Forensic Face Sketch Construction and Intelligent Recognition System

A Report on Major Project Part 1
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
NOV 2024

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 **Forensic Face Sketch Construction and Intelligent Recognition System** submitted by us to the **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 **Forensic Face Sketch Construction and Intelligent Recognition System** 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

Dr. K. Radhika — Head of the Department

Submitted for Semester Major-Project Part 1 viva-voce examination held on _______________

Examiner-1

Examiner-2

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.

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

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Facial recognition is a vital tool in forensic science, aiding law enforcement in identifying suspects based on eyewitness descriptions. Traditionally, forensic artists create hand-drawn sketches that are compared with photos in criminal databases. However, this process can be slow, subjective, and imprecise, particularly with non-frontal images or sketches that combine features from multiple sources. These limitations can hinder criminal investigations. To overcome these challenges, we propose an advanced dual-module system to optimize both sketch creation and recognition. The **Face Sketch Construction** module assists forensic artists in creating sketches, offering a drag-and-drop interface that allows users to select facial features from available options, enhancing precision and speed while reducing subjective bias. The module also uses an autoencoder with a recurrent neural network (RNN) to suggest features based on those already chosen, creating a more accurate and efficient translation of witness descriptions. The **Face Recognition** module improves sketch matching by combining FEGGAN, which generates high-fidelity sketches, with IPAM to ensure identity consistency. Once digitized, the ArcFace model, integrated with ResNet, performs feature embedding to enhance matching accuracy against criminal databases. To secure sensitive data, the platform includes machine locking, two-step verification, and centralized control. By modernizing forensic sketching and recognition, this system offers a faster, more accurate, and user-friendly tool for criminal investigations.

Keywords : Face Sketch Construction, Face Recognition, Autoencoder, Recurrent Neural Networks, General Adversarial Models, FEGGAN, IPAM, ArcFace, ResNet, Two-Step Verification, Machine-Locking

CHAPTER 1
INTRODUCTION

1.1 Overview

Facial recognition plays a pivotal role in modern forensic science, aiding law enforcement agencies in identifying suspects based on eyewitness accounts. Traditional methods rely on forensic artists to create hand-drawn sketches, which are then matched against criminal databases. While this technique has been used successfully for decades, it suffers from significant limitations in terms of accuracy, speed, and subjectivity. To address these challenges, this project proposes a comprehensive, dual-module system that integrates advanced machine learning technologies to streamline and enhance the process of facial sketch creation and recognition.

1.2 Motivation

The increasing complexity of criminal investigations necessitates more efficient and accurate tools for suspect identification. Eyewitnesses often struggle to recall detailed facial features, and forensic artists must rely heavily on subjective interpretation, leading to potential inaccuracies. Moreover, the time-consuming nature of creating detailed hand-drawn sketches can delay investigations. There is a clear need for an advanced system that can assist forensic artists and improve the overall reliability and speed of the sketch-to-recognition pipeline.

1.2.1 Drawbacks of Traditional Systems

Traditional facial recognition systems face several challenges:

1. **Subjectivity and Variability:** Hand-drawn sketches are highly dependent on the skill and interpretation of the artist, which can lead to inconsistencies.

2. **Time-Consuming:** Creating detailed sketches based on witness descriptions is a slow process, potentially delaying crucial investigation stages.

3. **Limited Accuracy:** Matching hand-drawn sketches with photographs in criminal databases often yields limited accuracy, especially when dealing with non-frontal images or composite features from multiple sources.

4. **Difficulty with Composite Features:** When sketches combine features described by multiple witnesses or different sources, traditional methods struggle to maintain fidelity and consistency.

1.2.2 Need for Better Systems

Traditional forensic sketching methods face several limitations that hinder their effectiveness in modern criminal investigations. These include:

1. **Complexity of Investigations:** Traditional methods struggle with non-frontal or low-quality images and conflicting witness descriptions, leading to inaccuracies and delays in identifying suspects.

2. **Time Sensitivity:** The time-consuming process of creating hand-drawn sketches delays investigations, which can be critical when rapid suspect identification is needed.

3. **Lack of Precision:** Hand-drawn sketches often lack the accuracy required to effectively match suspects with images in databases, especially when sketches combine features from multiple witnesses or sources.

4. **Scalability and Accessibility:** Existing systems are often inefficient in handling large databases, making it difficult for law enforcement to quickly access and compare relevant information across jurisdictions.

A more advanced system incorporating machine learning, such as GANs and IPAMs, can address these issues by automating and improving sketch generation, enhancing recognition accuracy, and speeding up the entire process, ultimately making forensic facial recognition more efficient and reliable.

CHAPTER 2

LITERATURE SURVEY

Forensic face sketch recognition by means of machine learning has progressed quickly, in part because of deep learning models that are equipped to handle the specific ways in which to compare sketches and photos. This section of the report reviews important models used previously, by grouping the tasks and dissecting each approach's solution to those problems, highlighting the reliability and efficiency of these methods where they exist.

2.1 BRIDGING THE DOMAIN GAP BETWEEN SKETCHES

The domain gap between sketches and photos is one of the main difficulties in forensic face recognition, and it's that sketches usually don't capture the fine details of real images, and that they are stylistically different. Most deep learning models overcome this lack of overlap by mapping sketches into some space closer to photos.

2.1.1 Domain Alignment Embedding Networks [1]

Another way to reduce the domain gap is the Domain Alignment Embedding Network. By projecting sketch and photo features into a common domain space, this technique enables more robust matching. Usually paired with CNN models, domain alignment has boosted accuracy on multiple databases, by synthesising the feature representations of sketches and photos.

2.1.2 Generative Adversarial Networks (GANs) [2]

Especially cyclic GANs, which have been very successful at changing sketches to look like photos. One example is the cyclic GAN architecture, which involves two generators and two discriminators that collaborate to impose cycle consistency. It preserves identity while mapping sketch features into a photorealistic space, with strong results on multiple datasets.

Ref. No.	Keywords	Methods and Datasets Used	Evaluation Metrics	Future Work
[1]	Sketch face recognition, feature embedding network, deep metric learning, small sample problem.	**S1:**UoM-SGFSv2-set A, **S2:**UoM-SGFSv2-set B, **S3:**PRIP-VSGC; **Method Used:** Domain Alignment Embedding Network	**Rank-1:** S1-68.53%, S2-74%; **Rank-10:** S1-92.4%, S2-95.2%, S3-63.2%; **Rank-50:** S1-97.47%, S2-99.07%	Enhance model robustness by incorporating various artistic and cartoon sketch styles; Improve recognition accuracy across different ethnicities, age groups and genders; Optimize DAEN for deployment in real-time applications, like surveillance systems and mobile apps.
[2]	Face photo-sketch, synthesis, image-to-image translation, generative adversarial network, encoder.	**CUFS-** Chinese University of Hong Kong Face Sketch Database, **CUFSF-** CUHK Face Sketch FERET Database; **Method used:** A cyclic GAN framework which consists of two generators and discriminators	**Face Recognition Accuracy:** CUFS-95.34, CUFSF- 100	Improve detail by gradually increasing resolution; Balance realism and identity retention dynamically; Adapt EGGAN for varied sketch styles and conditions,Train on diverse datasets for unbiased synthesis, Enhance efficiency for real-time forensic applications.

Table 2.1: Papers related to Bridging the Domain Gap between sketches and photos

2.2 IDENTITY PRESERVATION ACROSS TRANSFORMATIONS

Preservation of facial features in sketch morphing is essential for forensic accuracy. Those models that maintain identity information across domains have been more successful at producing recognisable faces, a need in criminal investigations.

Ref. No.	Keywords	Methods and Datasets Used	Evaluation Metrics	Future Work
[3]	Face sketch-photo synthesis, generative adversarial networks, image generation, image-to-image translation.	**CUFS**-Chinese University of Hong Kong Face Sketch Database, **CUFSF**-CUHK Face Sketch FERET; **Method Used:** A U-Net generator with two discriminators, optimized by leaky ReLU and Adam, alongside identity-verifying models (ResNet-50 and LightCNN-29v2)	**Rank-1:** CUFS(ResNet-50)-0.8, CUFS(LightCNN-29v2)-0.78, CUFSF(ResNet-50)-0.737, CUFSF(LightCNN-29v2)-0.701	Adapt model to various sketch types, Broaden demographic coverage and inclusivity, Enable instant face matching from sketches, Combine with infrared or thermal imaging, Create data for underrepresented populations.
[4]	Low-resolution face recognition, face super- resolution, identity-aware learning, magnitude loss.	LFW, CelebA Datasets; **Method Used:** Feature decoupling learning to separate identity-related (angle), Quality-related (magnitude) components, LightCNN_v9 for feature extraction, Cosine similarity evaluation for down-scaled faces, and Identity-aware super-resolution network architecture to enhance identity-related details in low-resolution faces.	**Accuracy(LFW):** LightCNN-v9-98.46%, LightCNN-v29-98.98%, **Rank-1(CelebA):** LightCNN-v9-30.39%, LightCNN-v29-41.36%	Exploring adaptive resolution scaling techniques

Table 2.2: Papers related to Identity Preservation across Transformations

2.2.1 Identity-Preserving Attribute Models (IPAMs) [3]

Identity-preserving modules, for example those in GAN networks, are incorporated to keep important facial features stable amid transformations. For instance, IPAMs paired with cyclic GANs have reached recognition rates of more than 85 per cent on difficult datasets, demonstrating that retaining distinctive identity characteristics greatly helps recognition performance.

2.2.2 Decoupling Learning with LightCNN [4]

Feature decoupling learning extracts identity-related features from the rest of the face, so that transformations do not change important identity features. LightCNN coupled with ArcFace as a feature extractor uses decoupled features, which maintain over 90 per cent accuracy in controlled studies by minimizing variance introduced by non-identity factors, such as lighting or slight facial characteristics.

2.3 CROSS-DOMAIN FEATURE ALIGNMENT

In forensic sketch recognition, a key challenge is the variability in sketch style and quality, largely influenced by the artist's skill and the witness's memory accuracy. These factors introduce significant inconsistencies in sketches, which can lead to mismatches when compared to standard photo databases. Recent advancements have aimed to standardize or normalize these variations to make sketch-based recognition more effective, particularly in cross-domain applications where photos and sketches differ greatly in visual features.

2.3.1 Dual-View Normalization with Light-CNN [5]

This method addresses, in particular, the stylistic and qualitative variation in sketches produced by different artists. By making these variations into the norm, the technique makes facial feature representations more robust, so that cross-view recognition is more consistent and reliable.

2.3.2 Graph-Regularized Locality-Constrained Joint Dictionary Learning [6]

For low-resolution forensic sketches, preserving essential identity cues becomes challenging due to limited facial details. Graph-Regularized Locality-Constrained Joint Dictionary Learning addresses this by combining graph regularization with residual learning to capture subtle textural and structural features. This approach enhances the visibility of identity cues that would otherwise be lost with traditional methods, preserving key characteristics such as eyebrow arch or jawline structure. By refining these subtle details, this method ensures that vital identity information is retained, improving recognition accuracy for lower-quality images typically found in forensic settings.

Ref. No.	Keywords	Methods and Datasets Used	Evaluation Metrics	Future Work
[5]	Face recognition, face normalization, face synthesis.	Multi-PIE, CASIA-WebFace, IJB-A, IJB-C Datasets; **Method Used:** Dual-View Normalization with Light-CNN and ArcFace as Feature Extractors	**VR@FAR=0.1:** IJB-A(Light-CNN): 95.7, IJB-A (ArcFace):97.2, IJB-C(Light-CNN):92.4, IJB-C(ArcFace):96.5 **Rank-5(IJB-A):** Light-CNN:98.7, ArcFace:98.8	Use unsupervised clustering to tailor normalization, Incorporate more views beyond dual poses, Improve generalization using unlabeled data.
[6]	Face sketch synthesis, residual learning, joint dictionary learning, local geometry structure, rare characteristics	CUHK, AR, XM2VTS Datasets; **Method Used:** Graph-regularized locality-constrained joint dictionary learning algorithm and residual learning algorithm	**SSIM:** ***1.CUHK:***RL-0.6104,Dict+RL-0.6228; ***2.AR:***RL-0.6212,Dict+RL-0.6256; ***3.XM2VTS:***RL-0.4908,Dict+RL-0.5089	Apply a parsing model to differentiate between facial and non-facial regions, Utilize distinct prediction models for improved synthesis accuracy, Implement a luminance recapping algorithm to standardize lighting conditions.

Table 2.3: Papers related to Cross Domain Feature Alignment

2.4 ADAPTING TO COMPLEX REAL-WORLD SCENARIOS

For forensic face recognition systems to be effective in real-world applications, they must handle complex scenarios like inconsistent lighting, occlusions (e.g., hats, glasses), and partial facial visibility. Hybrid models, which combine multiple types of neural networks or integrate machine learning with deep learning techniques, have shown strong potential for capturing facial details across varied conditions. These models, together with multi-feature fusion techniques, allow systems to integrate data from different facial features—such as texture, shape, and color—creating a flexible and comprehensive representation of each face. This approach enables forensic systems to compensate for missing or obscured features and adjust to lighting variations, maintaining recognition accuracy even in challenging situations.

Ref. No.	Keywords	Methods and Datasets Used	Evaluation Metrics	Future Work
[7]	Forensic, Face sketch, recognition, histogram of gradient, facial region	PRIP-HDC; **Method Used:** HOG with utilizing a pre-selected facial region of both sketches and the photos using Adobe Photoshop CS6 to remove the occlusion on the sketch dataset and photo dataset.	**Accuracy @Rank-1:** 6.38%(Existing- 2.13	Use deep learning models (e.g., GANs or U-Nets) to automatically detect occlusions and inpaint missing facial areas, providing cleaner inputs for recognition models, thus imporving accuracy
[8]	Domain adaptation, cross-modal image retrieval, sketch, person re-identification	UT-Zap50K, CelebFaces, Market-1501; **Method Used:** Instance-level Heterogeneous Domain Adaptation (IHDA) framework	**Sketch-to-photo retrieval: Rank-1:** UT-Zap50K- 68.7, CelebFaces- 95.7, Market-1501- 85.6	Explore effective methods to leverage rich-labeled datasets for limited-labeled tasks, Investigate the IHDA framework's applicability in scenarios beyond instance-level retrieval tasks, Develop strategies to identify shared attributes for category-level retrieval tasks across datasets like Sketchy, PACS, and M3SDA.

Table 2.4: Papers related to Adapting to Complex Real-world Scenarios

2.4.1 Histogram of Oriented Gradient and Gabor Wavelet Fusion with Canonical Correlation Analysis [7]

Combining HOG and GW features using CCA, matches are refines under different lighting conditions, as well as occlusions. Also, Patches of Interest (PoI) are employed to deal with partial faces, increasing accuracy by working on the most recognizable areas. This multi-feature fusion model performs at a level of 75-80 % accuracy

2.4.2 Instance-Level Heterogeneous Domain Adaptation (IHDA) [8]

IHDA models carry out instance-level domain adaptation to correct for variations in the quality of the sketch and photo at the instance level, so that the matches are more accurate in difficult conditions. With reported accuracies reaching as high as 80 per cent in cross-domain tasks, IHDA is highly effective at normalising away variations and bringing instances into alignment, which makes it particularly useful in large forensic databases

2.5 CNN-BASED TECHNIQUES

CNNs are quite effective at dealing with the subtle changes found in forensic sketches, which include age, style, and resolution. By leveraging deep hierarchical layers, CNNs can capture complex facial patterns and adapt to variations that are typical in forensic scenarios, such as different sketching techniques or artist-specific styles. Furthermore, advanced CNN architectures can incorporate attention mechanisms, enabling the model to focus on distinctive facial regions, thereby enhancing accuracy even when sketches lack detail or have partial occlusions.

2.5.1 A Fresh Approach to Matching Forensic Composite Sketches to Digital Photographs [9]

This method uses a six-layer convolutional neural network (CNN) with Swish activation functions to achieve higher recognition accuracy in sketches of different quality and age. On the E-PRIP and Composite Sketch with Age Variation (CSA) datasets, the model reached accuracies of 78.26 per cent and 69.57 per cent respectively. The approach emphasises the value of shallow CNNs and dataset augmentation in enhancing the ability to match sketches and digital images in forensic applications.

2.5.2 Forensic Face Photo-Sketch Recognition via a Deep Learning Architecture [10]

This technique uses a 3D Morphable model to capture facial feature changes by using a deep convolutional neural network (DCNN) with transfer learning. By combining synthetic with original sketch distance measures, the model improves identity matching. Evaluated on CUFS and Color FERET, this method shows good performance in photo-sketch matching, achieving improvements in mean rank retrieval with the best results in terms of mean rank, suggesting it as a potential approach for forensic photo-sketch recognition.

Ref. No.	Keywords	Methods and Datasets Used	Evaluation Metrics	Future Work
[9]	Convolution Neural Network (CNN), Dropout, E-PRIP Dataset, Exponential Linear Unit, Face Sketch Recognition, Leaky Rectified Linear Unit, Sigmoid, Swish Activation Function	E-PRIP, IIIT Delhi Composie Sketch with Age variation dataset; **Method Used:** Swish Activation Function with 6-layer CNN	**Accuracy:** E-PRIP - 78.26%, CSA - 69.57	Test multiple compact CNN architectures to improve efficiency, Merge additional datasets to increase image count by enhancing model performance, Increase dataset size to improve deep network performance, Explore using more sophisticated neural network models for training on larger datasets to enhance sketch and digital image matching accuracy.
[10]	Augmentation, convolutional neural network, deep learning, fusion, hand-drawn sketch, morphological model	CUFS, XM2VTS, AR, CUHK, CUFSF, Color FERET, PRIP-HDC(Testing); **Method Used: DEEPS:** uses a 3-D Morphable model to generate variations of facial features and attributes for face images; **DEEPS-M:** enhances identity determination by fusing distance measures from synthetic and original sketches for improved photo matching.	**Mean Rank Retrieval Rate:** DEEPS:325.02, DEEPS-M:312.11	Developing a more flexible 3-D Morphable model for better facial feature variation, Applying proposed methods to additional human face recognition tasks.

Table 2.5: Papers related to CNN-based Techniques

2.6 Comparative Analysis

In this section, we present a comparative analysis of the models discussed in the literature review, categorizing them based on the primary challenges they address in forensic face sketch recognition. This comparison in Table-7 offers clues about the strengths, weaknesses and performance of each model, revealing trade-offs between accuracy, computational cost and applicability in real-world forensic settings.

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	~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	~85%
	Feature Decoupling Learning + LightCNN	Separation of identity and non-identity features	Effective separation for better identification	Reduced performance with lower-resolution sketches	90%+
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 and 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%+

Figure 2.1: Comparative Analysis of Forensic Face Sketch Recognition Models Based on Primary Challenges

2.6.1 Summary of Comparative Analysis

This comparison highlights the advantages and disadvantages of each category of model:

• GAN-style model are very good at domain gap reduction (accuracy up to 85%) need a lot of computation.

• Identity-preserving methods, for example, feature decoupling with LightCNN also 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 per cent 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.

2.7 Limitations found in previous works

- **High Computational and Data Requirements:** While effective in domain gap reduction and specialize in generalizing across diverse datasets or sketch styles, Cyclic GAN Frameworks need large datasets and substantial computational resources, making them impractical for real-time or data-limited forensic settings.
- **Sensitivity to Sketch Quality, Style, and Resolution:** U-Net with Dual Discriminators and Dual-View Normalization with ArcFace models adapt to some sketch quality and style variations, but they perform poorly with extreme stylistic changes or low-resolution sketches, limiting their robustness.
- **Limited Generalization Across Datasets:** Identity-Preserving GANs (IPAM) and Feature Decoupling Learning + LightCNN models retain identity traits but struggle to generalize across diverse datasets or sketch styles, reducing their effectiveness in real-world applications with varied data.
- **Performance and Efficiency Trade-offs:** Feature Decoupling Learning + LightCNN and Dual-View Normalization with LightCNN models achieve high accuracy by focusing on identity features but at the cost of computational efficiency, resulting in slower performance and high training costs.
- **Inconsistent Performance in Real-World Conditions:** HOG + Gabor Wavelet Fusion with CCA and Instance-Level Heterogeneous Domain Adaptation models handle occlusions and complex environments but show reduced performance in low-resolution cases and require substantial training data, limiting their adaptability to real-world forensic cases.

In summary, while each model (e.g., Cyclic GAN, IPAM, U-Net, LightCNN-based methods, IHDA) offers solutions for specific challenges, their combined limitations underscore the need for models that are both adaptable and efficient across various forensic scenarios.

2.8 Proposed Solutions for Limitations:

- **Domain Gap Reduction**

  Existing Limitation: Traditional GAN models effectively reduce the domain gap between sketches and photos but are often computationally intensive, making them impractical for real-time forensic use.

  Proposed Solution: We propose optimizing GAN elements within a hybrid model to reduce domain gaps more efficiently. By improving computational efficiency without sacrificing accuracy, our model becomes accessible for forensic teams with limited resources, enabling faster and more reliable sketch-to-photo matching.

- **Identity Preservation**

  Existing Limitation: Identity-preserving methods like Light-CNN are accurate but computationally expensive, limiting their use in resource-constrained forensic settings.

  Proposed Solution: We can streamline identity-preserving techniques using Auto-ML to prioritize key identity features, reducing unnecessary processing. This approach retains essential identity characteristics while lowering computational costs, ensuring fast and accurate matching.

- **Data Efficiency in Hybrid Models**

  Existing Limitation: Hybrid models and methods like IHDA often require large datasets for robustness, limiting adaptability where data access is restricted.

  Proposed Solution: We propose a dual-module design, enhanced by Auto-ML, reduces data requirements while maintaining model robustness. This setup adapts to diverse cases with minimal data, making it practical for real-world forensic applications.

CHAPTER 3
METHODOLOGY

The proposed system for forensic sketch recognition operates in two major stages: **Sketch Construction** and **Sketch Recognition.**

1. **Sketch Construction:** This phase enables users to either upload a sketch or use an interface that leverages a machine learning recommendation model to guide the creation of a sketch.

2. **Sketch Recognition:** Once a sketch is prepared, it is processed through three specialized models to:

- Convert the sketch into a realistic digital image (Model-2).
- Generate feature embeddings from the digital image (Model-3).
- Match these embeddings with entries in the database of criminals to identify the most similar matches (Model-4).

Each model within these phases has a specific purpose and is architected with unique components to meet the system's accuracy, identity preservation, and computational efficiency goals as shown in the Figure 3.1.

3.1 Phase 1: Sketch Construction Module

3.1.1 Model-1: Interactive Recommendation Model

- **Purpose:** Model-1 is a recommendation model based on AutoML and RNNs, specifically designed to assist forensic artists in sketch creation. This model serves as a guide, recommending facial features, proportions, and details based on witness descriptions.

- **Mechanism:**

  - AutoML automates the process of selecting the best algorithm and parameters for generating features that align with the witness-provided description.

  - Recurrent Neural Networks (RNNs) are used to capture sequential dependencies in witness descriptions, helping to dynamically adjust recommendations for sketch details in real-time.

- **Output:** This model ensures that the sketches created are consistent with the description while retaining critical features that aid in later stages of recognition. Once the sketch is finalized, it is passed to the Recognition phase for further processing.

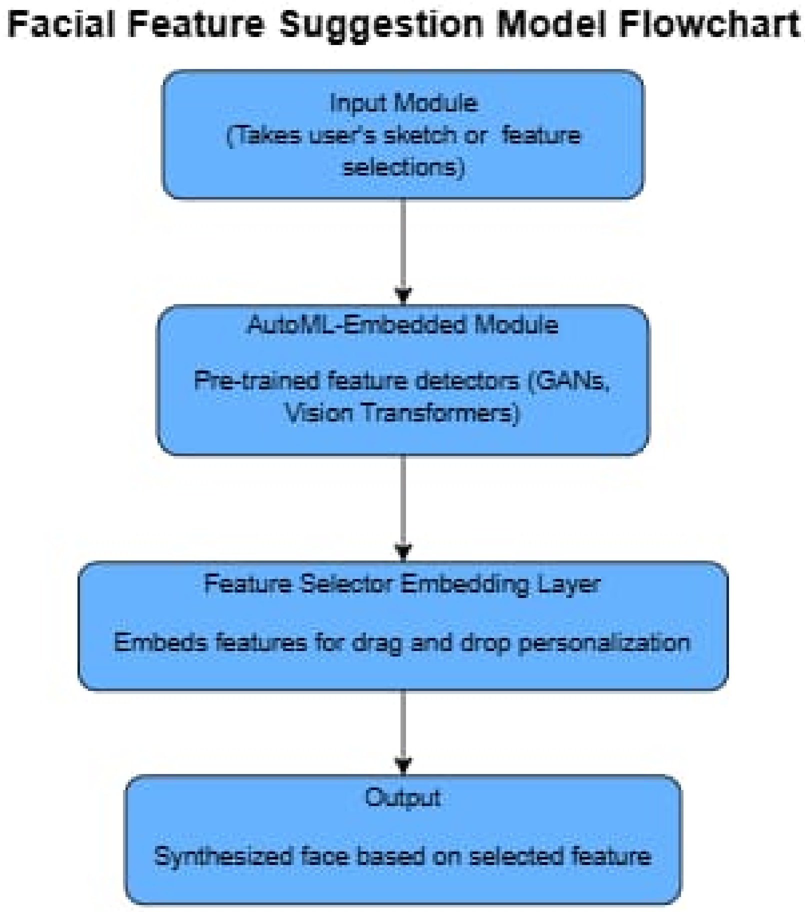

Figure 3.1: Workflow of Facial Feature Recommendation model (Model-1)

3.2 Phase 2: Sketch Recognition Module

The Recognition phase is divided into three key steps, each powered by a specific model to achieve an accurate match with existing records.

3.2.1 Model-2: Converting Sketch to Digital Image Architecture:

- **Generator Network:** The sketch is passed through a U-Net or ResNet generator, which is responsible for converting the sketch into a realistic digital image. These architectures are known for their ability to handle image-to-image translation tasks.
- **Discriminator Network:** This module uses adversarial training with a discriminator that focuses on identity and feature matching, enforcing that the generated image retains critical identity features from the sketch.

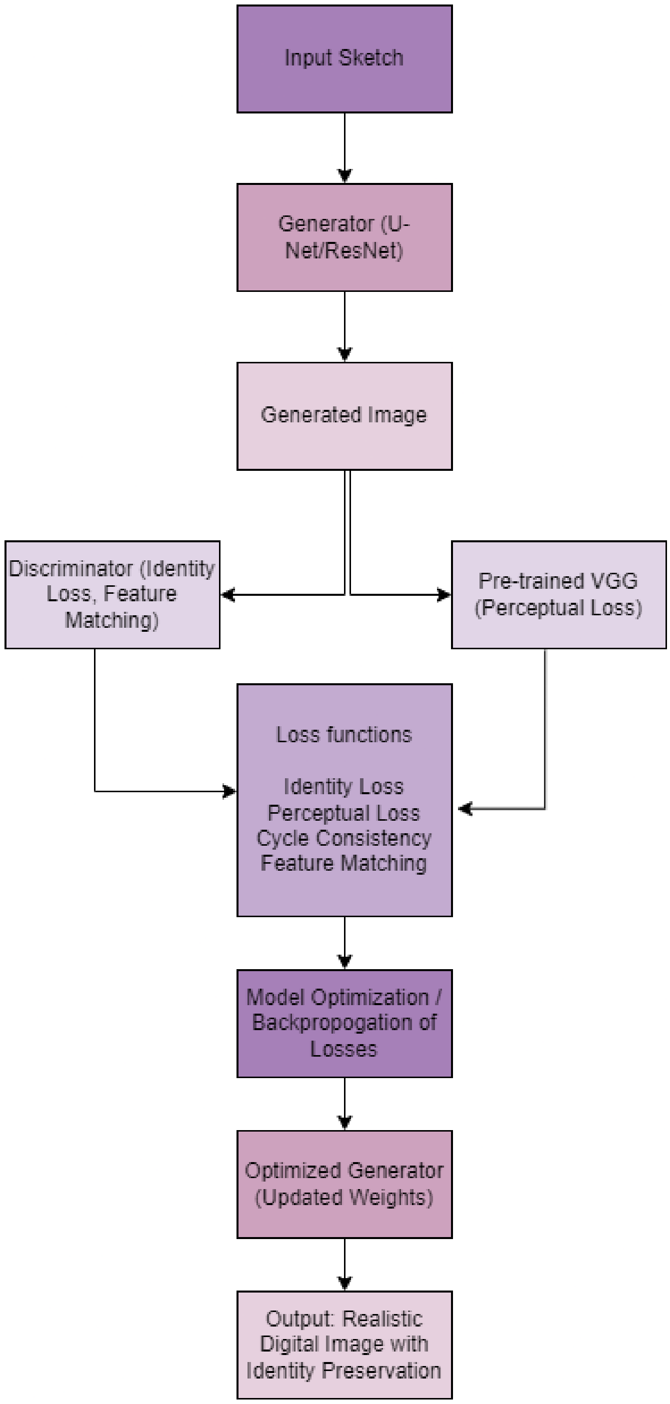

Figure 3.2: Workflow of Model-2

- **Loss Functions:**
  - Identity Loss: Ensures that the digital image maintains identifiable features that match the original sketch.
  - Perceptual Loss: Calculated using a pre-trained VGG model, this loss maintains high-level similarity, ensuring the generated image looks realistic.

  - Cycle Consistency: Helps the model produce consistent results by reducing variance in the outputs.
  - Feature Matching: Ensures that essential features are maintained during transformation.

- **Optimization:** The model iteratively updates weights through backpropagation, optimizing these loss functions to improve image quality and identity retention.

- **Output:** The final output of Model-2 is a high-quality digital image with preserved identity characteristics, making it suitable for embedding generation in the next step.

3.2.2 Model-3: Converting Digital Image to Embeddings Architecture:

- **Face Alignment Module:** The face alignment process begins by standardizing the digital image using a Multi-Task Cascaded Convolutional Network (MTCNN). This network performs face detection, landmark localization, and facial alignment to ensure that the face is correctly oriented. By aligning the face to a canonical pose, MTCNN removes variations caused by head tilt, scale differences, or misalignments, thus providing a consistent foundation for further processing.

- **Feature Extractor Backbone:** Once the face is aligned, an EfficientNet or ResNet-101 network is employed as the feature extractor backbone. These networks, which are fine-tuned on a facial dataset, are designed to capture high-level, identity-rich features from the image. EfficientNet is known for its efficiency in terms of parameter count and computational cost, while ResNet-101 leverages deep residual connections to effectively extract deep features, even from complex or noisy images.

- **Siamese Network with Triplet Loss:** The embedding generated by the feature extractor is refined using a Siamese Network, with Triplet Loss ensuring that embeddings are closer for similar identities and distant for different identities. This alignment process enhances matching accuracy.

- **Output:** The final output of Model-3 is a normalized identity-preserving embedding, which serves as the input for the similarity matching module.

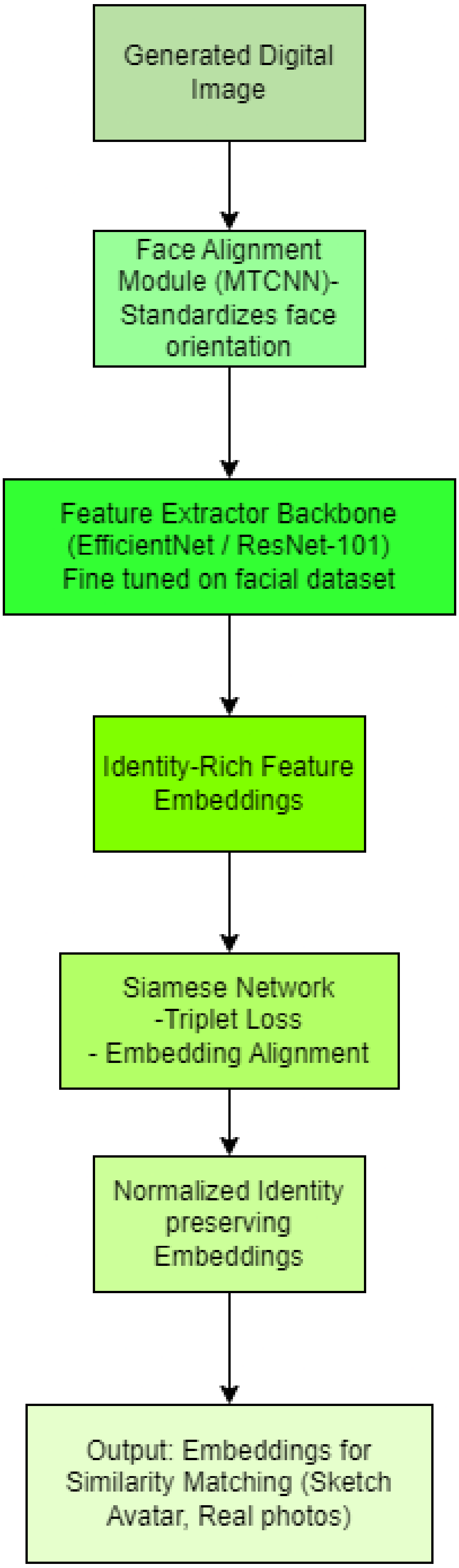

Figure 3.3: Workflow of Model-3

3.2.3 Model-4: Matching Embeddings to Database Architecture:

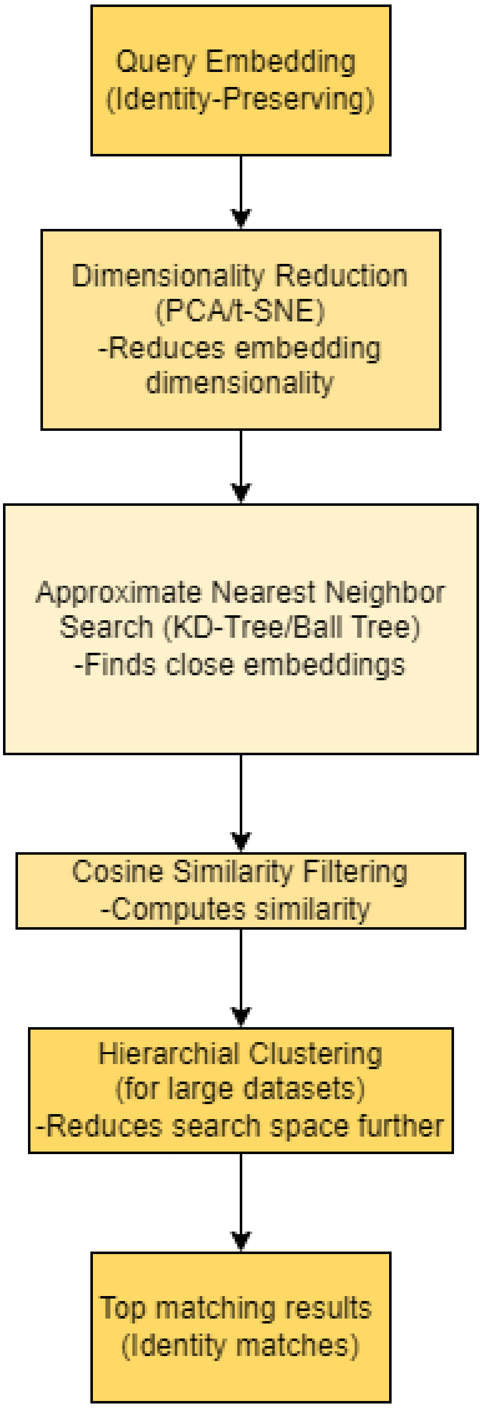

Figure 3.4: Workflow of Model-4

- **Query Embedding (Identity-Preserving):** The query embedding generated by Model-3 is used as the baseline for similarity matching.
- **Dimensionality Reduction (PCA/t-SNE):** To enhance efficiency, the embedding's dimensionality is reduced using Principal Component Analysis (PCA) or t-SNE, simplifying the data while preserving critical identity features.

- **Approximate Nearest Neighbor Search (KD-Tree/Ball Tree):** This component searches for similar embeddings in the database using KD-Tree or Ball Tree structures, which are optimized for finding approximate nearest neighbors in high-dimensional spaces.
- **Cosine Similarity Filtering:** To quantify similarity between embeddings, cosine similarity filtering is applied. This metric ensures that only highly similar embeddings proceed to the next step.
- **Hierarchical Clustering (for large datasets):** For cases with extensive databases, hierarchical clustering is employed to further refine the search space, grouping similar embeddings and improving retrieval efficiency.
- Output: The final output of Model-4 is a list of top matching results from the database, ranked by similarity percentage, with the highest scores representing the most likely matches.

3.3 Summary and Benefits of the System

This dual-module system brings several advantages to forensic sketch recognition:

- **Accuracy:** By using specialized models for each phase, the system ensures high fidelity in both sketch construction and recognition. The use of GANs, RNNs, and CNNs in a complementary setup helps retain identity features, which is crucial for accurate identification.
- **Efficiency:** Dimensionality reduction and clustering techniques in the matching phase reduce computational requirements, allowing for faster processing, even with large databases.
- **Identity Preservation:** Through identity-preserving loss functions and a Siamese Network, the system maintains critical facial features throughout each step, minimizing mismatches.
- **Flexibility:** The system is robust to variations in sketch quality and style, making it adaptable to real-world forensic applications.

The combination of recommendation, sketch-to-image conversion, embedding generation, and similarity matching provides a powerful solution to the challenges in forensic sketch recognition. This approach not only improves the process of sketch generation and matching but also contributes to more reliable and efficient forensic investigations, ultimately supporting law enforcement in the identification of suspects.

3.4 About Datasets

- **[11] CUFS Dataset:** CUHK Face Sketch database (CUFS) is for research on face sketch synthesis and face sketch recognition. It includes 188 faces from the Chinese University of Hong Kong (CUHK) student database, 123 faces from the AR database, and 295 faces from the XM2VTS database. There are 606 faces in total. For each face, there is a sketch drawn by an artist based on a photo taken in a frontal pose, under normal lighting condition, and with a neutral expression.

- **[12] CUFSF Dataset:** The CUHK Face Sketch FERET (CUFSF) is a dataset for research on face sketch synthesis and face sketch recognition. It contains two types of face images: photo and sketch. Total 1,194 images (one image per subject) were collected with lighting variations from the FERET dataset. For each subject, a sketch is drawn with shape exaggeration.

- **[13] FERET Dataset:** The Facial Recognition Technology (FERET) dataset is a widely used database in facial recognition research. It includes a diverse set of facial images captured under various lighting conditions, expressions, and angles. The dataset is valuable for training and evaluating facial recognition algorithms, particularly in handling variations in facial appearance.

- **[14] LFW Dataset:** The Labelled Faces in Wild (LFW) dataset contains over 13,000 labeled images of faces collected from the web. It is developed by researchers from University of Masachussets. The images feature a wide range of poses, expressions, and lighting conditions, making it a valuable resource for studying unconstrained face recognition.

CHAPTER 4

Conclusion

In this project, we addressed the challenges of forensic face sketch recognition, a field that faces considerable technical hurdles, including high computational demands and difficulties in accurately capturing and matching facial details. Through a comprehensive literature review, we identified limitations in existing technologies, such as the reliance on computationally intensive models like GANs and the inability of some systems to handle variations in sketch quality and style. By analyzing these challenges, we observed the need for a more efficient, adaptable, and accurate approach that could enhance the reliability of sketch recognition in real-world forensic applications.

Our proposed methodology consists of a dual-module system designed to optimize each stage of the sketch recognition pipeline. The first module focuses on assisting forensic artists in constructing high-quality facial sketches, using an ML recommendation model built on AutoML and RNNs to guide the sketching process and capture key facial features accurately. The second module is structured for the recognition phase, leveraging a combination of advanced models tailored for different tasks. Model-2 converts sketches into realistic digital images, Model-3 translates these images into identity-rich embeddings, and Model-4 matches these embeddings to a photo database, outputting the top matches based on similarity scores. This systematic approach aims to preserve identity characteristics effectively while maintaining computational efficiency.

Our proposal outlines a structured framework that addresses existing limitations, combining identity-preserving transformations, feature embeddings, and dimensionality reduction techniques. This integrated workflow is intended to enhance accuracy and adaptability in forensic applications, potentially enabling law enforcement agencies to achieve faster and more reliable sketch recognition. Our project sets a strong foundation for developing an advanced tool in forensic sketch recognition, with the potential for ongoing improvements and enhancements.

CHAPTER 5

Future Scope

The proposed forensic face sketch recognition system offers several promising future directions:

- **Advanced Deep Learning Models:** Future work could include more sophisticated models, like transformer-based architectures, for improved sketch-to-image realism and accuracy.
- **Expanded Dataset Diversity:** A broader and more varied training dataset would enhance the model's generalization to diverse facial features and sketch styles, improving robustness in real-world applications.
- **Multi-View Recognition:** Incorporating multi-angle sketches could increase accuracy by capturing different perspectives of a suspect's face.
- **Real-Time Capabilities:** Optimizing the system for real-time recognition could enable instant identification, valuable for surveillance and urgent forensic applications.
- **Handling Partial or Occluded Faces:** Adapting the model to recognize incomplete sketches would make it more practical for cases with limited information.
- **Cross-Modal Retrieval:** Expanding to match sketches with various image sources, including surveillance footage, would increase the system's versatility in identifying suspects.
- **Privacy and Ethics:** Future versions should emphasize strong privacy measures and ethical safeguards to ensure responsible use.

These advancements would greatly enhance the system's effectiveness, adaptability, and ethical application in real-world forensic investigations.

BIBILOGRAPHY

[1] C. Chen K. Du Y. Guo L. Cao and C. Fu. “Domain Alignment Embedding Network for Sketch Face Recognition”. In: *IEEE Access* 9 (2021). DOI: 10.1109/ACCESS.2020.3047108.

[2] Y. Wu R. Xu J. Zheng W. Song and F. Liu. “Feature Encoder Guided Generative Adversarial Network for Face Photo-Sketch Synthesis”. In: *IEEE Access* 7 (2019). DOI: 10.1109/ACCESS.2019.2949070.

[3] K. Fu Y. Lin S. Ling and P. Cheng. “An Identity-Preserved Model for Face Sketch-Photo Synthesis”. In: *IEEE Signal Processing Letters* 27 (2020). DOI: 10.1109/LSP.2020.3005039.

[4] Jun Wang Zheng Liang Chao Lin Chia-Wen Chen Jin Chen. “Identity-Aware Face Super-Resolution for Low-Resolution Face Recognition”. In: *IEEE Signal Processing Letters* (2020). DOI: 10.1109/LSP.2020.2986942.

[5] G. S. Hsu and C. H. Tang. “Dual-View Normalization for Face Recognition”. In: *IEEE Access* 8 (2020). DOI: 10.1109/ACCESS.2020.3014877.

[6] Z. Wang X. Liu J. Jiang Y. Yu and J. Ma. “Graph-Regularized Locality-Constrained Joint Dictionary and Residual Learning for Face Sketch Synthesis”. In: *IEEE Transactions on Image Processing* 28 (2019). DOI: 10.1109/TIP.2018.2870936.

[7] E. A. Saidon N. A. Othman N. N. Bahrum S. Setumin and M. F. Abdullah. “Forensic Face Sketch Recognition based on Pre-Selected Facial Regions”. In: *2022 IEEE 12th International Conference on Control System, Computing and Engineering (ICCSCE), Penang, Malaysia* (2022). DOI: 10.1109/ICCSCE54767.2022.9935651.

[8] Z. Wang X. Li S. Sakti F. Yang Y. Wu and S. Nakamura. “Instance-Level Heterogeneous Domain Adaptation for Limited-Labeled Sketch-to-Photo Retrieval”. In: *IEEE Transactions on Multimedia* 23 (2021). DOI: 10.1109/TMM.2020.3009476.

[9] Trisiladevi Nagavi. “A New Framework for Matching Forensic Composite Sketches With Digital Images”. In: *International Journal of Digital Crime and Forensics* 24 (2021). DOI: 10.4018/IJDCF.20210901.oa1.

[10] C. Galea and R. A. Farrugia. “Forensic Face Photo-Sketch Recognition Using a Deep Learning-Based Architecture”. In: *IEEE Signal Processing Letters* 24 (2017). DOI: 10.1109/LSP.2017.2749266.

[11] *CUFS Dataset*. URL: https://www.kaggle.com/datasets/arbazkhan971/cuhk-face-sketch-database-cufs.

[12] *CUFSF Dataset*. URL: https://mmlab.ie.cuhk.edu.hk/archive/cufsf/#Downloads.

[13] *FERET Dataset*. URL: https://www.nist.gov/itl/products-and-services/color-feret-database.

[14] *LFW Dataset*. URL: https://www.kaggle.com/datasets/jessicali9530/lfw-dataset.