Forensic Face Sketch Construction and Recognition using Deep Learning Techniques

Pasunuri Kathyayini, Cheera Rohan, Gillala Vaishak Reddy, T. Sai Sree

Department of Artificial Intelligence and Data Science, Chaitanya Bharathi Institute of Technology, Telangana, Hyderabad.

Abstract: In the domain of forensic science, facial sketch construction and recognition play a pivotal role, particularly when photographic evidence is unavailable. This literature review surveys the latest developments in forensic face sketch synthesis and recognition, covering a range of approaches that address fundamental issues such as cross-domain compatibility, identity preservation, and robustness to different sketch styles. We systematically survey a variety of deep learning techniques including convolutional neural networks (CNNs), generative adversarial networks (GANs), transfer learning models, feature transformation methods, and novel uses of embedding networks and domain adaptation techniques. Some of the prominent methods discussed are Swish activation with convolutional neural networks (CNN) layers, U-Net architectures with dual discriminators, domain alignment embedding networks, cyclic GAN architectures, graph-regularized dictionary learning and instance-level domain adaptation frameworks. We assess these methods with respect to important forensic constraints - accuracy in face-matching, generalisation across different image sets, and computational efficiency for real-time use. A comparative study shows that GAN-based approaches, especially when combined with identity-preserving modules and attention mechanisms, are very effective in producing recognizable, realistic sketches. While CNN models achieve high accuracy and generalise well across viewpoints, they struggle to maintain subtle facial characteristics that are crucial for forensic applications. In addition, hybrid systems that combine transformer models and attribute-specific learning paradigms, including feature decoupling and domain balancing, show promise for fine-tuning identity retention for low-resolution images. While improvements are notable, there are still issues of computational cost, dealing with varied sketch styles, and preserving reliability in messy real-world environments. This paper offers a complete overview of these developments, chronologically examining the important techniques and innovations that have been developed to address the issues of forensic face sketch recognition. We assess the strengths and weaknesses of each, noting their contributions to the field as well as areas where further research and development might be useful.

Keywords: Forensic Science, Facial Sketches, Face recognition, Generative Adversarial Networks, Convolutional Neural Networks, Identity-preserving modules, Domain alignment embedding networks

1. Introduction:

Forensic science is essential to today's criminal justice system, providing means that are indispensable to investigating crimes and identifying perpetrators. Of these, facial sketches are particularly useful when photographic proof of a suspect is not available. Based on eyewitness accounts, forensic sketches are used as a picture clue to identify suspects via database matches. These sketches, though, rely heavily on how accurately the artist captures the image, how well the witness remembers, and how reliable automated recognition procedures are in identifying the match between sketch and photograph. Classic sketch recognition has long struggled with the domain gap between hand-drawn sketches and photographs - the former lacking in style, shading and detail compared with the latter. Such discrepancies make it difficult for automated systems to reliably compare sketches and photographs. In addition, the quality of the sketch, which depends on the skill of the artist, the clarity of the witness's perception, and the time elapsed since the event, also creates inconsistencies that affect accuracy.

Machine learning has made great strides in the last 10 years, especially deep learning. Convolutional neural networks (CNNs) have improved facial feature extraction from drawings, while generative adversarial networks (GANs) allow sketch-to-photo synthesis, effectively closing the domain gap. And, more recently, new techniques - domain alignment and identity-preserving models - strive to further align sketch and photo representations, making recognition more robust. This paper reviews these developments in detail, examining the primary approaches and innovations for overcoming forensic face sketch challenges. We evaluate each approach's advantages and disadvantages, identifying opportunities for future improvement.

2. Literature Review:

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 part reviews important models, 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 and Photos

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.

Table 1: Literature Survey of papers related to bridging the domain gap between sketches and photos

Ref. No.	Keywords	Domain	Datasets used	Methods used	Evaluation Metrics	Future Work
1	Sketch face recognition, feature embedding network, deep metric learning, small sample problem.	Sketch-based Face Recognition	S1. UoM-SGFSv2 - set A S2. UoM-SGFSv2 - set B S3. PRIP-VSGC	Domain Alignment Embedding Network	**Rank-1:** 1. S1 - 68.53% 2. S2 - 74% **Rank-10:** 1. S1 - 92.4% 2. S2 - 95.2% 3. S3 - 63.2% **Rank-50:** 1. S1 - 97.47% 2. S2 - 99.07%	-**Diverse Sketch Styles:** Enhance model robustness by incorporating various artistic and cartoon sketch styles. -**Cross-Demographic Recognition:** Improve recognition accuracy across different ethnicities, age groups, and genders. -**Real-Time Implementation:** 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.	Face-Photo Sketch Synthesis, Image-to-image translation, Face Recognition	1. CUFS - Chinese University of Hong Kong Face Sketch Database - 606 faces 2. CUFSF - CUHK Face Sketch FERET - 1194 images	A cyclic GAN framework which consists of two generators and discriminators	***Photo->Sketch:*** **1. SSIM Score:** CUFS: 0.5182 CUFSF: 0.3478 **2. LPIPS:** CUFS- 0.2212 CUFSF- 0.2499 ***Sketch->Photo:*** **1. SSIM Score:** CUFS- 0.6325 CUFSF- 0.5842 **2. LPIPS:** CUFS- 0.1830 CUFSF: 0.2523 **Face Recognition Accuracy:** 1. CUFS- 95.34 2. CUFSF- 100	- **Apply Progressive Training:** Improve detail by gradually increasing resolution. - **Use Adaptive Loss Functions:** Balance realism and identity retention dynamically - **Multi-Domain Adaptability:** Adapt EGGAN for varied sketch styles and conditions. - **Cross-Demographic Inclusivity:** Train on diverse datasets for unbiased synthesis. - **Real-Time Optimization:** Enhance efficiency for real-time forensic applications.

2.2. Identity Preservation Across Transformations

Maintaining the integrity of facial features during sketch transformation is crucial for forensic accuracy. Models that preserve identity information during cross-domain transformations have shown greater success in generating recognizable faces, which is essential in criminal investigations.

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

Identity-preserving modules, such as those included in GAN frameworks, help ensure that core facial features remain consistent through transformations. For example, combining IPAMs with cyclic GANs has achieved recognition rates of over 85% on challenging datasets, illustrating that preserving unique identity attributes significantly boosts recognition outcomes.

2.2.2 Feature Decoupling Learning with LightCNN [4]

Feature decoupling learning separates identity-related features from other facial components, ensuring that transformations do not distort critical identity characteristics. LightCNN, paired with ArcFace as a feature extractor, leverages decoupled features, maintaining over 90% accuracy in controlled studies by reducing variability caused by non-identity elements such as lighting or minor facial details.

Table 2: Literature Survey of papers related to Identity Preservation Across Transformations

Ref. No.	Keywords	Problem Domain	Datasets Used	Methods Used	Evaluation Metrics	Future Work
3	Face sketch-photo synthesis, generative adversarial networks, image generation, image-to-image translation.	Face Synthesis, Image to image translation	1. CUFS - Chinese University of Hong Kong Face Sketch Database - 606 faces 2. CUFSF - CUHK Face Sketch FERET - 1194 images	A U-Net generator with two discriminators, optimized by leaky ReLU and Adam, alongside identity-verifying models (ResNet-50 and LightCNN-29v2)	**Rank-1:** 1.CUFS : ResNet-50- 0.8 LightCNN-29v2 - 0.78 2. CUFSF: ResNet-50- 0.737 LightCNN-29v2- 0.701 **VR@FAR=0.1%:** 1. CUFS: ResNet-50 - 0.58 LightCNN-29v2 - 0.48 2. CUFSF: ResNet-50-0.608 LightCNN-29v2- 0.598	- **Handling diverse sketch styles:** Adapt model to various sketch types. - **Cross-Racial and Age-inclusive models:** Broaden demographic coverage and inclusivity. - **Real-time synthesis and recognition:** Enable instant face matching from sketches. - **Integration with other modalities:** Combine with infrared or thermal imaging. **- Synthetic Data Generation for Low-Resource contexts:** Create data for underrepresented populations.
4	Low-resolution face recognition, face super-resolution, identity-aware learning, magnitude loss.	Low-resolution face recognition	1. LFW 2. CelebA	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.	**1. Accuracy:** LFW: LightCNN-v9- 98.46% LightCNN-v9- 98.98% **2. Rank-1:** CelebA: LightCNN-v9- 30.39% LightCNN-v29- 41.36%	1. Exploring adaptive resolution scaling techniques

2.3. Cross-Domain Feature Consistency and Enhancement Techniques

Variability in sketch style and quality, influenced by the artist's skill and the witness's memory, presents another challenge in forensic sketch recognition. To address this, recent models integrate techniques that normalize or standardize features across sketches.

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

This approach specifically tackles the **style and quality variance** among sketches created by different artists. By normalizing these variations, the method improves the robustness of facial feature representations, leading to more consistent and accurate cross-view recognition.

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

This approach focuses on **low-resolution images** with sparse facial details. Through graph regularization and residual learning, it enhances the texture and structural details in such images, allowing for better retention of identity cues that standard techniques may overlook.

Table 3: Literature Survey of papers related to Cross-Domain Feature Consistency and Enhancement Techniques

Ref. No.	Keywords	Problem Domain	Datasets Used	Methods Used	Evaluation Metrics	Future Work
5	Face recognition, face normalization, face synthesis.	Face Recognition	1. Multi-PIE - 75,000 images of 337 people 2. CASIA-WebFace - 494,414 images of 10,575 people 3. IJB-A - 5396 images and 20,412 video frames of 500 subjects 4. IJB-C - 3134 still images and 117542 frames from natural scene video of 3531 different individuals	Dual-View Normalization with Light-CNN and ArcFace as Feature Extractors	**VR@FAR=0.1:** 1. IJB-A: Light-CNN: 95.7 ArcFace: 97,2 2. IJB-C: Light-CNN: 92.4 ArcFace: 96.5 **VR@FAR=0.01** 1. IJB-A: Light CNN: 91.3 ArcFace: 94.2 2. IJB-C: Light-CNN: 87.96 ArcFace: 92.76 **Identification @Rank-1:** 1. IJB-A: Light-CNN: 96.8 ArcFace: 97.4 **@Rank-5:** 1. IJB-A: Light-CNN: 98.7 ArcFace: 98.8	- **Attribute-Based Normalization:** Use unsupervised clustering to tailor normalization. - **Multi-View Expansion:** Incorporate more views beyond dual poses. - **Self-Supervised Learning:** Improve generalization using unlabeled data.
6	Face sketch synthesis, residual learning, joint dictionary learning, local geometry structure, rare characteristics. Face Sketch Synthesis	Face Sketch Synthesis	1. CUHK 2. AR 3. XM2VTS	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	1. Apply a parsing model to differentiate between facial and non-facial regions. 2. Utilize distinct prediction models for improved synthesis accuracy. 3. Implement a luminance recapping algorithm to standardize lighting conditions.

2.4. Adapting to Complex Real-World Scenarios

For forensic face recognition systems to be reliable in the field, they must be resilient to complex scenarios involving inconsistent lighting, occlusions, and partial faces. Hybrid models and multi-feature fusion approaches have demonstrated considerable success in adapting to these variables.

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

By fusing HOG and Gabor Wavelet features through CCA, this method refines matches across varying lighting conditions and occlusions. Additionally, Patches of Interest (PoI) are used to handle partial faces, boosting accuracy by focusing on highly recognizable regions. This multi-feature fusion model achieves around 75–80% accuracy on datasets with complex variations, showing promise for real-world forensic application.

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

IHDA frameworks perform instance-level domain adaptation to account for differences in sketch and photo quality at the instance level, ensuring more accurate matches in challenging conditions. With accuracies reported to exceed 80% in cross-domain tasks, IHDA effectively mitigates variations and aligns instances, proving especially valuable in large forensic databases.

Table 4: Literature Survey of papers related to adaptation to complex real-world scenarios

Ref. No.	Keywords	Problem Domain	Datasets Used	Methods Used	Evaluation Merics	Future Work
7	Forensic, Face sketch, recognition, histogram of gradient, facial region	Face Recogniti on	1. PRIP-HDC	Histogram of Gradient 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.	Sketch-to-photo retrieval	Source: 1. UT-Zap50K 2. CelebFaces 3. Market-1501 Target: 4. QMUL-Shoes 5. IIIT-D Viewed Sketch 6. PKU-Sketch	Instance-level Heterogeneous Domain Adaptation (IHDA) framework	**Sketch-to-photo retrieval:** **Rank-1:** 1. UT-Zap50K- 68.7 2. CelebFaces- 95.7 3. Market-1501- 85.6 **Rank-5:** 3.Market-1501- 94.8 **Rank-10:** 1. UT-Zap50K- 95.7 3. Market-1501- 98.0 **Rank-20:** 3.Market-1501- 100 **Photo-to-sketch retrieval:** **Rank-1:** 1. UT-Zap50K- 69.6 2. CelebFaces- 96.2 3. Market-1501- 88.2 **Rank-10:** 1. UT-Zap50K- 97.4 2. CelebFaces- 98.6 3. Market-1501- 100 **Rank-20:** 1. UT-Zap50K- 99.1 2. CelebFaces- 99.2 3. Market-1501- 100	1. Explore effective methods to leverage rich-labeled datasets for limited-labeled tasks. 2. Investigate the IHDA framework's applicability in scenarios beyond instance-level retrieval tasks. 3. Develop strategies to identify shared attributes for category-level retrieval tasks across datasets like Sketchy, PACS, and M3SDA.

2.5 CNN-Based Techniques for Forensic Face Sketch Recognition

CNN-based methods have become central to forensic face sketch recognition due to their powerful feature extraction capabilities. These techniques are well-suited to handle the complex variations present in forensic sketches, such as differences in age, style, and resolution.

2.5.1 A New Framework for Matching Forensic Composite Sketches with Digital Images [9]

This approach leverages a six-layer CNN paired with Swish activation functions to improve recognition accuracy across sketches of varying quality and age. Using the E-PRIP and Composite Sketch with Age Variation (CSA) datasets, the model achieved accuracies of 78.26% and 69.57%, respectively. The framework highlights the benefits of lightweight CNNs and dataset expansion for improved matching between sketches and digital images in forensic contexts.

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

Utilizing a deep convolutional neural network (DCNN) combined with transfer learning, this method applies a 3D Morphable model to capture variations in facial features. By fusing synthetic and original sketch distance measures, the model enhances identity matching. Tested on datasets such as CUFS and Color FERET, this method demonstrates effective photo-sketch matching capabilities with improvements in mean rank retrieval rates, making it a promising option for forensic photo-sketch recognition tasks.

Table 5: Literature Survey of papers related to CNN-Based Techniques for Forensic Face Sketch Recognition

Ref. No.	Keywords	Problem Domain	Datasets Used	Methods Used	Evaluation Merics	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	Forensic Composite Face Sketch Matching with Digital Images	1. E-PRIP dataset (123 composite sketches and its respective digital images) 2.Composite Sketch with Age Variation dataset collected from IIIT, Delhi (3529 sketches and digital images from 150 individuals)	Swish Activation Function with 6 layer CNN	**Accuracy:** 1. E-PRIP - 78.26% 2. CSA - 69.57% **Precision:** 1. E-PRIP - 63.64% 2. CSA - 54.55% **Recall (TPR):** 1. E-PRIP - 87.50% 2. CSA - 75% **F1-score:** 1. E-PRIP - 73.68% 2. CSA - 63.16% **FPR:** 1. E-PRIP - 0.2667 2. CSA - 0.3333	- Experiment with Lightweight CNN Models: Test multiple compact CNN architectures to improve efficiency. - Combine Datasets: Merge additional datasets to increase image count, enhancing model performance. - Test on Larger Datasets: Increase dataset size to improve deep network performance. - Advanced Neural Networks: 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.	FacePhoto-Sketch Recognition	1. CUFS 2. XM2VTS 3. AR 4. CUHK 5. CUFSF 6. Color FERET 7. PRIP-HDC(Testing)	A deep convolutional neural network (DCNN) on which transfer learning is applied DEEPS: uses a 3-D Morphable model to generate variations of facial features and attributes for face images DEEPS-M: enhances identity determination in forensic sketches by fusing distance measures from synthetic and original sketches for improved photo matching.	**Mean Rank Retrieval Rate:** **1. DEEPS:** 325.02 **2. DEEPS-M:** 312.11	1. Developing a more flexible 3-D Morphable model for better facial feature variation. 2. Applying proposed methods to additional human face recognition tasks.

Models like these represent tremendous strides in overcoming challenges in forensic face sketching. GAN-type methods, domain-alignment techniques, and identity-preserving modules show great success in minimizing domain gaps, and accuracies typically fall between 75 and above 90 per cent depending on the dataset used and the complexity of the sketches involved. Style and resolution are handled by dual normalisation and super-resolution models, which allow recognition accuracy to remain high under diverse conditions. Other models that were included in the study, and that have demonstrated marked effects, are illustrated in the table below:

Table 6: Literature Survey of papers related to Face Recognition

Ref. No.	***Paper Title***	***Keywords***	***Problem Domain***	***Datasets Used***	***Methods Used***	***Evaluation Merics***	***Future Work***
11	SP-Net: A Novel Framework to Identify Composite Sketch	Composite sketch, hand-drawn sketches, convolutional neural network, contrastive loss.	Facial Recognition, Composite sketch identification	E-PRIP CUHK FEI CASPEAL MGDB AR FERET SCface	A coupled deep convolutional neural network, Sketch-Photo Net (SP-Net), incorporating VGG-Face as a base, Siamese network architecture, contrastive loss function, and elastic learning with dimensionality reduction	**1. E-PRIP** rank-1 (28.3%) rank-5 (53.1%) rank-10 (80.0%).	1. Adapting this approach to hand-drawn sketch recognition and composites generated through the Identikit software
12	Toward Realistic Face Photo–Sketch Synthesis via Composition-Aided GANs	Deep learning, face parsing, face photo–sketch synthesis, generative adversarial network (GAN), image-to-image translation.	Face-sketch photo synthesis, Image-to-image translation	1. CUFS 2. CUFSF	1. CA-GAN: generates the sketch portrait based on the face photo and composition masks: 2. SCA-GAN: refine the generated sketch portraits	**1. Avg. performance on FID Criterion:** CA-GAN: 32.7 SCA-GAN: 30.5 **2. Avg. performance on FSIM Criterion:** CA-GAN: 81.1 SCA-GAN: 82.0 **3. Avg. performance on NLDA Criterion:** CA-GAN: 99.2 SCA-GAN: 99.7	1. Using dimensionality reduction before computing FID might be a solution to th indeterminate reliability of FID as the dimension of deep features is dramatically higher than the number of photo/sketch samples
13	Multi-Task Explainable Quality Networks for Large-Scale Forensic Facial Recognition	Face image quality, explainable AI, multi-task learning, forensics.	Forensic Facial recognition	1. UTKFace 2. LFW 3. XQLFW 4. SCFace 5. ForenFace	1. XQNEt-ConvNet 2. XQNet-EfficientNet		Optimizing XQNet for edge devices, improving robustness for low-quality images, expanding explainability to reduce bias, and generalizing across diverse datasets and forensic tasks
14	Domain Balancing: Face Recognition on Long-Tailed Domains	Long-tailed domains, Domain balancing, Residual Balancing Mapping, Domain Frequency Indicator Domain Balancing Margin	Face-Recognition	Training: 1. CASIA-Webface 2. MS-Celeb-1 Testing: 1. LFW 2. CPLFW 3.CALFW 4. AgeDB 5. RFW 6. CACD 7. MegaFace	1. Domain Balancing mechanism	**1. Rank-1:** 96.35% **2. VA@FAR=10^(-6):** 96.56%	1. Developing a dynamic domain adaptation approach that continuously learns and adapts to new or underrepresented domains in real-time

15	Canonical Correlation Analysis Feature Fusion with Patch of Interest: A Dynamic Local Feature Matching for Face Sketch Image Retrieval	Identity of Interest, Patch of Interest, Sketch-to-Photo, Face Sketch, Forensic, Image Retrieval, CCA Fusion, Score Fusion, Deep Learning.	Face Sketch Image Retreival	1. CUFS 2. CUFSF 3. IIIT-Delhi Semi-Forensic Sketch Database 4. PRIP-HDC (Forensic Sketches)	1. Uses Histogram of Oriented Gradient (HOG) and Gabor Wavelet (GW) features fused via Canonical Correlation Analysis (CCA) to shortlist the top k similar photos. 2. Another method fortackling illumination problem is to use CCA Fusion with image represented by Difference of Gaussian Oriented Gradient Histogram (Do-GOGH) 3.The second block refines matches using local features on Patches of Interest (PoI).	**Rank-1 Accuracy:** 1. CCA Fusion: 93.89% 2. CCA Fusion+D-DoGOGH: 95.23% 3. CCA Fusion+D-DoGOGH on POI: 95.48%	1. Use attention mechanisms to dynamically select and weigh the most informative facial regions for matching, increasing robustness to variations in sketches and photos. 2. Apply cross-modal embedding techniques, like contrastive or triplet loss, to better align sketch and photo representations, reducing modality gaps caused by differences in texture, lighting, and detail.
16	Image Retrieval Using Convolutional Autoencoder, InfoGAN, and Vision Transformer Unsupervised Models	Feature extraction, InfoGAN, sketched-real image retrieval, object matching, spatial distance measurement, vision transformer.	Content-Based Image Retrieval, Facial Sketched-Real Image Retrieval	1. ESRIR	1. Convolutional Autoencoder 2. Information Maximizing GANs 3. Vision Transformer	**F1-Score:** ESRIR: 1. CA- 0.93 2. InfoGAN- 1.272 3. ViT- 1.183 QuickDraw-Extended: 1. CA- 0.64 2. InfoGAN- 0.534 3. ViT- 0.81 256_Object Categories: 1. CA- 0.22 2. InfoGAN- 0.533 3. ViT- 0.497	- Explore distance metrics beyond Euclidean distance for retrieval systems. - Use combined algorithms like capsule networks with existing methods to evaluate efficacy. - Investigate various AI methods on different datasets for improved performance.

3. 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.

Table-7: Comparative Analysis of Forensic Face Sketch Recognition Models Based on Primary Challenges

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%+

3.1 Summary of Comparative Analysis

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

- 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 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.

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

4. Future Work

The state of the art in sketch recognition is still subject to many difficulties, including the high computational expense of Generative Adversarial Networks (GANs) and a tendency to miss important facial details, leading to inaccuracies in both sketch construction and recognition. To overcome these limitations, our solution is a dual-module system, with two separate, but complementary, phases.

In the Face Sketch Construction module, we use a convolutional neural network (CNN) or similar machine learning model that serves as a kind of helper to forensic artists by suggesting facial details from witness descriptions. This model concentrates on the efficient capture and reconstruction of key facial characteristics, allowing for a more accurate and expedient creation of the sketches. Using a highly organised method for feature extraction and selection, this module improves the accuracy of the sketches generated, something of great importance in forensic work where accuracy is everything.

In the Recognition module, we use a hybrid method of GANs and Identity-Preserving Adversarial Matching (IPAM). This coupling enables us to create high-fidelity sketches and at the same time maintain salient identity characteristics that are usually sacrificed in standard GAN architectures. The GAN part then generates realistic sketches from the derived features, and the IPAM module makes sure that the person's identity is recognisable throughout the matching process. This double duty improves the system's ability to match sketches with real images by preserving essential facial features that could otherwise be degraded.

By combining these advanced techniques, our method significantly enhances adaptability, accuracy, and computational efficiency in sketch recognition systems. This integrated approach not only improves the reliability of sketch recognition in practical forensic applications but also positions itself as a crucial advancement for law enforcement. In the end, it gives forensic artists a powerful tool that makes the whole business of sketch creation and identification a lot more efficient, and the whole business of investigation, and ultimately the whole business of justice, a lot more effective.

5. Conclusion

In summary, this review of the literature on forensic face sketch construction and recognition identifies important progress in alleviating some of these challenges, including domain gap reduction, identity preservation, sketch quality, and robustness to real-world conditions. These methods, such as Cyclic GANs and Domain Alignment Embedding Networks, successfully alleviate domain discrepancies, and Identity-Preserving GANs (IPAM) and feature decoupling learning maintain identity-preserving consistency in recognition. Other models like the U-Net with two discriminators also increase the quality of the sketches, and Identity-Aware Super-Resolution Networks make low-res sketches clearer. This adaptability is also tackled by techniques such as HOG Gabor Wavelet Fusion and Instance-Level Heterogeneous Domain Adaptation (IHDA). Together, these developments highlight the importance of developing a dual-module system that improves both the construction and recognition accuracy of sketches, and gives law enforcement more robust tools for suspect identification and forensic investigation.

6. References

[1] Y. Guo, L. Cao, C. Chen, K. Du and C. Fu, “Domain Alignment Embedding Network for Sketch Face Recognition”, IEEE Access, Vol. 9, PP. 872-882, 2021, DOI: 10.1109/ACCESS.2020.3047108

[2] J. Zheng, W. Song, Y. Wu, R. Xu and F. Liu, “Feature Encoder Guided Generative Adversarial Network for Face Photo-Sketch Synthesis”, *IEEE Access*, Vol. 7, PP. 154971-154985, 2019, DOI: 10.1109/ACCESS.2019.2949070.

[3] Y. Lin, S. Ling, K. Fu and P. Cheng, “An Identity-Preserved Model for Face Sketch-Photo Synthesis”, IEEE Signal Processing Letters, Vol. 27, PP. 1095-1099, 2020, DOI: 10.1109/LSP.2020.3005039.

[4] Chen, Jin & Chen, Jun & Wang, Zheng & Liang, Chao & Lin, Chia-Wen, “Identity-Aware Face Super-Resolution for Low-Resolution Face Recognition”, IEEE Signal Processing Letters. PP. 1-1, . (2020), DOI: 10.1109/LSP.2020.2986942.

[5] G. -S. Hsu and C. -H. Tang, “Dual-View Normalization for Face Recognition”, IEEE Access, vol. 8, pp. 147765-147775, 2020, DOI: 10.1109/ACCESS.2020.3014877.

[6] J. Jiang, Y. Yu, Z. Wang, X. Liu and J. Ma, "Graph-Regularized Locality-Constrained Joint Dictionary and Residual Learning for Face Sketch Synthesis," IEEE Transactions on Image Processing, vol. 28, no. 2, pp. 628-641, Feb. 2019, DOI: 10.1109/TIP.2018.2870936.

[7] N. N. Bahrum, S. Setumin, E. A. Saidon, N. A. Othman and M. F. Abdullah, "Forensic Face Sketch Recognition based on Pre-Selected Facial Regions," 2022 IEEE 12th International Conference on Control System, Computing and Engineering (ICCSCE), Penang, Malaysia, 2022, pp. 174-179, DOI: 10.1109/ICCSCE54767.2022.9935651.

[8] F. Yang, Y. Wu, Z. Wang, X. Li, S. Sakti and S. Nakamura, "Instance-Level Heterogeneous Domain Adaptation for Limited-Labeled Sketch-to-Photo Retrieval," IEEE Transactions on Multimedia, vol. 23, pp. 2347-2360, 2021, DOI: 10.1109/TMM.2020.3009476.

[9] Nagavi, Trisiladevi., “A New Framework for Matching Forensic Composite Sketches With Digital Images. International Journal of Digital Crime and Forensics. 13. 1-19., 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," IEEE Signal Processing Letters, vol. 24, no. 11, pp. 1586-1590, Nov. 2017, DOI: 10.1109/LSP.2017.2749266.

[11] H. Cheraghi and H. J. Lee, "SP-Net: A Novel Framework to Identify Composite Sketch," IEEE Access, vol. 7, pp. 131749-131757, 2019, DOI: 10.1109/ACCESS.2019.2921382.

[12] Yu, Jun & Xu, Xingxin & Gao, Fei & Shi, Shengjie & Wang, Meng & Tao, Dacheng & Huang, Qingming, “Toward Realistic Face Photo-Sketch Synthesis via Composition-Aided GANs”, IEEE Transactions on Cybernetics. PP. 1-13., 2020, DOI:10.1109/TCYB.2020.2972944.

[13] A. M. Rodriguez, L. Unzueta, Z. Geradts, M. Worring and U. Elordi, "Multi-Task Explainable Quality Networks for Large-Scale Forensic Facial Recognition," IEEE Journal of Selected Topics in Signal Processing, vol. 17, no. 3, pp. 612-623, May 2023, DOI:10.1109/JSTSP.2023.3267263.

[14] D. Cao, X. Zhu, X. Huang, J. Guo and Z. Lei, "Domain Balancing: Face Recognition on Long-Tailed Domains," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020, pp. 5670-5678, DOI: 10.1109/CVPR42600.2020.00571.

[15] Setumin, Samsul & Che Aminudin, Muhamad Faris & Suandi, Shahrel Azmin, “Canonical Correlation Analysis Feature Fusion With Patch of Interest: A Dynamic Local Feature Matching for Face Sketch Image Retrieval”, IEEE Access. PP. 1-1, 2020, DOI:10.1109/ACCESS.2020.3009744.

[16] E. S. Sabry et al., "Image Retrieval Using Convolutional Autoencoder, InfoGAN, and Vision Transformer Unsupervised Models," IEEE Access, vol. 11, pp. 20445-20477, 2023, DOI: 10.1109/ACCESS.2023.3241858.