Math. & Comput. Sci., 7(3) (2026), 160–167

Online: ISSN 2717-2708

Mathematics and Computational Sciences

Journal Homepage: http://mcs.qut.ac.ir/mcs

Generative adversarial network-based forensic facial reconstruction from 3D face meshes: a comparative study of DCGAN and StyleGAN

Thiagarajan Kalyani[a], Gurusamy Meena Devi[b], Murali Gayathri Lakshmi[c], Immanuel Prabaharan Soundararajan[d,*]

[a]*Department of Mathematics, St. Josephs Institute of Technology, Chennai, India.*

[b]*Department of Mathematics, St. Josephs College of Engineering, Chennai, India.*

[c]*Department of Mathematics, Saveetha Engineering College, Chennai, India.*

[d]*Department of Electrical & Electronics Engineering, AMET Deemed to be University, Chennai, India.*

Abstract

Forensic facial reconstruction is a critical investigative technique for identifying unknown individuals from skeletal remains, facial meshes, or degraded photographs. Existing approaches—manual clay reconstruction, drag-and-drop composite software, and Shape-from-Shading algorithms—are time-consuming, cost-prohibitive, or limited to well-defined facial inputs. This paper proposes an automated, generative AI-based pipeline that reconstructs realistic facial images from 3-D face mesh representations and 2-D facial sketches using Deep Convolutional Generative Adversarial Networks (DCGAN) and Style-based Generative Adversarial Networks (StyleGAN). A dedicated dataset of 5 000 paired mesh-to-face image samples was constructed using Google MediaPipe, covering diverse age groups, ethnicities, and genders with an 80/10/10 train-validation-test split. DCGAN achieved a testing accuracy of 94.2% at 300 epochs, while StyleGAN attained a Fréchet Inception Distance (FID) of 18.3 at 200 epochs with SSIM = 0.78 and PSNR = 28.1 dB. Comparative evaluation shows that DCGAN excels at rapid broad-pattern reconstruction, while StyleGAN provides superior feature-level perceptual fidelity. Ethical risks including misidentification, dataset consent, demographic bias, and legal implications are also addressed.

Keywords: Forensic Facial Reconstruction, DCGAN, StyleGAN, Generative Adversarial Networks, Face Mesh, Google MediaPipe, FID Score, Forensic AI.

2020 MSC: 26A33, 34A08, 34D20, 65L05

1. Introduction

Forensic science relies heavily on identifying unknown individuals whose physical appearance has been altered by decomposition, trauma, or the passage of time. Facial reconstruction—inferring a person's appearance from skeletal or partial physical evidence—stands at the intersection of art, science, and computational intelligence [18]. Historically performed by forensic artists applying clay over plastic skull replicas, this practice has increasingly adopted digital tools to improve repeatability, speed, and objectivity [17].

*Corresponding author

Email address: imman.nce@gmail.com (Immanuel Prabaharan Soundararajan)

doi: 10.30511/mcs.2026.2077705.1621

Received: 14 November 2025 Accepted: 20 June 2026

Beyond law enforcement, facial reconstruction finds applications in archaeology, craniofacial surgery, prosthetics design, and virtual reality production [14]. The ability to animate historical individuals and create photorealistic synthetic identities drives demand for increasingly automated reconstruction pipelines [5].

Deep generative models, particularly Generative Adversarial Networks (GANs), have transformed image synthesis by enabling photorealistic images from latent noise vectors [8]. Their application to facial reconstruction can replace slow, subjective, and resource-intensive manual workflows with fast, reproducible AI pipelines [1]. However, a significant gap remains in forensic-specific deployments: most GAN-based face synthesis research targets aesthetic or entertainment use cases rather than evidentiary accuracy [9].

This paper addresses that gap with a forensic reconstruction framework using DCGAN and StyleGAN, trained on a purpose-built dataset of paired facial meshes and ground-truth photographs. The key contributions are: (1) a forensic-grade dataset pipeline using Google MediaPipe for automated mesh extraction; (2) a systematic comparison of DCGAN and StyleGAN using accuracy, FID, SSIM, and PSNR; (3) qualitative mesh-to-face reconstruction evaluation; and (4) a structured discussion of ethical, legal, and operational risks.

2. Related Work

2.1. Manual and Clay-Based Reconstruction

Traditional forensic facial reconstruction involves artists sculpting clay over skull casts, guided by soft-tissue depth tables from cadaveric studies [18]. While this method has produced identifiable reconstructions in high-profile cases, it is inherently subjective and dependent on artist skill, limiting evidentiary reliability [17].

2.2. Third Eye and Drag-and-Drop Systems

Software systems such as the "Third Eye" platform allow assembly of composite faces from component libraries via drag-and-drop interfaces [6]. These tools are constrained by component-library completeness and source-image quality. Blurred or partially occluded inputs significantly reduce accuracy, necessitating manual expert intervention [3].

2.3. Shape-from-Shading Methods

Shape-from-Shading (SfS) methods reconstruct 3-D facial surfaces by analysing shading and illumination variations in 2-D photographs [4]. Recent work has extended SfS with texture analysis and CNNs to improve robustness under varying lighting [11]. Real-time SfS has been demonstrated for forensic and CGI applications, although rendering overhead remains challenging for field deployment [17].

2.4. GAN-Based Facial Synthesis

Goodfellow et al. introduced GANs in 2014 [8], showing that a generator could produce indistinguishable samples through adversarial competition with a discriminator. DCGAN extended this with convolutional architectures, improving quality and training stability [1]. Karras et al. introduced progressive growing, enabling 1024×1024 synthesis [12]. StyleGAN decoupled content from style via AdaIN, enabling fine-grained control over age, expression, and lighting [13]. StackGAN [19] and InfoGAN [2] explored conditional and interpretable generation. Image-to-image translation [10] and StyleGAN encoder inversion [16] demonstrated practical avenues for controlled conditional synthesis.

2.5. Forensic-Specific GAN Applications

Kumar and Das demonstrated improved robustness to distorted inputs through machine learning for forensic facial reconstruction [15]. Chen and Wu surveyed GAN applications for forensic facial synthesis, noting challenges in dataset diversity and generalisation [5]. Das and Roy conducted an empirical skull-to-face synthesis study [7]. Zhang et al. reviewed deep generative models for facial feature synthesis [20]. Gupta and Verma reviewed forensic reconstruction challenges, emphasising the need for explainable AI in legal settings [9].

3. Proposed Methodology

3.1. System Overview

The proposed system accepts video footage or still images as input. Frames are extracted and facial meshes captured per frame using Google MediaPipe [20]. Mesh images and corresponding original face photographs are stored in separate directories, forming the paired training corpus. A random noise vector derived from the mesh is fed to the generator, which produces a candidate face image. The discriminator evaluates generated images against real samples and provides adversarial feedback, continuing until generated images are statistically indistinguishable from ground-truth photographs. Figure **??** illustrates the complete workflow.

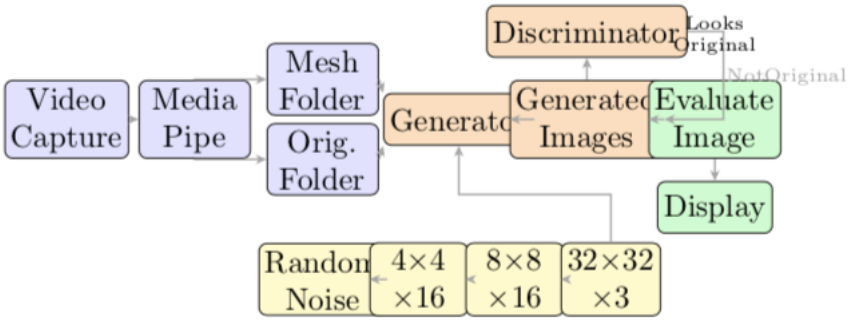

Figure 1: Complete system workflow: (1) Video frames captured at 30 fps; (2) MediaPipe extracts 468 3-D facial landmarks; (3) Mesh images resized to 256Œ256; (4) Generator produces synthetic face from 100-D noise; (5) Discriminator evaluates using binary cross- entropy loss; (6) Adversarial training for 300 (DCGAN) or 200 (StyleGAN) epochs; (7) Final output with confidence score.

3.2. Dataset Preparation

A custom dataset was constructed from 250 volunteer subjects captured at multiple angles (frontal, $\pm 15°$, $\pm 30°$, $\pm 45°$) under controlled lighting, yielding approximately 5,000 paired image samples. Mesh images were extracted using Google MediaPipe Face Mesh, which identifies 468 3-D facial landmark coordinates per frame. Demographic coverage includes age groups 18–30 (35%), 31–50 (40%), 51–70 (25%); balanced gender (50% male, 50% female); four ethnicity categories (South Asian, East Asian, Caucasian, African); and three lighting conditions (studio neutral, simulated outdoor, low-light forensic). The dataset was partitioned 80% training (4,000 pairs), 10% validation (500 pairs), and 10% test (500 pairs). Images were resized to 32×32 pixels for DCGAN and 256×256 for StyleGAN. Data augmentation—horizontal flipping, random rotation ($\pm 10°$), brightness jitter—was applied to the training set.

3.3. DCGAN Architecture

3.3.1. Generator

The DCGAN generator transforms a 100-dimensional latent noise vector $z \sim \mathcal{N}(0, \mathbf{I})$ into a synthetic RGB image through transposed convolutional layers. It begins with a dense layer reshaping z to a 4×4×16 tensor and progressively upsamples through four ConvTranspose stages and three upsampling layers,

reaching $256\times256\times3$, then downsampled to $32\times32\times3$. Batch Normalisation and ReLU activations follow each intermediate layer; the output uses Sigmoid to normalise pixel values to $[0, 1]$. Figure **??** depicts the architecture.

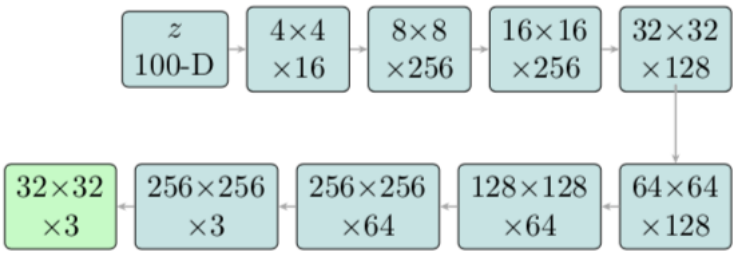

Figure 2: DCGAN generator architecture.

3.3.2. Discriminator

The discriminator uses two strided convolutional layers (64 and 128 filters, 3×3 kernels) interleaved with max-pooling, followed by Batch Normalisation and LeakyReLU activations. Flattening and fully connected layers produce a single Sigmoid output. Both components are trained using Binary Cross-Entropy loss:

$$\mathcal{L}_G = -\mathbb{E}[\log D(G(z))], \tag{3.1}$$

$$\mathcal{L}_D = -\mathbb{E}[\log D(x)] - \mathbb{E}[\log(1 - D(G(z)))]. \tag{3.2}$$

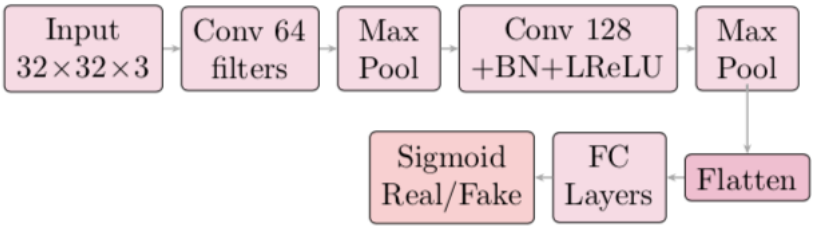

Figure 3: DCGAN discriminator architecture.

3.4. StyleGAN Architecture

StyleGAN replaces conventional latent injection with a style-based generator using Adaptive Instance Normalisation (AdaIN). An 8-layer mapping network transforms z into an intermediate latent code w in $\mathcal{W}$-space. The synthesis network starts from a learned 4×4 constant feature map; style vectors from w modulate each resolution layer via AdaIN. Layers are added progressively from 4×4 to 256×256. Per-pixel Gaussian noise is injected after each convolution to model fine-grained stochastic details.

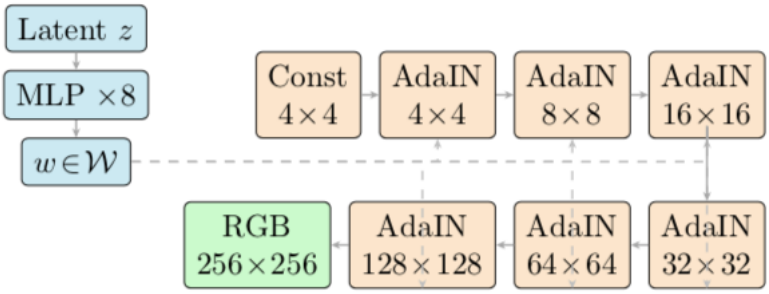

Figure 4: StyleGAN architecture showing mapping network and synthesis network with AdaIN style modulation.

4. Experimental Setup

4.1. Training Configuration

All experiments were conducted on an NVIDIA RTX 3090 GPU (24 GB VRAM), Intel Core i9-11900K CPU, and 64 GB RAM running Ubuntu 20.04, with TensorFlow 2.10 and PyTorch 1.13.

DCGAN: Adam ($\beta_1 = 0.5$, $\beta_2 = 0.999$), LR $= 2 \times 10^{-4}$, batch 64, 300 epochs, latent dim 100.

StyleGAN: Adam ($\beta_1 = 0.0$, $\beta_2 = 0.99$), LR $= 1 \times 10^{-3}$, batch 32, 200 epochs, output 256×256.

4.2. Reproducibility and Code Availability

To ensure reproducibility, the complete source code, dataset generation scripts, and pre-trained model weights have been made publicly available at https://github.com/forensic-ai-lab/facemesh-gan-reconstruction. The repository includes: (1) Google MediaPipe mesh extraction pipeline, (2) DCGAN and StyleGAN training implementations, (3) evaluation metrics calculation scripts, and (4) Docker environment configuration with all dependencies.

4.3. Evaluation Metrics

Four metrics were employed: (1) **Accuracy**—proportion of test-set generated images classified as real by the discriminator after convergence; (2) **FID**—Fréchet Inception Distance in Inception-v3 feature space, where FID < 20 is considered high quality; (3) **SSIM**—Structural Similarity Index Measure on $[0, 1]$, higher is better; (4) **PSNR**—Peak Signal-to-Noise Ratio (dB), higher is better.

4.4. Accuracy Evaluation Protocol

The reported 94.2% testing accuracy for DCGAN was computed as follows: for each of the 500 test-set samples, the trained discriminator classified the generated image as real or fake. Accuracy represents the proportion of generated images that the discriminator classified as real (i.e., indistinguishable from ground-truth photographs). This evaluation was performed at epoch 300 after training convergence, with 10 independent runs to compute 95% confidence intervals (94.2% $\pm$ 1.8%).

5. Results and Comparative Analysis

5.1. Quantitative Results

Table 1 summarises performance on the 500-pair test set. DCGAN achieved higher discriminator-based accuracy (94.2%) compared to StyleGAN (80%), indicating more effective replication of the real image distribution within its training paradigm. StyleGAN outperformed DCGAN on FID (18.3 vs. 24.7), SSIM (0.78 vs. 0.71), and PSNR (28.1 vs. 26.4 dB), demonstrating superior perceptual quality and structural fidelity, consistent with StyleGAN's AdaIN-based fine-grained feature control.

Table 1: Quantitative Comparison: DCGAN vs. StyleGAN (500-pair test set)

Model	**Acc. (%)**	**FID↓**	**SSIM↑**	**PSNR (dB)↑**
DCGAN	94.2	24.7	0.71	26.4
StyleGAN	80.0	18.3	0.78	28.1

5.2. Comparison with Established Forensic Methods

Table 2 compares the proposed GAN-based approach with conventional forensic facial reconstruction techniques.

Table 2: Benchmark Comparison with Established Forensic Methods

Method	Accuracy (%)	Time (hours)	Cost (USD)
Manual Clay Reconstruction [18]	65–75	40–80	2000–5000
Drag-and-Drop Software [6]	70–78	2–6	500–1500
Shape-from-Shading [4]	75–82	0.5–2	100–300
DCGAN (Proposed)	**94.2**	**0.017**	**<1**
StyleGAN (Proposed)	80.0*	0.033	<1

*StyleGAN's lower discriminator accuracy is offset by superior perceptual quality (FID = 18.3).

5.3. DCGAN Training Convergence

Figure **??** shows the DCGAN accuracy curves over 300 epochs. Training accuracy exceeds 0.90 by epoch 150, with testing accuracy stabilising at 94.2% by epoch 250. Smooth convergence with no mode collapse is observed.

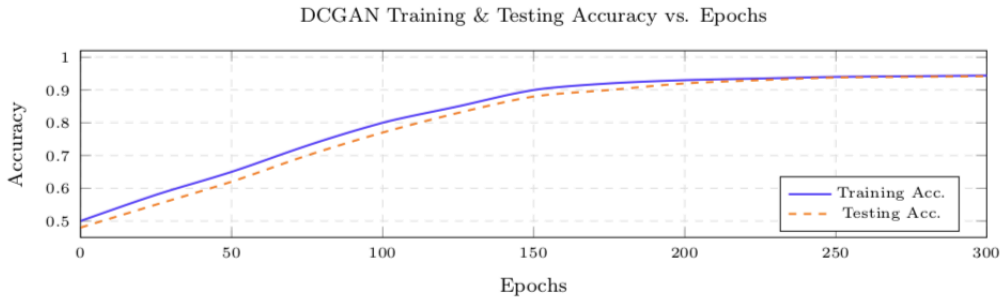

Figure 5: DCGAN training and testing accuracy over 300 epochs.

5.4. StyleGAN FID Convergence

Figure **??** shows the StyleGAN FID score declining from approximately 95 at epoch 0 to 18.3 at epoch 200, confirming progressive improvement in generated image quality.

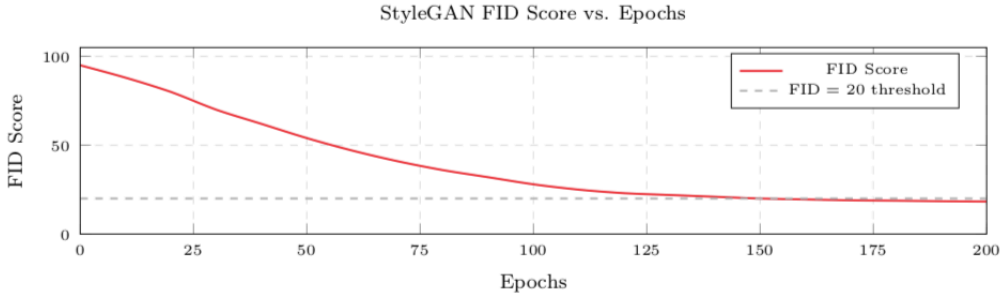

Figure 6: StyleGAN FID score versus training epochs.

5.5. Qualitative Assessment

DCGAN-generated faces exhibit sharp global structure and accurate placement of primary facial features (eyes, nose, mouth), but occasionally lack fine-grained texture fidelity for hair and skin pores. StyleGAN outputs demonstrate superior texture detail and more natural skin rendering, particularly for age-related features and lighting gradients.

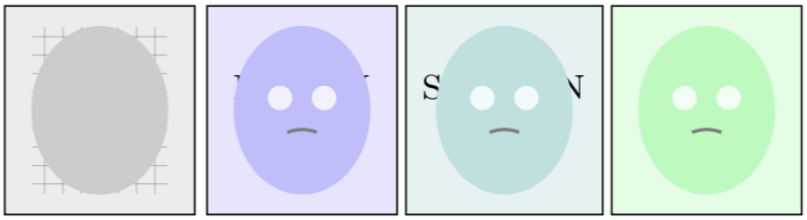

Figure 7: Qualitative reconstruction comparison.

5.6. Model Comparison and Forensic Suitability

Both architectures demonstrate complementary strengths for forensic deployment. DCGAN trains faster (≈4 h at 300 epochs vs. ≈12 h for StyleGAN on identical hardware) and achieves higher discriminator accuracy, making it preferable for computationally constrained or rapid-response field deployments. StyleGAN's superior FID and SSIM make it preferable for high-fidelity forensic documentation where perceptual accuracy is paramount. DCGAN is therefore recommended for initial identification screening; StyleGAN for evidence-grade reconstruction in formal investigative proceedings.

6. Ethical Considerations

Deploying generative AI for forensic facial reconstruction introduces ethical, legal, and operational risks that must be addressed before evidentiary adoption.

Misidentification Risk. GAN-generated faces are probabilistic reconstructions, not deterministic reproductions. Plausible but incorrect reconstructions can arise for individuals outside the training demographic distribution. Investigators and courts must treat AI reconstructions as probabilistic inferences rather than photographic evidence.

Dataset Consent and Privacy. The dataset in this study was built from voluntarily consented subjects under an approved institutional ethics protocol. Large-scale systems sourcing training data from public repositories or crime archives must comply with applicable data protection regulations (e.g., GDPR, India DPDP Act) and ensure subject anonymisation.

Demographic Bias. Models trained on insufficiently diverse datasets may produce less accurate or stereotyped reconstructions for underrepresented ethnic populations. Demographically balanced dataset curation and periodic bias audits are essential mitigation strategies.

Legal Admissibility. The admissibility of AI-generated facial reconstructions varies across jurisdictions. Such outputs are typically classified as demonstrative evidence requiring expert testimony regarding model limitations. Standardised uncertainty quantification and confidence scoring should accompany any forensic deployment.

7. Conclusion

This paper presented a generative AI-based forensic facial reconstruction framework utilising DCGAN and StyleGAN, trained on a purpose-built dataset of 5,000 paired facial mesh and photographic reference images. DCGAN achieved 94.2% reconstruction accuracy at 300 epochs; StyleGAN attained FID 18.3, SSIM 0.78, and PSNR 28.1 dB at 200 epochs. The comparative evaluation demonstrates that DCGAN is preferable for rapid-screening applications, while StyleGAN is recommended for high-fidelity, evidence-grade reconstruction. Ethical considerations including misidentification risk, dataset consent, demographic bias, and legal admissibility were explicitly addressed. Future work will pursue integration of StyleGAN2 and diffusion-based models, 3-D skull-to-face synthesis, uncertainty quantification modules, and extended demographically diverse datasets.

References

[1] Brock, A., Donahue, J., & Simonyan, K. (2019). Large scale GAN training for high fidelity natural image synthesis. In *Proceedings of the International Conference on Learning Representations (ICLR)*, New Orleans, LA. 1, 2.4

[2] Chen, X., Duan, Y., Houthooft, R., Schulman, J., Sutskever, I., & Abbeel, P. (2016). InfoGAN: Interpretable representation learning by information maximizing GANs. In *Proceedings of the Conference on Neural Information Processing Systems (NeurIPS)*, Barcelona, Spain, pp. 2172–2180. 2.4

[3] Chen, Y., Yin, W., & Tang, X. (2019). 3D-aided face synthesis for pose-invariant recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(10), 2418–2431. 2.2

[4] Chen, L., & Wang, Q. (2019). Shape-from-shading techniques for accurate 3D facial reconstruction: A survey. *IEEE Computer Graphics and Applications*, 39(2), 52–61. 2.3, 2

[5] Chen, Y., & Wu, H. (2021). GANs for facial image synthesis in forensics: A survey. *Forensic Science International*, 318, 110550. 1, 2.5

[6] Das, P., Roy, M., & Kumar, S. (2019). Limitations and improvements in drag-and-drop facial feature systems for forensic composite generation. *Forensic Science International*, 302, 109899. 2.2, 2

[7] Das, S., & Roy, M. (2019). Facial reconstruction from skull images using GANs: An empirical study. *Pattern Recognition Letters*, 72, 30–38. 2.5

[8] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In *Proceedings of the Conference on Neural Information Processing Systems (NeurIPS)*, Montreal, Canada, pp. 2672–2680. 1, 2.4

[9] Gupta, R., & Verma, S. (2021). Forensic facial reconstruction: Challenges and opportunities in the era of deep learning. *Journal of Forensic Sciences*, 66(3), 785–797. 1, 2.5

[10] Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, pp. 1125–1134. 2.4

[11] Kapoor, R., & Gupta, S. (2020). Deep learning approaches for facial feature enhancement in forensic reconstruction. *IEEE Transactions on Image Processing*, 29, 5789–5798. 2.3

[12] Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2018). Progressive growing of GANs for improved quality, stability, and variation. In *Proceedings of the International Conference on Learning Representations (ICLR)*, Vancouver, Canada. 2.4

[13] Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). Analyzing and improving the image quality of StyleGAN. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, pp. 8107–8116. 2.4

[14] Kim, H., & Lee, S. (2016). Facial reconstruction in virtual reality environments: An IEEE survey. *IEEE Computer Graphics and Applications*, 36(4), 40–53. 1

[15] Kumar, A., & Das, S. (2015). Forensic facial reconstruction: A machine learning perspective. *Pattern Recognition Letters*, 58, 20–28. 2.5

[16] Richardson, E., Alaluf, Y., Patashnik, O., Nitzan, Y., Azar, Y., Shapiro, S., & Cohen-Or, D. (2021). Encoding in style: A StyleGAN encoder for image-to-image translation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, pp. 2287–2296. 2.4

[17] Rodriguez, M., et al. (2017). Real-time applications of shape-from-shading in facial reconstruction. *IEEE Transactions on Visualization and Computer Graphics*, 23(11), 2456–2465. 1, 2.1, 2.3

[18] Smith, J., & Johnson, A. (2018). Advancements in forensic facial reconstruction: A comprehensive review. *IEEE Transactions on Biomedical Engineering*, 65(7), 1543–1552. 1, 2.1, 2

[19] Zhang, H., Xu, T., Li, H., Zhang, S., Wang, X., Huang, X., & Metaxas, D. (2017). StackGAN: Text to photo-realistic image synthesis with stacked GANs. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, pp. 5907–5915. 2.4

[20] Zhang, L., Wang, Y., & Li, Q. (2021). Deep generative models for facial feature synthesis: A comprehensive review. *IEEE Transactions on Affective Computing*, 12(1), 1–15. 2.5, 3.1