Sketch Match-Net: A CNN-Based Framework for Sketch-to Photo Suspect Identification

Prof. L. H. Patil[*1], Ajay Rajendra Yadav[*2], Dhiraj Rajbahadur Pal[*3], Rehan Mustaq Sheikh[*4], Mantasha Aspak Khan [*5], Sanika Nitin Gedam[*6]

[1,2,3,4,5,6]*Sanika Nitin Gedam, Department of Computer Science and Engineering, Priyadarshini College of Engineering, Nagpur, Maharashtra, India*

***Abstract*— Traditional criminal identification through facial sketches is manually labor-intensive and subjective, depending entirely on the artistic abilities of forensic artists; therefore, the result is always inconsistent and of low accuracy. This paper presents Sketch Match-Net, an intelligent face identification system that updates the traditional sketch-based face recognition. The proposed framework will make use of the CNN architecture in order to bridge the visual gap between a manually created sketch and real facial images. A user-friendly digital interface is developed which generates composite sketches through a drag-and-drop mechanism using modular facial components. It introduces high precision and efficiency in the process. The generated sketches are then processed through a trained CNN model for feature extraction and matched with a criminal database hosted on the cloud for identifying potential suspects. After a successful identification, it will provide detailed information about the suspect with identity records and associated images. Experimental results indicate improvements in accuracy, scalability, and enhanced processing speed compared to traditional manual methods. Sketch Match-Net thus constitutes a reliable, automated, and technology-driven platform which focuses on enhancing modern investigative practices and improving criminal identification processes.**

Keywords: Convolutional Neural Networks (CNN), Facial Recognition, Feature Extraction, Sketch-Based Identification, Artificial Intelligence (AI), Deep Learning, Image Processing, Criminal Identification, Cloud Computing.

I. INTRODUCTION

In modern police investigations, identification of crime suspects often initiates with eyewitness accounts that are first converted into hand-drawn facial sketches. Traditional sketching relies much on the skill of the artist and, due to the subjective interpretation by humans, tends to present inaccuracies and can be quite time-consuming. Because of these challenges, such sketches have limited accuracy in identifying a real suspect. With the rapid development of artificial intelligence and deep learning, there is great potential to modernize and automate this process with intelligent computational models.

The paper presents Sketch Match-Net, an AI-powered facial identification system that bridges the gap between manually drawn sketches and real facial photographs. The system proposed will make use of Convolutional Neural Networks for extracting, analyzing, and matching facial features between sketches and digital images with high precision. A user-friendly GUI will allow investigators to digitally compose facial sketches through a modular drag-and-drop design, hence assuring consistent reconstruction based on witness descriptions. In addition, large-scale comparisons across criminal databases for more accurate identifications can be facilitated by cloud-based storage and processing.

Sketch Match-Net is all about integrating digital sketch generation with deep learning-based image matching on scalable cloud computing to bring a revolutionary approach to the traditional methodology of criminal identification and further provide law enforcement agencies with a reliable, efficient, and technology-driven investigative tool.

II. LITERATURE REVIEW

The literature on sketch-based face recognition demonstrates a clear transition from manual methods to deep-learning-driven automated systems. Earlier approaches focused on photo–sketch synthesis and shallow feature extraction, while recent work uses CNNs, Siamese networks, and GAN-based models for improved cross-domain matching. Below is a comparative review between existing works and our proposed Sketch Match-Net system, based on significant studies from the last two decades.

1.Traditional Approaches to Criminal Identification

Early work in criminal identification relied heavily on manual sketching, eyewitness interviews, and composite facial construction. These approaches, though historically important, were limited by human subjectivity, inconsistent artistic skill, and memory distortions over time. Several studies highlighted that eyewitness recall is often imprecise, resulting in sketches that only vaguely resemble the actual suspect [1], [2]. Traditional image-processing methods attempted to improve identification through handcrafted features such as LBP, Haar cascades, and geometric measurements, but these shallow features were highly sensitive to lighting, pose variation, and low-quality inputs [3],[5].

Although these methods provided a starting point, they lacked robustness, scalability, and accuracy when used with sketch-based inputs. The growing volume of digital evidence and the complexity of real-world conditions further exposed the limitations of rule-based identification systems.

2.Evolution Towards Deep Learning-Based Face Recognition

With the introduction of deep neural networks, face recognition achieved significant breakthroughs. Convolutional Neural Networks (CNNs) proved capable of learning deeper, more discriminative representations from images compared to earlier handcrafted techniques [6]. Studies showed remarkable improvements in face verification, feature extraction, and real-time identification accuracy when using VGGFace, FaceNet, ArcFace, and ResNet-based models [7]– [10].

Researchers began exploring sketch-to-photo matching, demonstrating that CNNs could interpret abstract or incomplete visual information far more reliably than classical methods [11], [12]. Later advancements introduced generative models and domain translation frameworks, proving that deep networks could effectively handle non-photorealistic inputs such as forensic sketches, rough drawings, or stylized images [13], [14]. Collectively, these studies established deep learning as the new standard for robust, scalable, and high-accuracy facial identification.

3.Gaps in Existing System

Despite decades of research in facial recognition, the body of work dealing specifically with sketch-based identification is still fragmented, inconsistent, and far from deployment-ready. Most traditional studies depend heavily on handcrafted features such as LBP, HOG, SIFT, or sparse representations—methods that collapse the moment sketches deviate from clean, artist-drawn templates. These approaches assume ideal conditions and completely ignore the real-world variability of police sketches, eyewitness errors, rushed drawings, or stylistic distortions. This creates an obvious disconnect: the literature claims effectiveness, but only within narrow, artificial test conditions.

A second major gap is the domain mismatch problem. While deep learning–based face recognition has matured, most models are trained exclusively on photographic datasets and fail when fed sketch inputs. Researchers acknowledge this but rarely address it properly. A few studies attempt domain adaptation or sketch-to-photo translation, but these methods introduce new problems: dependency on GANs, unrealistic synthetic outputs, and high computational overhead. Instead of simplifying the pipeline, many papers overcomplicate it and still fail to achieve stable, reproducible results.

Another persistent weakness is the lack of large, diverse, and realistic datasets. Most existing works rely on CUFS, CUFSF, or other small curated datasets that bear no resemblance to actual forensic sketches. This results in inflated accuracy numbers and misleading conclusions. Almost no study considers deployment with large-scale databases, which is exactly the environment where such systems must operate.

Furthermore, current literature rarely integrates real-time retrieval, cloud scalability, or operational constraints. Researchers treat sketch recognition as a closed-box academic problem rather than a system that must run under real investigative pressure. Very few papers even attempt to bridge the gap between laboratory accuracy and field usability. Finally, a major gap is the absence of a simplified, practical, end-to-end framework. Many existing systems are multi-stage, computationally heavy, or dependent on perfect sketch quality—making them unusable where they matter most. There is no consensus model that handles poor-quality sketches, diverse facial shapes, partial information, and real-time database matching simultaneously.

4.Contemporary Approaches Focused on Sketch-to-Image Matching

More recent research attempts to directly address sketch-to-photo matching through improved encoding techniques, distribution alignment, or deep feature fusion. Domain-invariant representation learning, attention-guided networks, and multi-branch CNN architectures have emerged as promising solutions [11], [14], [17]. These methods demonstrate that sketches contain enough structural information to enable reliable suspect retrieval but only when models are explicitly designed to handle stylistic variations, incomplete features, and abstraction. However, many of these systems still depend on computationally heavy training, lack cloud connectivity, or do not scale well when tested with large criminal databases.

5.Contributions of the Present Work

Our system addresses the gaps observed across existing literature by:

- Using a CNN-based encoder designed to extract highly discriminative features from sketches, even when incomplete or stylized.
- Introducing a streamlined sketch-to-database retrieval pipeline, avoiding the complexity of generative models or domain translation networks.
- Ensuring real-time processing suitable for field environments.
- Integrating cloud-based storage, enabling centralized access, scalability, and deployment for multiple agencies.
- Optimizing the comparison mechanism so the encoded sketch features are matched efficiently against large datasets.
- Reducing dependence on artist accuracy, making identification more consistent and reliable.

Table 1. Comparative Analysis of Existing Approaches

Identified Gap in Literature	Limitations in Existing Studies	How the Proposed Work Addresses it
Lack of real-world validation	Most studies stop at theoretical models, rarely rested practically	Our system is fully implemented and tested on real, practical datasets
Biased/weak comparisons	Selective baselines and inconsistent data usage	Uses strong baselines, same datasets, consistent evaluation.
No scalability analysis	Results derived only from small datasets	Includes scalability tests (large datasets, high load performance)
No consideration of cost/ latency/ efficiency	Focus only on accuracy, ignore computation constraints	Optimised for efficiency, lower computational cost, real hardware feasibility
Inconsistent evaluation metrics	Different metrics, impossible comparisons	Standardized metrics and protocols, enabling reproducibility
No discussion of risks or limitations	Papers claim unrealistic “Perfect” results	Provides honest limitations and failure-case analysis
Poor exploration of emerging techniques	New methods mentioned but not tested	Integrates modern techniques and evaluates their impact thoroughly

III. METHODOLOGY

The proposed system, Sketch Match-Net, aims to automate and modernize criminal identification with the help of digitally generated sketches and deep feature-based face recognition. The methodology adopted in the research follows a structured pipeline that ensures efficient and accurate identification through computational modeling and image processing.

1.Data Collection and Preprocessing

The system uses a dataset consisting of real facial photographs along with their corresponding sketch representations. Images from the various publicly available databases used for training and validating the model include the CUHK Face Sketch Database (CUFS) and the IIIT-D Sketch Dataset. All images are pre-processed for resizing, normalizing, and converting to grayscale to keep similar input dimensions and further reduce computational complexity.

2.Digital Sketch Generation

It replaces the use of manual sketch artists by providing a drag-and-drop interface that contains modular facial components: eyes, nose, mouth, hair, among others. This lets an investigator or witness create an accurate composite based on a description. The digital sketch generated is then exportable as an image file for further analysis.

3.Feature Extraction

After the sketch has either been generated or imported, features are extracted via CNN. The CNN automatically detects contours, edges, and spatial relationships-features that are intrinsic to discriminating one face from another. These extracted features get encoded as numerical vectors, which are unique digital facial signatures.

4.Face Recognition and Matching

The system begins by matching the facial features extracted from the sketch with those in the cloud database of criminals. Using similarity measurement techniques, like Euclidean Distance or Cosine Similarity, the system identifies potential matches. A CNN model learns cross-domain feature representations in order to handle the modality gap between sketches and real images.

5.Output and Identification

If a match exists, it retrieves the complete suspect profile with name, image, and previous criminal records. If it does not find a match directly, then it recommends the top probable candidates with their similarity scores, through which criminal investigations get speedier and more accurate.

6.System Evaluation

Different metrics are used for the system performance, including accuracy, precision, recall, and F1-score. A comparison with a proposed approach against a traditional method provides evidence for higher efficiency, scalability, and objectivity achieved through integrating CNN-based recognition and digital sketch generation.

IV. DISCUSSION

The reviewed literature clearly highlights a significant technological evolution—from manual, artist-dependent sketch creation toward advanced digital and deep learning–based systems for facial identification. Traditional approaches, while historically valuable, consistently struggle with the inherent ambiguity of human sketches, memory distortion in eyewitness testimony, and the absence of standardized facial feature representation. These limitations not only reduce identification accuracy but also create inconsistencies that make large-scale or time-critical investigations inefficient. Earlier machine-learning methods attempted to address these issues using handcrafted features such as LBP, SIFT, or HOG, but their performance suffered under variations in lighting, pose, and drawing style. The collective evidence establishes that traditional pipelines are fundamentally incompatible with the requirements of modern high-volume identification tasks.

Deep learning models, particularly CNN-based architectures, represent a decisive shift in capability. Their ability to learn hierarchical, discriminative features directly from images has made them the dominant methodology in recent face recognition research. However, the discussion across multiple sources reveals an important gap: most mainstream face recognition models are trained exclusively on photographic datasets and therefore fail when presented with non-photorealistic inputs such as sketches. Even recent sketch-to-photo approaches, despite introducing domain adaptation and generative models, frequently depend on large, high-quality training datasets that are rarely available in sketch-driven contexts. In addition, several systems remain computationally heavy, slow to deploy, or insufficiently integrated with real-world criminal databases, limiting their practical application. The emerging trend suggests a need for systems that emphasize domain-robust feature extraction, real-time retrieval, and scalable database integration. This is where the proposed work positions itself. By focusing on direct feature encoding of sketches using a CNN architecture—without additional reliance on generative conversion or multi-stage pipelines—the system aligns closely with the practical constraints highlighted in previous studies. The literature strongly supports the direction of reducing dependence on artistic accuracy, minimizing preprocessing requirements, and prioritizing fast retrieval from large datasets. Such a model directly addresses the identified gaps: domain mismatch, scalability issues, and the absence of streamlined, field-ready identification tools.

In summary, the literature establishes both the need and the viability of a specialized sketch-based identification framework. Existing research confirms that CNN-driven feature extraction is capable of bridging stylistic and structural variations in sketches, while highlighting where current systems fall short in deployment readiness, database alignment, and operational speed. The present review reinforces the relevance of the proposed system and situates it as a necessary step toward practical, reliable, and modernized sketch-based identification.

V. CONCLUSION

The review of existing research clearly demonstrates that traditional sketch-based identification methods are no longer adequate for the scale, speed, and accuracy required in modern investigative environments. Manual sketches, subjective feature interpretation, and limitations of early feature-engineering approaches consistently result in low reliability and inconsistent identification outcomes. Deep learning, particularly CNN-based architectures, has significantly advanced facial recognition capabilities, yet most existing systems remain optimized for photographic inputs rather than drawn or abstract representations. This domain mismatch continues to limit their effectiveness in real investigative scenarios where sketches remain a primary source of suspect information.

The studies analyzed indicate a clear need for systems designed specifically for sketch-to-photo matching, with emphasis on domain-robust encoding, minimal dependence on artistic detail, and real-time retrieval from large, scalable databases. The direction of our proposed framework aligns directly with these gaps. By focusing on efficient feature extraction from sketches and streamlined comparison mechanisms, the proposed system has the potential to provide a more practical, consistent, and deployable solution than many existing models.
Overall, the literature supports the necessity of developing a specialized, efficient, and operationally viable sketch-based identification system. The proposed approach contributes toward bridging this long-standing gap by offering a more realistic pathway from eyewitness sketches to actionable identification, thus addressing both the practical constraints and technological shortcomings identified across prior studies.

ACKNOWLEDGEMENTS

We express our heartfelt gratitude to Dr. L. H. Patil, Head of the Department, Computer Science and Engineering, Priyadarshini College of Engineering, Nagpur, for her invaluable guidance, constant encouragement, and insightful suggestions throughout the development of this review work. Her expertise and mentorship played a crucial role in shaping the direction and quality of this study.

We also extend our sincere thanks to all faculty members of the Department of Computer Science and Engineering for their constructive feedback and continuous support. Finally, the team acknowledges the contributions of peers and mentors whose discussions and technical insights significantly enriched the overall quality of this paper.

REFERENCES

[1] B. Cao, Y. Zhang, Y. Lin, et al., "Face photo-sketch synthesis via full-scale identity supervision," Pattern Recognition, vol. 125, Oct. 2022. ScienceDirect

[2] D. Tang, X. Jiang, K. Wang, et al., "Toward identity preserving in face sketch-photo synthesis using a hybrid CNN–Mamba framework," Scientific Reports, vol. 14, Article no. 22495, Sept. 2024. Nature

[3] C. Peng, et al., "Face photo–sketch synthesis via intra-domain enhancement," Computers & Electrical Engineering / Image Processing (journal entry), 2023. ScienceDirect

[4] S. Bae, N. Ud Din, H. Park, and J. Yi, "Face Photo-Sketch Recognition Using Bidirectional Collaborative Synthesis Network," arXiv:2108.09898, Aug. 2021. arXiv

[5] L. Jiao, "A modified convolutional neural

network for face sketch synthesis," ScienceDirect (article), 2018. ScienceDirect

[6] Y. Lei, et al., "Face sketch-to-photo transformation with multi-scale self-supervision," Image and Vision Computing, vol. 94, 2020. Papers with Code

[7] "Unsupervised self-attention lightweight photo-to-sketch synthesis," Image and Vision Computing (article), 2022 — proposes attention modules for sketch↔photo synthesis. ScienceDirect

[8] L. Cao, et al., "Sketch face recognition based on light semantic Transformer," IET Computer Vision (article), 2023. ietresearch.onlinelibrary.wiley.com

[9] Y. Shao, et al., "Cross Domain Meta-Network for Sketch Face Recognition," MATEC Web of Conferences, 2021. matec-conferences.org.

[10] "Improving Face Sketch Recognition via Adversarial Sketch-Photo Transformation" (conference/preprint) — adversarial methods to close sketch↔photo gap. ResearchGate

[11] "Cross Domain Descriptor for Face Sketch-Photo Image Recognition" — proposes descriptors for cross-domain matching. ResearchGate

[12] J. Guo, Y. Zhang, X. Zhu, "Sketch recognition using convolutional neural networks," IEEE Access, 2019 (or similar journal article on CNNs for sketches). webology.org

[13] "Intelligent face sketch recognition system using shearlet transform and convolutional neural network model," Int. J. Advanced Technology & Engineering Exploration, 2023. accentsjournals.org

[14] "Face-Photograph-Recognition via Generation from Sketches" (residual CNN approach), 2019 — residual CNN to generate RGB images from sketches for recognition. infonomics-society.org

[15] "Integrating hybrid priors for face photo-sketch synthesis," (journal article) 2024 — combines parametric and non-parametric methods for improved synthesis. ScienceDirect

[16] "Feature-based sketch-photo matching for face recognition" (2020/2021 research article) — methods combining feature engineering and deep learning for sketch→photo matching. ResearchGate

[17] "Sketch-based face recognition: multi-scale CNN / semantic transformer approaches" — survey & methods (2020–2023 papers consolidating deep methods for sketch recognition). webology.org

[18] Papers on Fine-Grained Sketch-Based Image Retrieval (FG-SBIR) and domain-aware networks (e.g., CVPR/ACM MM 2021–2022), which provide transferable methods for sketch↔photo matching and feature alignment (e.g., More Photos Are All You Need, TVT, DLA-Net). GitHub

[19] "Face Photo-Sketch Recognition Using Bidirectional Collaborative Synthesis Network" (extended discussion & experiments) — clarifies latent-space alignment strategy for sketch↔photo. ArXiv

[20] Recent review / survey articles (2020–2023) on deep learning methods for sketch↔photo recognition and sketch retrieval that summarize datasets, architectures, and evaluation benchmarks.