Computational Engineering and Technology Innovations

www.ceti.reapress.com

Comput. Eng. Technol. Innov. Vol. 2, No. 1 (2025) 11–22.

Paper Type: Original Article

AI-Powered Forensic Face Drawing: An Overview of the Research on Methodological, Theoretical, and Real-World Defects in Suspect Identification Systems

Kannan Nagarajan[1,*], Jeeva Sakthivel[1], Rosenpranav Kogilavani Satheeshkumar[1], Mohan Raja Velusamy[1], Kamalaveni Vanjigounder[1]

[1] Department of Artificial Intelligence and Data Science, Sri Shakthi Institute of Engineering and Technology, Coimbatore-641062, Anna University, Chennai; kannannagarajan2004@gmail.com, imjeevasakthi12@gmail.com, rosenpranav1010@gmail.com, mohanraja2545@gmail.com, vvkamalaveni@gmail.com.

Citation:

Received: 07 September 2024 *Revised: 15 November 2024* *Accepted: 28 December 2024*	Kannan, N., Jeeva S., Rosenpranav, K. S., Mohan Raja V., Kamalaveni V. (2025). AI-powered forensic face drawing: An overview of the research on methodological, theoretical, and real-world defects in suspect identification systems. *Computational engineering and technology innovations*, *2*(1), 11-22.

Abstract

Artificial Intelligence (AI) has transformed forensic facial sketching, introducing advanced deep learning architectures for suspect identification in constrained-data environments. This literature survey systematically analyzes the state of the art in AI-driven forensic facial sketching, identifying critical gaps across methodological, theoretical, and practical dimensions. Methodologically, we highlight the lack of comparative studies across deep learning architectures (e.g., GANs, VAEs, diffusion models), the over-reliance on accuracy as the sole evaluation metric, and the insufficient investigation of algorithmic robustness to noise and distortions. Theoretically, we identify gaps in understanding how AI models interpret and reconstruct facial features from sparse witness descriptions, as well as in the limited research on inference mechanisms with incomplete data. Practically, we note deficiencies in real-world scenario testing, user-centric design for forensic practitioners, and system scalability for operational deployment. By synthesizing existing literature, this survey not only identifies these interconnected gaps but also proposes future research directions to develop more robust, efficient, and forensically applicable AI systems. Our analysis emphasizes the need for standardized benchmarks, comprehensive evaluation protocols, and interdisciplinary collaboration to advance the field.

Keywords: Forensic facial sketching, Deep learning architectures, Generative adversarial networks, Suspect identification, Evaluation metrics.

1 | Introduction

Forensic facial sketching represents a critical component of suspect identification systems, bridging the gap between eyewitness descriptions and the generation of recognizable facial representations [1]. The evolution

Corresponding Author: kannannagarajan2004@gmail.com

https://doi.org/10.48314/ceti.vi.54

from traditional artist-rendered sketches to AI-driven computational systems has opened new possibilities for law enforcement agencies worldwide [2]. These systems aim to transform verbal descriptions from eyewitnesses or victims into facial representations that can aid in identifying suspects, with recent advances in deep learning, particularly Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and diffusion models, significantly enhancing the photorealism and accuracy of computer-generated facial composites [3].

Despite these technological advancements, significant research gaps persist across methodological, theoretical, and practical dimensions, hindering the full potential of AI-driven forensic facial sketching systems [4]. These gaps limit the adoption of these systems in real-world forensic settings and impede the development of more effective suspect identification tools. Understanding these gaps is essential for directing future research efforts and developing systems that are not only technically sophisticated but also forensically valuable and practically implementable.

This literature survey provides a comprehensive analysis of the current research landscape in AI-driven forensic facial sketching, with particular focus on identifying critical gaps across methodological, theoretical, and practical domains. By systematically examining these gaps, this survey aims to establish a foundation for future research directions that could address these limitations and advance the state of the art in forensic facial sketching technology for suspect identification.

2 | Methodology

This literature survey employed a structured approach to identify and analyze relevant research on AI-driven forensic facial sketching systems. While not a systematic review following PRISMA guidelines, our methodology incorporated systematic elements to ensure comprehensive coverage and critical analysis of the literature.

2.1 | Search Strategy

We conducted a comprehensive search across multiple academic databases, including IEEE Xplore, ACM Digital Library, SpringerLink, ScienceDirect, and Google Scholar. The search was performed using combinations of keywords and phrases such as Artificial Intelligence (AI), forensic facial sketching, deep learning facial composite, GAN suspect identification, automated facial sketching, machine learning forensic art, and AI eyewitness composite. The search period covered publications from 2010 to 2024 to capture the evolution of the field from early machine learning approaches to current deep learning techniques.

2.2 | Inclusion and Exclusion Criteria

Studies were included based on the following criteria:

I. Focus on AI-driven approaches to forensic facial sketching or composite generation.

II. Published in peer-reviewed journals, conference proceedings, or academic books.

III. Written in English.

IV. Contained technical details about algorithms, architectures, or evaluation methodologies.

Exclusion criteria included:

I. Studies focusing solely on traditional artist-rendered sketches without AI components.

II. Non-technical articles or news reports.

III. Studies not directly related to forensic applications (E.g., general face generation without forensic context).

IV. Duplicate publications or preliminary versions of already included papers.

2.3 | Selection Process

The initial search yielded 327 publications. After removing duplicates (N = 42), 285 publications remained for title and abstract screening. Two reviewers independently screened these publications, with disagreements resolved through discussion. This process resulted in 156 publications for full-text review. Following detailed examination, 89 publications were selected for inclusion in this survey based on their relevance, technical contribution, and methodological rigor.

2.4 | Data Extraction and Analysis

For each included publication, we extracted information on:

I. AI architectures and algorithms employed.

II. Evaluation methodologies and metrics.

III. Dataset characteristics and size.

IV. Reported performance and limitations.

V. Identified research gaps and future directions.

The extracted data were synthesized thematically to identify patterns, trends, and gaps across methodological, theoretical, and practical dimensions. This synthesis serves as the basis for our analysis in the subsequent sections.

3 | Technical Overview of AI-Driven Forensic Facial Sketching Systems

3.1 | Evolution of AI Architectures

AI-driven forensic facial sketching has evolved significantly over the past decade, with several distinct generations of architectures emerging:

Early machine learning approaches (2010-2015): initial systems employed traditional machine learning techniques such as Principal Component Analysis (PCA), Independent Component Analysis (ICA), and Support Vector Machines (SVM) for facial feature extraction and reconstruction [5]. These systems typically operate by assembling pre-defined facial components based on witness descriptions, with limited ability to generate novel facial features.

Deep learning revolution (2015-2018): the introduction of deep neural networks marked a significant advancement, with Convolutional Neural Networks (CNNs) becoming the dominant architecture for feature extraction and representation [6]. These systems demonstrated improved ability to capture complex facial features and their relationships, though they often required large training datasets and struggled with generating high-quality outputs from sparse descriptions.

Generative model era (2018-present): the current generation of systems leverages generative models — particularly GANs, VAEs, and, more recently, diffusion models — to create photorealistic facial composites [7]. These architectures have significantly improved the quality and diversity of generated sketches, enabling more accurate representation of subtle facial features and better handling of incomplete input data.

3.2 | Dominant AI Architectures

GANS: GANs have become the most widely adopted architecture for forensic facial sketching, accounting for approximately 65% of recent publications in the field [8]. The adversarial training process, involving a generator and discriminator network, enables the creation of highly realistic facial images. Key GAN variants applied in forensic contexts include:

I. Conditional GANs (CGANs): allow control over generated outputs by incorporating witness descriptions as conditioning information [9].

II. Progressive GANs: generate high-resolution images through a multi-scale training process [10].

III. Style-based GANs (StyleGAN): enable fine-grained control over facial features and attributes [11].

VAEs: VAEs represent approximately 20% of current systems, offering advantages in terms of training stability and the ability to learn meaningful latent representations of facial features [12]. VAE-based systems typically employ encoder-decoder architectures, with the encoder mapping witness descriptions to a latent space and the decoder generating facial images from this representation.

Diffusion models: Emerging as a promising alternative, they have gained attention in the past two years and now comprise approximately 10% of recent publications [13]. These models generate images through an iterative denoising process, demonstrating superior performance in handling noisy or incomplete input data compared to earlier architectures.

Hybrid approaches: The remaining 5% of systems employ hybrid architectures that combine elements from multiple approaches to leverage their respective strengths [14]. Common combinations include GAN-VAE hybrids for improved training stability and diffusion-GAN hybrids for enhanced image quality.

3.3 | Technical Challenges and Innovations

Data scarcity: A persistent challenge in forensic facial sketching is the limited availability of paired witness descriptions and facial images. Recent innovations include:

I. Few-shot learning techniques: enabling model training with limited examples [15].

II. Data augmentation strategies: synthetic data generation to expand training datasets [16].

III. Transfer learning: leveraging pre-trained models on general facial datasets [17].

Feature control: achieving precise control over individual facial features based on witness descriptions remains challenging. Notable technical solutions include:

I. Attribute-based loss functions: penalizing deviations from specified facial attributes [18].

II. Attention mechanisms: focusing on relevant facial regions during generation [19].

III. Hierarchical generation: building facial composites feature-by-feature according to witness priorities [20].

Evaluation methodologies: the field lacks standardized evaluation protocols, leading to inconsistent performance reporting. Recent technical innovations include:

I. Multi-metric evaluation frameworks: Combining traditional metrics with forensic-specific measures [21]

II. Human-in-the-loop evaluation: Incorporating forensic artist assessments [22].

III. Adversarial evaluation: Testing system robustness against challenging inputs [23].

Table 1. Comparison of AI architectures for forensic facial sketching.

Architecture	Training Stability	Image Quality	Control Precision	Data Efficiency	Computational Cost
Traditional ML	High	Low	Medium	High	Low
CNN-based	Medium	Medium	Medium	Medium	Medium
VAE	High	Medium	Low	Medium	Medium
GAN	Low	High	High	Low	High
Diffusion	Medium	Very high	Medium	Low	Very high
Hybrid	Medium	High	High	Medium	High

4 | Methodological Gaps

4.1 | Inadequate Comparative Studies

The literature reveals a significant paucity of comparative studies evaluating the relative performance of different algorithmic approaches in AI-driven forensic facial sketching. Most existing research focuses on demonstrating the efficacy of specific algorithms in isolation rather than systematically comparing them against alternative approaches [24]. This limitation is particularly evident in the lack of studies that directly compare deep learning algorithms against traditional machine learning methods and conventional artist-rendered techniques under controlled conditions.

Several studies have examined individual algorithmic approaches, such as GAN-based systems [25], VAE implementations [26], or evolutionary algorithms [27]. However, these studies typically evaluate their proposed methods against baseline approaches or earlier versions of the same technology, rather than against fundamentally different algorithmic paradigms. This methodological limitation makes it difficult to determine which approaches are most suitable for specific forensic scenarios or to understand the trade-offs between different techniques in terms of accuracy, computational efficiency, and practical utility.

Table 2 summarizes the current state of comparative research in AI-driven forensic facial sketching, highlighting the scarcity of studies that directly compare different algorithmic approaches.

Table 2. Comparison of algorithmic approaches in AI-driven forensic facial sketching.

Algorithm Type	Number of Studies	Comparative Studies	Common Evaluation Metrics	Key Limitations
GAN-based systems	42	3 (7%)	Accuracy, SSIM, FID	Limited comparison with non-GAN methods
VAE implementations	28	2 (7%)	Accuracy, reconstruction error	Small-scale evaluations only
Evolutionary algorithms	15	1 (7%)	Accuracy, convergence, rate	Outdated benchmarks used
Traditional ML methods	23	4 (17)	Accuracy, precision, recall	Limited modern comparisons
Artist rendered	19	5(26)	Recognition rate, similarity	Subjective evaluation methods

The absence of comprehensive comparative studies has practical implications for forensic practitioners who must select appropriate systems for their specific needs. Without precise comparative data, agencies may adopt systems based on marketing claims rather than empirical evidence of relative performance [28]. Furthermore, the lack of comparative evaluation hinders the identification of best practices and the development of standardized protocols for forensic facial sketching in suspect identification systems.

4.2 | Need for Improved Evaluation Metrics

Current research in AI-driven forensic facial sketching systems predominantly relies on accuracy as the primary metric for evaluating system performance [2]. While accuracy provides valuable information about a system's ability to generate recognizable facial representations, it fails to capture the multifaceted nature of forensic utility. The over-reliance on this single metric limits our understanding of system performance across relevant dimensions in forensic contexts.

A more comprehensive evaluation framework should incorporate additional metrics such as precision, recall, F1 score, and Area Under the Curve (AUC) to provide a more nuanced assessment of system performance [29]. These metrics could help distinguish between different types of errors and provide insights into how systems perform across various demographic groups, facial feature types, and input conditions.

The current approach focuses on a single metric (accuracy), whereas a comprehensive framework would incorporate multiple metrics, including accuracy, precision, recall, F1, AUC, robustness, efficiency, and usability. This multi-dimensional evaluation would provide a more complete assessment of system performance and forensic utility.

Furthermore, the field lacks standardized evaluation protocols that enable meaningful comparisons across studies and systems. The lack of standard benchmarks, datasets, and evaluation criteria makes it difficult to synthesize findings across studies and establish clear performance standards for forensic applications [30]. This methodological gap significantly impedes progress in the field by limiting the ability to build on previous research systematically.

4.3 | Limited Studies on Algorithm Robustness

The robustness of AI-driven forensic facial sketching systems against various forms of noise and distortions in input data represents another significant methodological gap in current research [31]. Eyewitness descriptions are inherently subjective, incomplete, and potentially distorted by factors such as stress, memory decay, and cross-racial identification challenges [32]. However, most existing systems are evaluated under relatively ideal conditions that do not adequately reflect the complexity and variability of real-world forensic scenarios.

Research examining how different algorithms perform under incomplete descriptions, contradictory information, or ambiguous input parameters is notably limited [33]. Similarly, studies investigating the effects of different types of noise in input data—such as imprecise spatial relationships, inconsistent feature descriptions, or temporal inconsistencies in witness accounts—are scarce in the literature.

Table 3 summarizes the types of robustness testing that are currently lacking in the literature and their importance for forensic applications.

Table 3. Robustness testing gaps in AI-driven forensic facial sketching.

Robustness Dimension	Current Research Coverage	Forensic Importance	Key Research Questions
Input noise tolerance	Limited (12% of studies)	High	How do systems perform with ambiguous witness descriptions?
Cross-demographic performance	Minimal (8% of studies)	Critical	Are systems equally effective across ethnicities, ages, and genders?
Temporal consistency	Very limited (5% of studies)	Moderate	Can systems maintain consistency in the face of evolving witness accounts?
Environmental variability	Limited (17% of studies)	High	How do lighting, angle, and image quality affect performance?
Adversarial resistance	Very limited (3% of studies)	Emerging	Can systems withstand intentional attempts to mislead?

This methodological gap has significant implications for the reliability of these systems in practical forensic applications. Without a thorough understanding of algorithm robustness under challenging conditions, practitioners cannot assess the confidence they should place in system outputs or determine appropriate protocols for verifying results [23]. Furthermore, the lack of robustness evaluation limits our understanding of how different algorithms might be combined or adapted to improve performance in real-world suspect identification scenarios.

5 | Theoretical Gaps

5.1 | Insufficient Understanding of the AI-Face Sketching Relationship

Despite significant advances in the technical implementation of AI-driven facial sketching systems, the theoretical understanding of how these systems interpret and recreate human facial features remains limited [34]. Most research focuses on empirical demonstrations of system performance rather than developing theoretical frameworks that explain the underlying principles of facial representation and generation in AI systems.

The literature reveals a particular lack of theoretical understanding regarding how AI systems capture and represent the complex interrelationships between facial features [35]. Human faces exhibit intricate correlations among features—such as the relationships between eye shape, nose structure, and jawline—that contribute to overall facial appearance and recognizability. However, current AI systems often treat facial features as relatively independent components, potentially missing critical holistic aspects of facial representation.

Current AI approaches typically employ feature-based processing, treating facial components (eyes, nose, mouth, ears, hair) independently. However, a theoretical gap exists in understanding holistic facial representation, which would involve modeling interconnected feature networks and contextual factors. This holistic understanding is needed to capture the complex relationships between facial features and their cultural and contextual influences.

Furthermore, there is limited theoretical work examining how different AI architectures—such as GANs, VAEs, and diffusion models—differentially capture and represent facial features and their relationships [36]. This theoretical gap hinders the development of more sophisticated algorithms that could better model the complex nature of human facial appearance and improve the forensic utility of generated sketches for suspect identification.

5.2 | Lack of Research on AI Inference

The process by which AI systems make inferences and predictions based on limited or incomplete data represents another significant theoretical gap in the literature [37]. In forensic facial sketching scenarios, AI systems must often generate complete facial representations from sparse, ambiguous, or potentially contradictory witness descriptions. However, the theoretical foundations of how these systems perform such inference tasks remain poorly understood.

Current research provides limited insight into how different AI architectures handle uncertainty, fill in missing information, or resolve inconsistencies in input data [38]. This theoretical gap is particularly problematic given the high-stakes nature of forensic applications, where the accuracy and reliability of generated sketches can have significant consequences for investigations and legal proceedings.

Furthermore, there is insufficient theoretical understanding of how AI systems balance between fidelity to input descriptions and adherence to anatomical and statistical norms of human facial appearance [39]. This balance is critical in forensic applications, where sketches must both accurately reflect witness descriptions and produce plausible facial representations. Without a stronger theoretical foundation in this area, it is challenging to develop systems that can reliably navigate this fundamental tension in suspect identification contexts.

6 | Practical Application Gaps

6.1 | Limited Research on Real-World Application

The literature reveals a significant gap between laboratory evaluations of AI-driven forensic facial sketching systems and their performance in real-world forensic scenarios [40]. Most existing studies evaluate systems under controlled conditions using standardized datasets and metrics that may not adequately reflect the complexity and variability of actual forensic applications.

Research examining system performance under different lighting conditions, camera angles, and reference image qualities is notably limited [41]. Similarly, studies investigating how these systems perform with witnesses from different demographic backgrounds, cultural contexts, or levels of description ability are scarce in the literature. This gap is particularly significant given the diverse range of conditions under which forensic facial sketching systems must operate in practice for suspect identification.

Table 4 summarizes the key real-world application gaps identified in the literature and their implications for forensic practice.

Table 4. Real-world application gaps and implications.

Application Dimension	Current Research Coverage	Practical Implications	Critical Research Questions
Environmental variability	Limited (18% of studies)	Reduced reliability in field conditions	How do systems perform with poor lighting, angles, and distances?
Cross-demographic performance	Minimal (10% of studies)	Potential bias in suspect identification	Are systems equally effective across ethnicities, ages, and genders?
Longitudinal performance	Very limited (5% of studies)	Unknown sustainability over time	How do systems perform with prolonged use and evolving cases?
Integration with workflows	Limited (15% of studies)	Implementation challenges	How do systems integrate with existing forensic processes?
Legal admissibility	Minimal (7% of studies)	Uncertain standing in court	What standards are needed for legal acceptance?

Furthermore, there is limited research examining the long-term performance and reliability of these systems in operational forensic settings [42]. Most studies focus on short-term technical evaluations rather than longitudinal assessments of how these systems integrate into forensic workflows, adapt to different case types, or evolve as practitioners gain experience with them.

6.2 | Absence of User-Centric Research

A critical gap in the literature is the lack of user-centric research examining the usability and practical utility of AI-driven forensic facial sketching systems from the perspective of forensic artists and law enforcement officials [43]. Most studies focus on technical performance metrics rather than human factors that determine whether and how these systems are actually used in practice.

Research examining how forensic artists interact with AI systems, how they integrate algorithmic outputs into their workflows, or how they balance between automated suggestions and their own expertise is notably limited [44]. Similarly, studies investigating how law enforcement officials interpret and use AI-generated sketches in investigations are scarce in the literature.

This user-centric research gap has significant implications for the design and implementation of AI-driven forensic facial sketching systems. Without a clear understanding of user needs, preferences, and workflows, system developers may create technically sophisticated tools that fail to address practical requirements or fit seamlessly into existing forensic processes [45]. This misalignment between technical capabilities and user needs can significantly limit the adoption and impact of these systems in real-world settings for suspect identification.

6.3 | Scalability and Efficiency

The scalability and efficiency of AI-driven forensic facial sketching systems represent another significant practical gap in the literature [43]. Most research focuses on demonstrating proof-of-concept systems or optimizing performance on relatively small datasets, with limited attention to how these systems would perform at scale in operational forensic environments.

Research examining the computational requirements of different algorithms —such as their processing times for different types of inputs or their resource utilization across different hardware configurations — is notably limited [46]. Similarly, studies investigating how these systems perform when accessing and processing large facial databases—such as those maintained by law enforcement agencies—are scarce in the literature.

Current research focuses on small-scale performance (single case, limited dataset, ideal conditions). In contrast, operational reality demands large-scale requirements (multiple cases, massive database, variable resources, real-time needs, resource constraints, network integration). Key efficiency gaps include processing time (current systems take minutes to hours vs. operational need of seconds to minutes), resource requirements (high-end hardware vs. standard forensic equipment), database integration (limited research on integration with existing law enforcement databases), and network performance (unstudied performance in distributed forensic environments).

This scalability and efficiency gap has significant implications for the practical deployment of AI-driven forensic facial sketching systems. Without a clear understanding of their computational requirements and performance characteristics at scale, forensic agencies cannot effectively plan for the infrastructure, training, and support needed to implement these systems in operational settings [47]. Furthermore, the lack of research on efficiency optimization limits our understanding of how these systems could be adapted for resource-constrained environments or real-time suspect identification applications.

7 | Critical Discussion

The methodological, theoretical, and practical gaps identified in this survey are interconnected and mutually reinforcing. The inadequate comparative studies and over-reliance on limited evaluation metrics (Methodological gaps) contribute to a poor understanding of how different algorithms perform under various conditions, which in turn limits the development of robust theoretical frameworks for AI-face sketching relationships (theoretical gap). Similarly, the lack of user-centric research and real-world application studies (practical gaps) means that theoretical insights are not adequately tested or refined in authentic contexts, perpetuating a cycle where systems may be technically sophisticated but forensically limited.

These gaps hinder the advancement of forensic facial sketching systems. Integrated research that addresses methodological, theoretical, and practical dimensions—such as comparative studies with comprehensive metrics—can yield deeper technical insights and stronger theoretical foundations.

The identified gaps underscore the importance of collaboration among computer scientists, forensic practitioners, cognitive psychologists, and human-computer interaction specialists. Such interdisciplinary efforts can ensure that technological developments are firmly grounded in theoretical understanding and practical needs, resulting in systems that are effective, reliable, and valuable for suspect identification.

8 | Conclusion and Future Research Directions

This literature survey has identified critical research gaps across methodological, theoretical, and practical dimensions in AI-driven forensic facial sketching systems for suspect identification. Addressing these gaps is essential for advancing the field and developing systems that can make meaningful contributions to forensic investigations.

Based on the gaps identified, several promising directions for future research emerge:

I. Comparative studies of various AI and traditional techniques: rigorous comparative evaluations of different algorithmic approaches under standardized conditions would provide valuable insights into their relative strengths, weaknesses, and suitability for different forensic scenarios.

II. Development of new evaluation metrics: creating comprehensive evaluation frameworks that incorporate multiple metrics beyond accuracy would enable more nuanced assessments of system performance and facilitate meaningful comparisons between different approaches.

III. Studies exploring the theoretical basis of the AI-face sketching relationship: theoretical research examining how AI systems interpret and represent facial features and their relationships could inform the development of more sophisticated algorithms that better capture the complex nature of human facial appearance.

IV. Research on the real-world application and usability: user-centered studies examining how these systems perform in actual forensic scenarios and how practitioners interact with them would help ensure that technical advances translate into practical value for suspect identification.

V. Studies focusing on optimization and scalability: research addressing the computational efficiency and scalability of these systems would facilitate their deployment in operational forensic settings and enable their use in resource-constrained environments.

By addressing these research directions, the field can develop AI-driven forensic facial sketching systems that are not only technically sophisticated but also theoretically grounded, forensically valuable, and practically implementable. Such systems have the potential to enhance law enforcement agencies' capabilities significantly and to contribute to more effective and just suspect identification processes.

Competing Interests

The authors declare that they have no competing interests.

Data Availability Statement

This is a literature survey and does not involve original data collection or analysis.

References

[1] Jain, A. K., & Li, S. Z. (2011). *Handbook of face recognition* (Vol. 1). Springer. https://doi.org/10.1007/978-0-85729-932-1

[2] Frowd, C. D., Hancock, P. J. B., & Carson, D. (2004). Evofit: A holistic, evolutionary facial imaging technique for creating composites. *ACM transactions on applied perception (tap)*, *1*(1), 19–39. https://doi.org/10.1145/1008722.1008725

[3] Dash, A., Gamboa, J. C. B., Ahmed, S., Liwicki, M., & Afzal, M. Z. (2017). Tac-gan-text conditioned auxiliary classifier generative adversarial network. *ArXiv preprint arxiv:1703.06412*. https://doi.org/10.48550/arXiv.1703.06412

[4] Frowd, C., Bruce, V., McIntyre, A., & Hancock, P. (2007). The relative importance of external and internal features of facial composites. *British journal of psychology*, *98*(1), 61–77. https://doi.org/10.1348/000712606X104481

[5] Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of cognitive neuroscience*, *3*(1), 71–86. https://doi.org/10.1162/jocn.1991.3.1.71

[6] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386

[7] Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in neural information processing systems*, *33*, 6840–6851. http://dx.doi.org/10.48550/arXiv.2006.11239

[8] Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. *ACM computing surveys (CSUR)*, *54*(1), 1–41. https://doi.org/10.1145/3425780

[9] Mirza, M., & Osindero, S. (2014). Conditional generative adversarial nets. *ArXiv preprint arxiv:1411.1784*. https://doi.org/10.48550/arXiv.1411.1784

[10] Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2017). Progressive growing of gans for improved quality, stability, and variation. *ArXiv preprint arxiv:1710.10196*. https://doi.org/10.48550/arXiv.1710.10196

[11] Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). Analyzing and improving the image quality of stylegan. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 8110-8119). http://dx.doi.org/10.1109/CVPR42600.2020.00813

[12] Kingma, D. P., Welling, M., & others. (2019). An introduction to variational autoencoders. *Foundations and trends®in machine learning*, *12*(4), 307–392. http://dx.doi.org/10.1561/2200000056

[13] Yang, L., Zhang, Z., Song, Y., Hong, S., Xu, R., Zhao, Y., … & Yang, M. H. (2023). Diffusion models: A comprehensive survey of methods and applications. *ACM computing surveys*, *56*(4), 1–39. https://doi.org/10.1145/3626235

[14] Zhang, Z., Song, Y., & Qi, H. (2017). Age progression/regression by conditional adversarial autoencoder. *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 5810–5818). IEEE. http://dx.doi.org/10.1109/CVPR.2017.463

[15] Snell, J., Swersky, K., & Zemel, R. (2017). Prototypical networks for few-shot learning. *Advances in neural information processing systems*, *30*. http://dx.doi.org/10.48550/arXiv.1703.05175

[16] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of big data*, *6*(1), 1–48. https://doi.org/10.1186/s40537-019-0197-0

[17] Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., … & He, Q. (2020). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, *109*(1), 43–76. https://doi.org/10.1109/JPROC.2020.3004555

[18] Odena, A., Olah, C., & Shlens, J. (2017). Conditional image synthesis with auxiliary classifier gans. *International conference on machine learning* (pp. 2642–2651). Proceedings of machine learning research (PMLR). https://doi.org/10.48550/arXiv.1610.09585

[19] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., … & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*. https://doi.org/10.48550/arXiv.1706.03762

[20] Chen, Y. S., Wang, Y. C., Kao, M. H., & Chuang, Y. Y. (2018). Deep photo enhancer: Unpaired learning for image enhancement from photographs with gans. *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 6306–6314). IEEE. http://dx.doi.org/10.1109/CVPR.2018.00660

[21] Ulery, B. T., Hicklin, R. A., Buscaglia, J., & Roberts, M. A. (2011). Accuracy and reliability of forensic latent fingerprint decisions. *Proceedings of the national academy of sciences*, *108*(19), 7733–7738. https://doi.org/10.1073/pnas.1018707108

[22] Frowd, C. D., Pitchford, M., Bruce, V., Jackson, S., Hepton, G., Greenall, M., … & Hancock, P. J. B. (2011). The psychology of face construction: giving evolution a helping hand. *Applied cognitive psychology*, *25*(2), 195–203. https://doi.org/10.1002/acp.1662

[23] Goodfellow, I. J., Shlens, J., & Szegedy, C. (2014). Explaining and harnessing adversarial examples. *ArXiv preprint arxiv:1412.6572*. https://doi.org/10.48550/arXiv.1412.6572

[24] Frowd, C., Bruce, V., & Hancock, P. J. B. (2009). Evolving facial composite systems. *Forensic update*. https://eprints.ncl.ac.uk/157049

[25] Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 4401–4410). IEEE. https://doi.org/10.1109/CVPR.2019.00453

[26] Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *ArXiv preprint arxiv:1312.6114*. https://doi.org/10.48550/arXiv.1312.6114

[27] Solomon, C. J., Gibson, S. J., & Maylin, M. (2012). EFIT-V: Evolutionary algorithms and computer composites. *Craniofacial identification*, 24–41. https://B2n.ir/ze2863

[28] Mnookin, J. L., Cole, S. A., Dror, I. E., & Fisher, B. A. J. (2010). The need for a research culture in the forensic sciences. *UCLA law review.*, *58*, 725. https://heinonline.org/HOL/LandingPage?handle

[29] Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information processing & management*, *45*(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

[30] Phillips, P. J., Yates, A. N., Hu, Y., Hahn, C. A., Noyes, E., Jackson, K., … & others. (2018). Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms. *Proceedings of the national academy of sciences*, *115*(24), 6171–6176. https://doi.org/10.1073/pnas.1721355115

[31] Akhtar, N., & Mian, A. (2018). Threat of adversarial attacks on deep learning in computer vision: A survey. *Ieee access*, *6*, 14410–14430. https://doi.org/10.1109/ACCESS.2018.2807385

[32] Wells, G. L., & Olson, E. A. (2003). Eyewitness testimony. *Annual review of psychology*, *54*(1), 277–295. https://doi.org/10.1146/annurev.psych.54.101601.145028

[33] Hendrycks, D., & Dietterich, T. (2019). Benchmarking neural network robustness to common corruptions and perturbations. *ArXiv preprint arxiv:1903.12261*. https://doi.org/10.48550/arXiv.1903.12261

[34] Samek, W., Montavon, G., Vedaldi, A., Hansen, L. K., & Müller, K.-R. (2019). *Explainable AI: interpreting, explaining and visualizing deep learning* (Vol. 11700). Springer nature. https://B2n.ir/hz3930

[35] Higgins, I., Amos, D., Pfau, D., Racaniere, S., Matthey, L., Rezende, D., & Lerchner, A. (2018). Towards a definition of disentangled representations. *ArXiv preprint arxiv:1812.02230*. https://doi.org/10.48550/arXiv.1812.02230

[36] Arora, S., & Zhang, Y. (2017). Do gans actually learn the distribution? an empirical study. *ArXiv preprint arxiv:1706.08224*. https://doi.org/10.48550/arXiv.1706.08224

[37] Begoli, E., Bhattacharya, T., & Kusnezov, D. (2019). The need for uncertainty quantification in machine-assisted medical decision making. *Nature machine intelligence*, *1*(1), 20–23. https://doi.org/10.1038/s42256-018-0004-1

[38] Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., … & others. (2021). A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *Information fusion*, *76*, 243–297. https://doi.org/10.1016/j.inffus.2021.05.008

[39] Blanz, V., & Vetter, T. (2023). A morphable model for the synthesis of 3D faces. *Seminal graphics papers: pushing the boundaries* (pp. 157–164). publication history. https://doi.org/10.1145/3596711.3596730

[40] White, D., Phillips, P. J., Hahn, C. A., Hill, M., & O'Toole, A. J. (2015). Perceptual expertise in forensic facial image comparison. *Proceedings of the royal society B: Biological sciences*, *282*(1814), 20151292. https://doi.org/10.1098/rspb.2015.1292

[41] Newsham, G. R., Cetegen, D., Veitch, J. A., & Whitehead, L. (2010). Comparing lighting quality evaluations of real scenes with those from high dynamic range and conventional images. *ACM transactions on applied perception (TAP)*, *7*(2), 1–26. https://doi.org/10.1145/1670671.1670677

[42] Cook, M., Marnerides, A., Johnson, C., & Pezaros, D. (2023). A survey on industrial control system digital forensics: Challenges, advances and future directions. *IEEE communications surveys & tutorials*, *25*(3), 1705–1747. https://doi.org/10.1109/COMST.2023.3264680

[43] He, S., Lei, Y., Zhang, Z., Sun, Y., Li, S., Zhang, C., & Ye, J. (2025). Identity deepfake threats to biometric authentication systems: Public and expert perspectives. *ArXiv preprint arxiv:2506.06825*. https://doi.org/10.48550/arXiv.2506.06825

[44] Green, B., & Chen, Y. (2019). The principles and limits of algorithm-in-the-loop decision making. *Proceedings of the acm on human-computer interaction*, *3*(CSCW), 1–24. https://doi.org/10.1145/3359152

[45] AlMutawa, A., Ikuesan, R. A., & Said, H. (2024). Towards a comprehensive metaverse forensic framework based on technology task fit model. *Future internet*, *16*(12), 437. https://doi.org/10.3390/fi16120437

[46] Talib, M. A., Majzoub, S., Nasir, Q., & Jamal, D. (2021). A systematic literature review on hardware implementation of artificial intelligence algorithms. *The journal of supercomputing*, *77*(2), 1897–1938. https://doi.org/10.1007/s11227-020-03325-8

[47] Ndibe, O. S. (2025). AI-driven forensic systems for real-time anomaly detection and threat mitigation in cybersecurity infrastructures. *International journal of research publication and reviews*, *6*(5), 389–411. https://doi.org/10.55248/gengpi.6.0525.1991