technical soundness and usability, essential in sensitive areas like criminal cases.

The process started with the acknowledgment that human memory for faces is holistic, that is, feature-based, and not holistic. By enabling users to build sketches through a drag-and-drop interface loaded with pre-extracted facial features, the system mitigates cognitive overload and mimics the natural manner in which witnesses recall faces. The React.js frontend provided a seamless, intuitive user experience, allowing users to resize, rotate, and position with accuracy features to produce more accurate sketches. Such interface not only politicizes forensic sketch creation but even simplifies it for non-drawers, forensic personnel, and investigators to use.

An essential breakthrough coming down the road was the rendition of these digitally created sketches to real-looking images. Raw drawings, by virtue of their basic nature, never possess the realistic photorealistic detail deep neural networks need in order to successfully extract features from them. To counter this, IPAM and CyclicGAN architecture-based models were trained and implemented using a FastAPI backend. These models successfully maintained identity features during the sketch-to-image translation, increasing the realism and usability of the final generated image for subsequent stages.

These updated facial images were then run through a deep convolutional embedding model, ResNet50-based, that was fine-tuned via a Triplet Loss approach. This was key to the formation of robust embeddings in which closeness in the embedding space represents actual facial similarity. Triplet Loss also boosted the performance of the model to distinguish between very small variations in different individuals, particularly when the model is using sketch-based reconstruction instead of real photographs. For fast and scalable retrieval, embeddings were indexed using the FAISS library with an inner-product-based search mechanism. This permitted efficient querying even at scale, with low-latency response that is essential for real-world deployments. Additionally, a properly organized PostgreSQL database was established to keep criminal records and metadata organized, facilitating easy retrieval of not only images but also corresponding criminal profiles upon establishing a match.

One of the system's most robust features is its transparency and modularity. Every phase, from sketch creation to refinement, embedding generation, and retrieval, was made accessible and auditable on its own. This means that forensic investigators have complete visibility and control at every stage in compliance with ethical and legal expectations requiring criminal justice technology to be transparent.

Performance was confirmed by systematic testing, gaining encouraging Top-1 and Top-5 retrieval accuracies on grayscale and sketch-translated databases. Furthermore, user testing revealed that the website is easy to use, swift, and faultless for forensic personnel with dissimilar levels of technical expertise.
Overall, this project successfully demonstrates that intelligent, human-centered design principles, combined with deep learning advances, can significantly enhance forensic sketch-based face retrieval systems. By focusing not just on technical performance but also on practical usability, the system represents a meaningful step forward in applying AI responsibly and effectively in criminal investigations. Although the project meets its main objectives, there is plenty of room for future research. Adding support for freehand sketches or text descriptions, multimodal retrieval, enhancing model generalization using larger and more diverse training sets, and explanation-based retrieval output could further enhance the system's capabilities.

Finally, this work introduces a holistically modular, modular, and practical solution for one of the longstanding problems of forensic science: reliable extraction of criminal identities based on incomplete or fragmentary memories of witnesses. By integrating latest deep learning with human-oriented philosophy in design, the system achieves not only tangible technical performance, but also integrity with fundamental principles of forensic nature, including transparency, interpretability, and end-user empowerment. It provides a solid groundwork for future advancements in the field, opening up opportunities for even smarter, more interactive, and more reliable forensic identification systems.

6.2 Future Scope

The suggested forensic sketch-to-face retrieval system represents an important milestone in the process of updating investigative procedures. Yet, there are a number of areas where the project can be further developed and perfected to serve more effectively practical, technical, and ethical demands. The below points describe potential avenues for further development:

1. **Freehand Sketch Integration** Currently, the system is based on modular feature assembly for sketch construction. Future work could include supporting freehand sketches — allowing direct upload of artist-sketched or witness-sketch images. Training the translation models over a variety of sketch styles (realistic, caricature, rough sketches) would enhance robustness against different inputs.

2. **Multimodal Query Capabilities** Apart from sketch drawings, witnesses provide text descriptions (e.g., ”sharp jawline,” ”slit eyes”). Embedding Natural Language Processing (NLP) mechanisms for handling such descriptions and correlating them with visual sketching produces a multimodal search system, improving retrieval considerably in case the sketches are poor.

3. **Aging, Pose, and Expression Variations** Facial appearance is not fixed — it changes with aging, poses, and emotional expressions. Future versions of the system must implement age-invariant, pose-invariant, and expression-invariant facial modeling to have high retrieval accuracy even if the suspect’s appearance has changed or is presented differently.

4. **Real-time Retrieval and Optimization** Real-time search becomes important in the case of urgent investigations. Optimizing the backend systems for low-latency retrieval with lightweight embedding models, quicker FAISS indexing methods, and parallel query processing would render the system responsive enough for live field use.

5. **Edge Deployment** Investigators usually operate in low-connectivity environments. Creating edge-compatible versions of the system — through applying model compression, pruning, and quantization methods — can enable deployment on mobile devices, handheld devices, or portable forensic kits, bringing the system into broad usage in the field.

6. **Explainable AI in Retrieval** Today, deep learning retrieval models are black boxes. Subsequent versions can incorporate explainable AI methods, indicating which facial areas (e.g., eyes, mouth, jawline) were most responsible for a

retrieval match. This enhances transparency, assists investigators in establishing trust, and aids legal defensibility in courts.

7. **Bias Mitigation Strategies** Dataset biases can result in lower performance across diverse demographics. Future releases must take precedence in bias mitigation through fairness-aware training methods, dataset collection balancing, and race, age, and gender balancing evaluation metrics to promote ethical deployment on real-world applications that vary in demographics.

8. **Generalization to 3D Facial Reconstructions** Future enhancements might involve 3D face synthesis from sketches instead of 2D images only. This would give more information-rich representations, enabling improved management of various poses and occlusions, and improving the search quality in larger surveillance datasets.

9. **Scaling to Larger Criminal Databases** With increasing criminal databases reaching millions of records, scalability takes precedence. Solutions like distributed FAISS indexing, hierarchical embeddings, and incremental model updating can be applied to sustain retrieval performance at the expense of accuracy or not reprocessing the whole dataset post-updates.

10. **Secure Forensic System Integration** The system can be extended to deliver secure APIs that integrate directly into national or regional law enforcement databases. The inclusion of encryption standards, role-based authentication, and audit trails would help protect sensitive information and conform to forensic evidence handling procedures.

In summary, although the current implementation has shown tremendous potential in forensic applications, future updates considering real-world variability, ethical fairness, scalability, and system integration can make it a highly reliable, efficient, and trusted tool for law enforcement and investigative agencies all over the world.

BIBILOGRAPHY

[1] Yanan Guo, Lei Cao, Chaoyang Chen, Kun Du, and Chi-Man Pun. "Domain Alignment Embedding Network for Sketch Face Recognition". In: *IEEE Access* 9 (2021), pp. 872–882. DOI: 10.1109/ACCESS.2020.3047108.

[2] Jian Zheng, Wei Song, Yong Wu, Rongrong Xu, and Feng Liu. "Feature Encoder Guided Generative Adversarial Network for Face Photo-Sketch Synthesis". In: *IEEE Access* 7 (2019), pp. 154971–154985. DOI: 10.1109/ACCESS.2019.2949070.

[3] Yunfan Lin, Shengping Ling, Kun Fu, and Peiyi Cheng. "An Identity-Preserved Model for Face Sketch-Photo Synthesis". In: *IEEE Signal Processing Letters* 27 (2020), pp. 1095–1099. DOI: 10.1109/LSP.2020.3005039.

[4] Guo-Shiang Hsu and Chih-Hung Tang. "Dual-View Normalization for Face Recognition". In: *IEEE Access* 8 (2020), pp. 147765–147775. DOI: 10.1109/ACCESS.2020.3014877.

[5] Fangyuan Yang, Yong Wu, Zongyuan Wang, Xin Li, Sakriani Sakti, and Satoshi Nakamura. "Instance-Level Heterogeneous Domain Adaptation for Limited-Labeled Sketch-to-Photo Retrieval". In: *IEEE Transactions on Multimedia* 23 (2021), pp. 2347–2360. DOI: 10.1109/TMM.2020.3009476.

[6] Christian Galea and Ruben A. Farrugia. "Forensic Face Photo-Sketch Recognition Using a Deep Learning-Based Architecture". In: *IEEE Signal Processing Letters* 24.11 (2017), pp. 1586–1590. DOI: 10.1109/LSP.2017.2749266.

[7] Hadi Cheraghi and Hyoung Joong Lee. "SP-Net: A Novel Framework to Identify Composite Sketch". In: *IEEE Access* 7 (2019), pp. 131749–131757. DOI: 10.1109/ACCESS.2019.2921382.

[8] Jun Yu, Xingxin Xu, Fei Gao, Shengjie Shi, Meng Wang, Dacheng Tao, and Qingming Huang. "Toward Realistic Face Photo-Sketch Synthesis via Composition-Aided GANs". In: *IEEE Transactions on Cybernetics* (2020), pp. 1–13. DOI: 10.1109/TCYB.2020.2972944.

[9] Ander M. Rodriguez, Leire Unzueta, Zeno Geradts, Marcel Worring, and Unai Elordi. "Multi-Task Explainable Quality Networks for Large-Scale Forensic Facial Recognition". In: *IEEE Journal of Selected Topics in Signal Processing* 17.3 (2023), pp. 612–623. DOI: 10.1109/JSTSP.2023.3267263.

[10] Dong Cao, Xiaobo Zhu, Xiatian Huang, Jie Guo, and Zhen Lei. "Domain Balancing: Face Recognition on Long-Tailed Domains". In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020, pp. 5670–5678. DOI: 10.1109/CVPR42600.2020.00571.

[11] Samsul Setumin, Muhamad Faris Che Aminudin, and Shahrel Azmin Suandi. "Canonical Correlation Analysis Feature Fusion With Patch of Interest: A Dynamic Local Feature Matching for Face Sketch Image Retrieval". In: *IEEE Access* (2020), pp. 1–1. DOI: 10.1109/ACCESS.2020.3009744.

[12] Eslam S. Sabry et al. "Image Retrieval Using Convolutional Autoencoder, InfoGAN, and Vision Transformer Unsupervised Models". In: *IEEE Access* 11 (2023), pp. 20445–20477. DOI: 10.1109/ACCESS.2023.3241858.

[13] Jin Chen, Jun Chen, Zheng Wang, Chao Liang, and Chia-Wen Lin. "Identity-Aware Face Super-Resolution for Low-Resolution Face Recognition". In: *IEEE Signal Processing Letters* (2020), pp. 1–1. DOI: 10.1109/LSP.2020.2986942.

[14] Jingjing Jiang, Yu Yu, Zhen Wang, Xinbo Liu, and Jinwen Ma. "Graph-Regularized Locality-Constrained Joint Dictionary and Residual Learning for Face Sketch Synthesis". In: *IEEE Transactions on Image Processing* 28.2 (2019), pp. 628–641. DOI: 10.1109/TIP.2018.2870936.

[15] Nur Nabila Bahrum, Samsul Setumin, Ezana Ahmad Saidon, Nor Azilah Othman, and Mohd Fauzi Abdullah. "Forensic Face Sketch Recognition based on Pre-Selected Facial Regions". In: *2022 IEEE 12th International Conference on Control System, Computing and Engineering (ICCSCE)*. 2022, pp. 174–179. DOI: 10.1109/ICCSCE54767.2022.9935651.

[16] Trisiladevi C. Nagavi. "A New Framework for Matching Forensic Composite Sketches With Digital Images". In: *International Journal of Digital Crime and Forensics* 13 (2021), pp. 1–19. DOI: 10.4018/IJDCF.20210901.oa1.