International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

FORENSIC SKETCH DETECTION BASED ON ANFIS AND MULTIVIEW LEARNING

Mahtab Alam[1], Umesh Kumar[2], Vikram Kaushik[3]

[1]Professor, Dept. of Computer Science & Engineering (AIML)

Parul University, Vadodara, Gujarat, India

mahtab.alam40472@paruluniversity.ac.in

[2]Assistant Professor,

Dept. of Computer Science & Engineering (AIML)

Parul University, Vadodara, Gujarat, India

Umeshkumar.kumar38823@paruluniversity.ac.in

[3]Professor, Dept. of Computer Science & Engineering

Parul University, Vadodara, Gujarat, India

vikram.kaushik40608@paruluniversity.ac.in

Abstract-

In forensic science, accurately linking a hand-drawn sketch to a real photo remains a challenging task due to the sensory gap between modalities and artistic exaggerations in sketches. Conventional techniques—whether handcrafted feature extraction or deep learning transformations—often struggle to generalize across varied facial structures. This paper introduces a sketch recognition framework that integrates Multiview Learning with an ANFIS for dynamic feature fusion. The method employs a GAN to generate photo-like images from sketches, reducing cross-modal disparity. Four complementary descriptors—HOG, ORB, LBP, and DoG—are then extracted as multiple views. Finally, ANFIS adaptively assigns feature weights to optimize the fusion process. Experimental evaluation on the CUFS dataset and the AR face dataset demonstrates that the proposed system surpasses state-of-the-art approaches, achieving 98.2% Rank-1 accuracy on CUFS and 97.65% on AR, with 100% precision at Rank-3 and Rank-5. Further analysis using SSIM, FSIM, and VIF metrics validates the strength of ANFIS-driven fusion. Unlike traditional fuzzy systems relying on fixed rules, ANFIS dynamically tunes feature selection, enhancing robustness and generalization. The results establish a new benchmark in forensic sketch recognition, highlighting the practical significance of combining Multiview Learning with adaptive fusion strategies.

***Key Words*—**Adaptive Neuro-Fuzzy Inference System (ANFIS), Feature Fusion in Face Recognition, Generative Adversarial Network, Multiview learning, Sketch-to-Photo Matching.

Received: August 02, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

I. INTRODUCTION

Facial sketches play a crucial role in forensic investigations, where witness descriptions are often converted into drawings by forensic artists. Such sketches can then be matched with mugshots or photo databases to aid in suspect identification. However, recognition accuracy is frequently hindered by two key challenges: **(1)** the *shape-exaggeration effect*, caused by inaccuracies in human memory and drawing, and **(2)** the *cross-modal gap* between hand-drawn sketches and photographic images, which makes pixel-level comparisons ineffective [1], [2].

Over the years, researchers have proposed a variety of techniques to address sketch-to-photo recognition. Linear and non-linear subspace learning models attempted to align the two modalities [3], while patch-based approaches such as **Markov Random Fields (MRF)** focused on localized similarity [4]. Despite their contributions, these methods suffer when applied to sketches drawn from verbal descriptions rather than direct photos [5], leading to reduced reliability in real-world forensic settings [6].

To overcome these issues, **Multiview Learning** has emerged as a promising direction. By representing data through multiple feature sets (or "views"), this approach exploits complementary information to strengthen recognition. For instance, Dalal et al. utilized **HOG** and **GLCM** features for improved matching [7], while Setumin and Suandi proposed **DoGOGH**, integrating DoG with HOG for salient feature extraction [8]. Extensions combining **Gabor Wavelets** and **Canonical Correlation Analysis** further enhanced feature representation [9]. Similarly, **Multiview Discriminant Analysis (MvDA)** aimed to unify features into a common subspace [10]. Nonetheless, these methods generally assume linearity, which limits their adaptability to complex forensic scenarios.

Recent research has shifted toward **non-linear representations**. You et al. explored non-linear subspace learning [11], and Chugh et al. applied transfer learning with evolutionary optimization [12]. Peng et al. further addressed data scarcity using domain translation across handcrafted and CNN-based descriptors [13].

Building upon these works, the present study introduces a **Multiview Learning framework enhanced with GAN-based modality bridging and ANFIS-driven feature fusion**. Four distinct descriptors—HOG, ORB, LBP, and DoG—are used to capture structural, key-point, texture, and contour information. GANs generate photo-realistic faces from sketches, reducing modality gaps, while ANFIS adaptively learns optimal fusion strategies.

Kan et al. proposed Multiview Discriminant Analysis (MvDA) to project multiple feature views onto a common subspace, maximizing intraclass variation while minimizing interclass variation [10]. However, these methods primarily assume a linear relationship between extracted sketch and photo features, which limits their real-world applicability, especially in forensic scenarios involving highly similar faces.

Recent efforts have explored non-linear methods to address these limitations. You et al. investigated nonlinear feature representations in subspace learning [11]. Chugh et al. utilized transfer learning with evolutionary optimization to fuse complementary sketch and photo features, although their approach relies on large-scale annotated data, which is often

Received: August 02, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

unavailable in forensic cases [12]. Similarly, Peng et al. introduced a Multiview domain translation technique using handcrafted descriptors alongside CNN-based representations, mitigating data availability constraints by reducing inter-domain discrepancies.

This paper introduces a Multiview learning approach to improve sketch-to-photo detection. Instead of relying on a single descriptor, the proposed system extracts four distinct views from sketch images using Histogram of Oriented Gradi- ents (HOG) [14], ORB [15], LBP [16], and DoG [17]. These feature descriptors capture diverse facial attributes, offering a more robust representation.

To further improve performance, a GAN is used to synthesize photorealistic images from sketches while preserving critical facial attributes such as beards, moles, and wrinkles [18], [19]. Additionally, an ANFIS is introduced to optimize feature selection, dynamically learning the best feature fusion strategy from training data. By integrating these components, the proposed system effectively bridges the modality gap and enhances sketch-to-photo matching accuracy.

A. Contributions

The key contributions of this work are:

- A novel Multiview learning framework that enhances sketch-to-photo identification by leveraging multiple feature representations.
- A GAN-based synthesis model that generates realistic photo images while addressing the shape-exaggeration phenomenon.
- An ANFIS that learns feature weighting dynamically, ensuring optimal view fusion.
- Extensive evaluation on CUFS [20] and AR datasets [4], demonstrating superior performance over existing methods.

The rest of this paper is organized as: Section II details the proposed methodology followed by Section III describes the experimental setup, Result and discussion is detailed in section IV. Finally, Section V concludes the study and discusses future work.

II. PROPOSED METHODOLOGY

This section presents the proposed Multiview learning approach for sketch-to-photo identification. The system leverages multiple feature descriptors to generate diverse representations, each capturing unique statistical properties to enhance model inference. The methodology is structured into three phases: (a) Inter-Modality Gap Bridge, (b) Multiview Representation and Learning, (c) View Controller.

A. Problem Formulation

Given a dataset of N sketch images

$\vec{x} = [x_1,x_2,x_3,\ldots\ldots..x_N] \in \Re^{N \times Di}$ and their corresponding photo images $\vec{y} = [y_1,y_2,_3,\ldots\ldots..x_N] \in \Re^{N \times Dj}$

The goal is to study a mapping function that accurately retrieves the most similar picture

Received: August 02, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

y_j for a give query sketch x_i

Step 1: Modality Gap Bridging A Generative Adversarial Network (GAN) is introduced to generate photo-like images from sketches, ensuring better structural and textural alignment:

$$\hat{x} = G(x_j) \approx y_j, \qquad \forall x_j \in \vec{x}, y_j \in \vec{y}\,. \tag{1}$$

Step 2: Multiview Feature Extraction To enhance retrieval accuracy, four feature descriptors extract different views of the generated photo $\hat{x}$: $a, b, c, d = F(f_1(\hat{x}), f_2(\hat{x}), f_3(\hat{x}), f_4(\hat{x}))$, (2)

where f_1, f_2, f_3, f_4 correspond to HOG, ORB, LBP, and DoG features.

Step 3: Feature Fusion Using ANFIS ANFIS is used to learn the importance of features and optimize Multiview fusion. ANFIS is trained using a hybrid learning approach combining gradient descent and least-squares estimation:

$$\text{ZANFIS}(a, b, c, d) = \sum_{i=1}^{m} wifi(a, b, c, d) \tag{3}$$

where w_i represents the learned weight for each feature view, and f_i are the membership functions optimized during training.

Step 4: Feature Matching and Retrieval The ANFIS output is used to compute similarity between the query sketch and stored feature vectors in the Codebook.

Bhattacharyya distance is used to retrieve the closest match:

$$F_{min} = \min_{i=1}^{N} [D_B(a, A_i), D_B(b, B_i), D_B(c, C_i), D_B(d, D_i)] \tag{4}$$

The final retrieval output $\hat{y}$ is given by: $\hat{y} = \text{argmin}\, F_{min}$. (5)

A. *Phase 1: Inter-Modality Gap Bridge*

To address the modality gap, a Generative Adversarial Network (GAN) is employed to synthesize photo-realistic images from sketches. Traditional Neural Style Transfer (NST) methods [21] fail to preserve salient facial attributes due to poor texture separation; Fig. 1 illustrate this failure. Instead, the proposed GAN leverages residual networks and skip connections to improve feature retention while reducing shape distortions [22]. Fig. 2 and Fig. 3 showcase the architecture of the generator and discriminator network for the proposed GAN.

The loss function incorporates three key components:

- **Adversarial Loss** [18]: Guides the generator to produce realistic images by minimizing:

$$L_G = E_y[\log D(y)] + E_x[\log(1 - D(G(x)))] \tag{6}$$

Received: August 02, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

- **Reconstruction Loss**: Encourages sketch-to-photo similarity while allowing flexibility for artistic distortions:

$$L_{\delta}(G,X) \begin{cases} = \frac{1}{2}\left|\left|\nabla x\right|\right|1, if\left|\left|\nabla x\right|\right| < \delta 1 \\ \text{else} \quad \delta||\nabla x||1 - \frac{1}{2}\delta^2 \end{cases} \tag{7}$$

- **Contextual Loss** [23]: Ensures that perceptual consistency is maintained between sketches and generated photos.

-

B. *Phase 2: Multiview Representation and Learning*

This phase extracts multiple representations of images using four distinct feature descriptors:

1 **Histogram of Oriented Gradients** (HOG): Captures gradient-based structural patterns, enhancing edge detection and object shape representation [14].

2 **Oriented FAST and Rotated BRIEF** (ORB): Extracts key point-based descriptors that remain invariant to rotation and illumination [15].

3 **Local Binary Pattern** (LBP): Encodes local texture features by thresholding neighborhood pixel intensities [16].

4 **Difference of Gaussian** (DoG): Enhances contrast variations and highlights facial contours [17].

Each feature descriptor provides a unique view of the image, enabling a more comprehensive facial representation. These extracted Multiview representations are stored in a Codebook as:

$$C = \{(\vec{A}_i, \vec{B}_i, \vec{C}_i, \vec{D}_i) \mid i = 1, 2, \ldots, N\} \tag{8}$$

This phase ensures that input image representations are rich and diverse, leading to improved retrieval accuracy in forensic face identification.

C. *Phase 3: View Controller Using ANFIS*

To ensure feature complementarity while avoiding conflicts, an ANFIS is introduced. Unlike a traditional Fuzzy Inference System (FIS), which requires predefined rules and membership functions, ANFIS integrates fuzzy logic with neural networks, allowing it to learn the best feature fusion strategy from training data. ANFIS is structured as a five-layer hybrid model, where:

- *Layer 1 - Fuzzification:* Converts input feature descriptors (HOG, ORB, LBP, DoG) into fuzzy membership values.
- *Layer 2 - Rule Learning:* Generates fuzzy IF-THEN rules based on input feature distributions.
- *Layer 3 - Rule Normalization:* Assigns rule weights based on importance learned from training data.

Received: August 02, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

- *Layer 4 - Weighted Output Computation:* Computes feature fusion using learned weights.
- *Layer 5 - Defuzzification:* Produces a final score indicating feature relevance for decision-making.

Finally, given a query sketch, ANFIS assigns optimized feature weights and fuses the four extracted views (a, b, c, d). The fused feature vector is then matched against stored representations using the Bhattacharyya distance:

$$F_{min}=\min_{i=1}^{N}[D_B(a, A_i), D_B(b, B_i), D_B(c, C_i), D_B(d, D_i)] \quad (9)$$

The final retrieval output $\hat{y}$ is given by: $\hat{y} = \text{argmin}\, F_{min}$ (10)

a b c

Fig. 1. Demonstrating the effect of Neural Style Transfer in this problem domain. (a) Sketch to Photo . (b) Sketch to Photo. (c) Photo to Sketch.

This phase ensures that the best combination of Multiview features contributes to sketch-to-photo identification while adapting to different feature relationships. The experimental setup and evaluation of this framework are discussed in Section III.

III. EXPERIMENTAL SETUP

The proposed system was implemented in a Python 3.11.0 programming language.

A. *Datasets*

Experiments were conducted using the Chinese University of Hong Kong Face Sketch (CUFS) dataset [4] and the AR face dataset [20]. The dataset consists of 311 paired images (188 from CUHK students and 123 from AR). Each sketch was drawn by an artist based on its corresponding photo. The dataset includes variations in ethnicity, lighting, and background, ensuring robustness in model evaluation.

For training and testing:

- *CUFS*: 88 faces for training, 100 for testing.
- *AR dataset*: 126 faces for training, 37 for testing.

B. *Pre-processing*

Received: August 02, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

To improve computational efficiency and accuracy, several pre-processing steps were applied:

- *Face Alignment*: Three fiducial points (left eye, right eye, chin) were detected to align faces, minimizing shape- exaggeration distortions.
- *Background Removal*: Mugshot photos typically have uniform backgrounds, which were removed to prevent unnecessary interference.
- *Image Cropping and Resizing*: Each image was resized to 250 × 250 pixels for consistency.
- *Data Augmentation*: Applied horizontal flipping and random cropping to improve model generalization.
- *Bilateral Filtering*: Used to enhance edges and reduce noise, preserving fine-grained facial details.

C. *Performance Evaluation Metrics*

Model evaluation was conducted using:

- *Rank-N Accuracy*: Measures correct retrieval of a target photo in Rank-3 or Rank-5.
- *Structural Similarity Index Metric (SSIM)* [24]: Evaluates perceptual similarity between generated and real images.

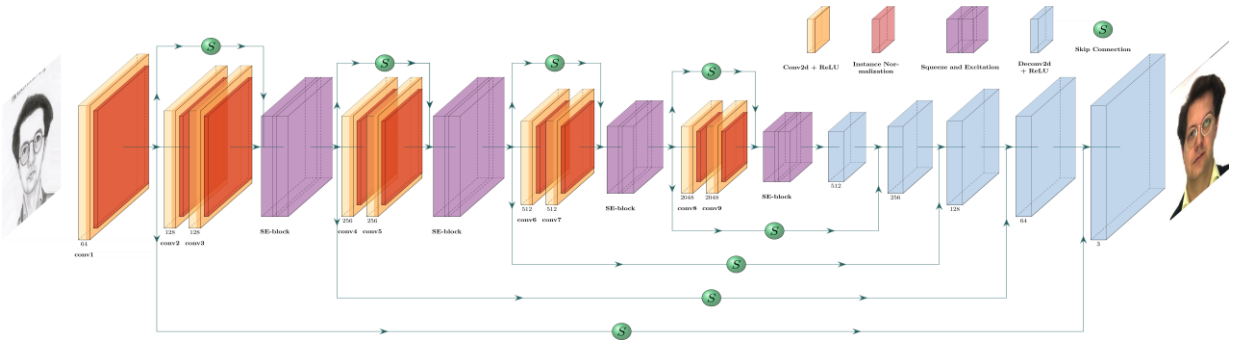
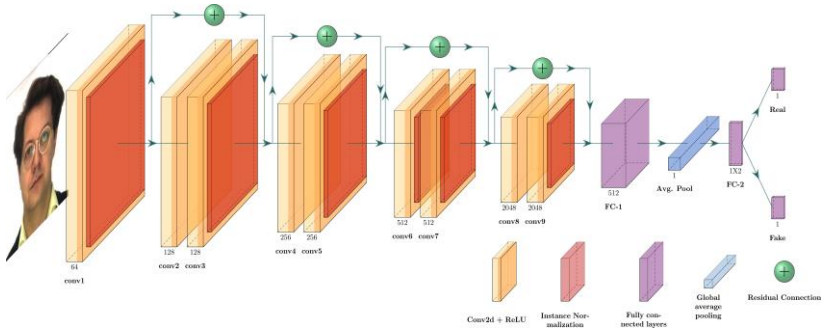


Fig. 2. The architecture of the proposed generator model for sketch to photo image.



Received: August 02, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Fig. 3. The architecture of the proposed discriminator model to tell if giving photo image is real or fake.

TABLE IQUANTITATIVE PERFORMANCE: USING TWO IQA METRICS (SSIM & VIF SCORE) ON CUFS DATASET

Metrics / Case	Case 1 - photo to sketch		Case 2 - sketch to photo	
Model	SSIM (%)	VIF (%)	SSIM (%)	VIF (%)
SpiralNet	54.42	-	-	-
FaceGAN	-	-	65.5	11.5
SU	54.23	-	56.70	-
MOST-Net	54.33	-	-	-
SD-GAN	54.93	-	68.22	-
LLE	52.58	12.64	67.07	17.47
MWF	53.93	12.99	68.15	18.00
cGAN	48.39	10.24	61.39	13.74
cycleGAN	49.63	9.94	62.16	14.79
AdaIN	51.20	9.44	48.85	12.08
MvDT	54.43	13.27	68.31	18.49
Proposed approach	**56.84**	**15.25**	**69.35**	**19.47**

TABLE II

QUANTITATIVE EVALUATION OF THE PROPOSED MODEL SHOWING THE AVERAGE ACCURACY SCORE FOR SKETCH TO PHOTO ON CUFS AND AR DATASET IN PERCENTAGE (%).

Dataset	Rank-1	Rank-3	Rank-5
CUFS	98.2	100	100
AR	97.65%	100	100

Received: August 02, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

- *Feature Similarity Index Metric (FSIM)* [25]: Assesses low-level feature correspondence.
- *Visual Information Fidelity (VIF)* [26]: Measures retained visual information between the generated and real photo.

D. *Baseline Models for Comparison*

The proposed model was benchmarked against existing state-of-the-art methods:

- SpiralNet [27]
- FaceGAN [19]
- SD-GAN [28]
- MvDT [13]

These comparisons assess the effectiveness of the Multiview Learning approach against alternative deep learning techniques.

E. *Experimental Protocol*

Each experiment was repeated five times, and average scores were reported to ensure robustness. The performance results and analysis are discussed in Section IV.

IV. Results and Discussion

This section presents the quantitative and qualitative evaluation of the proposed system. Performance is compared against state-of-the-art methods, including data-driven approaches (LLE [29], MWF [30], Simple Union [31]) and model-driven approaches (Spiral-Net [27], FaceGAN [32], MOST-Net [33], SD-GAN [28], AdaIN [34], cycleGAN [35], MvDT [13]). Experiments were conducted on the CUFS and AR datasets.

A. *Quantitative Evaluation*

The accuracy of the proposed system was measured over five runs, and average scores were reported. The proposed model achieved:

- *Rank-1 Accuracy*: 98.2% on CUFS, 97.65% on AR.
- *Rank-3 and Rank-5 Accuracy*: 100% on both datasets. Table II presents the quantitative evaluation results.

B. *Qualitative Evaluation*

Qualitative assessment was conducted using three Image Quality Assessment (IQA) metrics:

- Structural Similarity Index Metric (SSIM) [24]: Evaluates high-level perceptual similarity.
- Feature Similarity Index Metric (FSIM) [25]: Assesses low-level feature correspondence.
- Visual Information Fidelity (VIF) [26]: Measures retained visual information.

Received: August 02, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Table I reports the qualitative SSIM scores for CUFS, demonstrating that the proposed system outperforms existing methods for both Sketch-to-Photo and Photo-to-Sketch translation. While previous studies did not consider FSIM, this work provides FSIM results as a baseline for future research, emphasizing its relevance due to its closer alignment with the Human Visual System (HVS).

The proposed system achieved:

- *FSIM Scores on CUFS*: 61.34% (Sketches generated from Photos), 58.27% (Photos generated from Sketches).
- *FSIM Scores on AR*: 57.15% (Sketches generated from Photos), 56.71% (Photos generated from Sketches).

C. *Comparison with Baseline Models*

To assess model effectiveness, the proposed system was compared against state-of-the-art models (e.g., SpiralNet, FaceGAN, SD-GAN, MvDT). The proposed approach pro- duces sharper, more realistic images, preserving fine facial details better than alternative methods.

Discussion on Multiview Learning and Feature Fusion

The effectiveness of the proposed system is driven by its ability to extract and integrate multiple feature representations from sketches and photos. Traditional single-view methods often suffer from information loss, as they rely on a single feature descriptor, limiting their ability to capture the com- plex variations in sketches. In contrast, Multiview Learning ensures that diverse facial attributes are considered, such as structure, texture, and key points, leading to a more robust and discriminative representation.

1) *Multiview Representation Strengths:* Each feature descriptor contributes uniquely to the feature learning process:

- HOG: Captures structural and gradient-based information, focusing on facial contours.
- ORB: Extracts key-point descriptors that remain invariant to illumination and rotation.
- LBP: Encodes fine-textured details, improving noise resistance.
- DoG: Enhances contrast variations, making feature ex- traction more robust.

D. *Feature Fusion with ANFIS:*

The challenge with Multiview Learning is ensuring that the extracted views contribute collaboratively without introducing conflicts. Instead of using a predefined rule-based Fuzzy Inference System (FIS), this study adopts an ANFIS, which learns the optimal combination of diverse features. The introduction of ANFIS allows the system to refine feature importance in real time, adapting to variations in sketches and improving decision-making accuracy. The learned rules ensure that dominant features receive higher contributions, while redundant or conflicting features are down-weighted, improving robustness in challenging cases.

E. *Limitations and Future Directions*

Although the proposed model achieves high accuracy, certain challenges remain.

Received: August 02, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

- Small Datase00t Size: While CUFS and AR datasets contain diverse face variations, real-world forensic sketches are often more distorted. Future work will explore domain adaptation techniques.
- Computational Complexity: The GAN-based approach is computationally intensive. Optimization techniques such as lightweight neural architectures will be investigated.
- Potential Biases: Ethnic diversity in forensic datasets must be expanded to ensure fairness across demographic groups.

Despite these limitations, the proposed approach sets a strong baseline for future work in Multiview Learning and Sketch-to-Photo Matching. The next section presents the conclusion and future work.

V. Conclusion and Future Work

In conclusion, this paper introduces a Multiview Learning framework for forensic sketch-to-photo matching, incorporating an ANFIS feature fusion. The approach involves three main phases:

(1) a GAN-based transformation that generates photo-like images from sketches while preserving facial features,(2) Multiview representation using four distinct feature descriptors (HOG, ORB, LBP, DoG) to extract complementary facial attributes, and (3) ANFIS-based feature fusion to adaptively learn optimal feature weighting. Experimental results on the CUFS and AR datasets show 98.2% Rank-1 accuracy on CUFS, 97.65% on AR, and 100% accuracy in Rank-3 and Rank-5 retrieval. The proposed system also outperforms state-of-the-art methods in SSIM, FSIM, and VIF metrics, demonstrating its qualitative effectiveness.

For future work, several avenues warrant exploration. First, the system will be expanded to handle real-world forensic sketches, which typically exhibit more distortions than those in datasets. Second, the computational cost of the GAN-based transformation will be addressed by investigating lightweight neural architectures and knowledge distillation. Third, testing on larger, more diverse datasets is essential for ensuring broader real-world applicability. Additionally, combining ANFIS with deep metric learning could enhance retrieval robustness, and addressing bias and fairness in face matching will be prioritized by expanding training data and implementing fairness-aware learning strategies.

Declaration

I declare there is no conflict of interest regarding this manuscript.

Abbreviation Table	
ANFIS	***Adaptive Neuro-Fuzzy Inference System***
GAN	***Generative Adversarial Network***
HOG	***Histogram of Oriented Gradients***
DOG	***Difference of Gaussian***

Received: August 02, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

ORB	***Oriented FAST and Rotated BRIEF***
LBP	***Local Binary Patterns***

REFERENCES

[1] B. Klare, Z. Li, and A. K. Jain, "Matching forensic sketches to mug shot photos," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 3, pp. 639–646, 2010.

[2] K. B. A. Hassen, J. J. Machado, and J. M. R. Tavares, "Convolutional neural networks and heuristic methods for crowd counting: A systematic review," *Sensors*, vol. 22, no. 14, p. 5286, 2022.

[3] N. Wang, D. Tao, X. Gao, X. Li, and J. Li, "A comprehensive survey to face hallucination," *International journal of computer vision*, vol. 106, no. 1, pp. 9–30, 2014.

[4] X. Wang and X. Tang, "Face photo-sketch synthesis and recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 11, pp. 1955–1967, 2008.

[5] X. Tang and X. Wang, "Face sketch synthesis and recognition," *Proceed- ings Ninth IEEE International Conference on Computer Vision*, vol. 31, no. 11, pp. 687–694, 2003.

[6] X. Gao, N. Wang, D. Tao, and X. Li, "Face sketch–photo synthesis and retrieval using sparse representation," *IEEE Transactions on circuits and systems for video technology*, vol. 22, no. 8, pp. 1213–1226, 2012.

[7] S. Dalal, V. P. Vishwakarma, and S. Kumar, "Feature-based sketch-photo matching for face recognition," *Procedia Computer Science*, vol. 167, pp. 562–570, 2020.

[8] S. Setumin and S. A. Suandi, "Difference of gaussian oriented gradient histogram for face sketch to photo matching," *IEEE Access*, vol. 6, pp. 39 344–39 352, 2018.

[9] S. Setumin, M. F. C. Aminudin, and S. A. Suandi, "Canonical correlation analysis feature fusion with patch of interest: A dynamic local feature matching for face sketch image retrieval," *IEEE Access*, vol. 8, pp. 137 342–137 355, 2020.

[10] M. Kan, S. Shan, H. Zhang, S. Lao, and X. Chen, "Multi-view dis- criminant analysis," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 1, pp. 188–194, 2015.

[11] X. You, J. Xu, W. Yuan, X.-Y. Jing, D. Tao, and T. Zhang, "Multi-view common component discriminant analysis for cross-view classification," *Pattern Recognition*, vol. 92, pp. 37–51, 2019.

[12] T. Chugh, M. Singh, S. Nagpal, R. Singh, and M. Vatsa, "Transfer learn- ing based evolutionary algorithm for composite face sketch recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 117–125, 2017.

[13] C. Peng, N. Wang, J. Li, and X. Gao, "Universal face photo-sketch style transfer via multiview domain translation," *IEEE Transactions on Image Processing*, vol. 29, pp. 8519–8534, 2020.

Received: August 02, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[14] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, vol. 1, pp. 886–893, 2005.

[15] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "Orb: An efficient alternative to sift or surf," *2011 International conference on computer vision*, pp. 2564–2571, 2011.

[16] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[17] D. G. Lowe, "Object recognition from local scale-invariant features," *Proceedings of the seventh IEEE international conference on computer vision*, vol. 2, pp. 1150–1157, 1999.

[18] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, pp. 2672– 2680, 2014.

[19] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image trans- lation with conditional adversarial networks," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125–1134, 2017.

[20] A. M. Martinez, "The ar face database," *CVC Technical Report24*, 1998.

[21] L. A. Gatys, A. S. Ecker, and M. Bethge, "A neural algorithm of artistic style," *arXiv preprint arXiv:1508.06576*, 2015.

[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[23] I. D. Mastan and S. Raman, "Deepcfl: Deep contextual features learning from a single image," *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 2897–2906, 2021.

[24] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[25] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "Fsim: A feature similarity index for image quality assessment," *IEEE transactions on Image Processing*, vol. 20, no. 8, pp. 2378–2386, 2011.

[26] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 430–444, 2006.

[27] I. Azhar, M. Sharif, M. Raza, M. A. Khan, and H.-S. Yong, "A decision support system for face sketch synthesis using deep learning and artificial intelligence," *Sensors*, vol. 21, no. 24, p. 8178, 2021.

[28] X. Qi, M. Sun, Q. Li, and C. Shan, "Biphasic face photo-sketch synthesis via semantic-driven generative adversarial network with graph representation learning," *arXiv preprint*

Received: August 02, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

arXiv:2201.01592, 2022.

[29] Q. Liu, X. Tang, H. Jin, H. Lu, and S. Ma, "A nonlinear approach for face sketch synthesis and recognition," *2005 IEEE Computer Society conference on computer vision and pattern recognition (CVPR'05)*, vol. 1, pp. 1005–1010, 2005.

[30] H. Zhou, Z. Kuang, and K.-Y. K. Wong, "Markov weight fields for face sketch synthesis," *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1091–1097, 2012.

[31] I. Azhar, M. Raza, M. Sharif, S. Kadry, and S. Rho, "Union is strength: Improving face sketch synthesis by fusing outcomes of fully-convolutional-networks and random sampling locality constraint," *Alexandria Engineering Journal*, vol. 61, no. 12, pp. 10 727–10 741,2022.

[32] N. M. Farid, M. S. Fard, and A. Nickabadi, "Face sketch to photo translation using generative adversarial networks," *arXiv preprint arXiv:2110.12290*, 2021.

[33] F. Ji, M. Sun, X. Qi, Q. Li, and Z. Sun, "Most-net: A memory oriented style transfer network for face sketch synthesis," *arXiv preprint arXiv:2202.03596*, 2022.

[34] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2414–2423, 2016.

[35] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," *Proceedings of the IEEE international conference on computer vision*, pp. 2223–2232, 2017.

Received: August 02, 2025