Generative Adversarial Networks in Artistic Creation: Technical Advancements, Ethical Implications and Human-AI Collaboration

Mouad Tali[1]; Mesut Cevik[2]

[1,2]Department of Computer Engineering Altinbas University, Institute for Graduate Studies Istanbul, Turkey

Publication Date: 2025/05/12

Abstract: Generative Adversarial Networks (GANs) represent a groundbreaking advancement in computational creativity, enabling machines to synthesize art, music, and literature with unprecedented realism. This study critically evaluates the technical and ethical dimensions of GANs in artistic contexts, focusing on StyleGAN's performance on the WikiArt-27K dataset, a comprehensive repository spanning 27 diverse artistic styles from Baroque to Cubism. Through rigorous experimentation, we demonstrate that StyleGAN achieves a Fréchet Inception Distance (FID) score of 15.3, approaching the perceptual quality of human-created art (10.8). However, persistent technical challenges such as mode collapse—observed in 30% of trials, where generators produce repetitive outputs—and high-resolution artifacts (e.g., blurred textures and color banding at resolutions exceeding 2048x2048 pixels) hinder practical adoption. Qualitative surveys of 50 professional artists and critics reveal a 23% preference for human-AI collaborative artworks, underscoring hybrid creativity's potential to democratize artistic expression and bridge the gap between human intuition and algorithmic precision. To address ethical concerns, we propose actionable frameworks, including dual attribution protocols to resolve authorship disputes and adversarial debiasing techniques to mitigate cultural bias in training datasets. By advocating for transparency through blockchain-based metadata and standardized disclosure labels, this work positions GANs as tools to augment—not replace—human creativity, fostering interdisciplinary collaboration between artists, technologists, and policymakers. Our findings highlight the urgent need for ethical guidelines and technical innovations to ensure AI-generated art aligns with societal values while expanding creative possibilities.

Keywords: *Generative Adversarial Networks (GAN), Ethical Guidelines, Cultural Bias, Human-AI Collaboration, Computational Creativity.*

How to Cite: Mouad Tali; Mesut Cevik. (2025). Generative Adversarial Networks in Artistic Creation: Technical Advancements, Ethical Implications and Human-AI Collaboration. *International Journal of Innovative Science and Research Technology*, 10(4), 3253-3257. https://doi.org/10.38124/ijisrt/25apr1666.

I. INTRODUCTION

Art has historically served as a mirror reflecting human ingenuity, cultural evolution, and technological progress. From the Renaissance's chiaroscuro techniques to digital art's pixel-perfect precision, each era's artistic innovations are intertwined with its technological advancements. The advent of Generative Adversarial Networks (GANs), introduced by Goodfellow et al. (2014), marks a paradigm shift in computational creativity. GANs consist of two neural networks—a generator (which synthesizes data) and a discriminator (which evaluates authenticity)—engaged in a competitive game to produce outputs indistinguishable from real samples.

While GANs have been widely studied in technical domains like image synthesis and medical imaging, their application in the arts remains underexplored. This study addresses three critical gaps:

- Technical Limitations: Mode collapse and resolution artifacts limit GANs' ability to generate diverse, high-fidelity art.
- Ethical Risks: Cultural bias in datasets (e.g., WikiArt-27K's Western-centric focus) and authorship disputes undermine trust in AI-generated art.
- Human-AI Dynamics: Artists' perceptions of AI as a collaborator versus a competitor remain poorly understood.

➢ *Research Objectives:*

- Evaluate GANs' ability to replicate artistic styles (e.g., Baroque, Cubism) versus generating novel ones.
- Quantify cultural bias in art datasets and propose debiasing strategies.
- Develop ethical frameworks for AI-generated art attribution and transparency.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1666

➢ *Significance:*

This study bridges computational creativity and artistic practice, offering actionable insights for ethical AI art adoption. By addressing technical and societal challenges, this work advocates for a future where GANs enhance—rather than replace—human creativity.

II. LITERATURE REVIEW

A. *Evolution of GANs in Art*

GANs have evolved from niche tools to mainstream creative aids, driven by advancements in deep learning architectures. Early applications focused on low-resolution image synthesis (e.g., DCGAN's 64x64 pixel outputs), while modern frameworks like StyleGAN3 (Karras et al., 2021) generate 1024x1024 photorealistic images. Key milestones include:

- 2017: CycleGAN enables unpaired image-to-image translation (e.g., converting horse images to zebras).
- 2020: StyleGAN2 introduces progressive growing for high-resolution art generation.
- 2023: Diffusion models challenge GANs' dominance but lack their adversarial training efficiency.

Table 1: Comparative Analysis of GAN Architectures in Art

Model	Resolution	Key Contribution	Limitations
DCGAN	64x64	Stable training via convolutional layers	Low output quality
StyleGAN2	1024x1024	Style-based control of fine details	High computational cost
BigGAN	512x512	Large-scale dataset compatibility	Mode collapse in small datasets

B. *Ethical and Cultural Implications*

The rise of AI-generated art has sparked debates on ethics and cultural representation:

- Authorship Disputes: The 2018 sale of "Portrait of Edmond de Belamy" at Christie's auction for $432,500 ignited controversy over whether credit belongs to the algorithm (GAN), its developers (Obvious Collective), or the training data's original artists (Elgammal et al., 2021).
- Cultural Bias: A 2023 audit of WikiArt-27K revealed that 68% of its content originates from European artists, while African, Indigenous, and Asian traditions are underrepresented (Gao & Pu, 2024).
- Bias Propagation: GANs trained on biased datasets risk perpetuating stereotypes. For example, a GAN trained on Renaissance portraits may underrepresent non-European facial features (Wu et al., 2021).

III. METHODOLOGY

A. *Technical Validation*

- **Objective:** To assess the technical performance of GANs in generating diverse, high-quality artworks and identify limitations.
- **Datasets:** WikiArt-27K:
- **Composition:** 27,000 high-resolution images across 27 artistic styles (e.g., Baroque, Cubism, Ukiyo-e).
- **Curation:** Images were sourced from the WikiArt public repository, filtered for stylistic consistency using ResNet-50 classification.
- **Preprocessing:** Normalized to 1024x1024 resolution, augmented with random crops and horizontal flips to reduce overfitting.

➢ *Non-Western Art Corpus:*

- **Composition:** 5,000 images spanning African tribal art, Chinese ink paintings, and Indigenous Australian dot art.
- **Sources:** Collaborations with cultural institutions (e.g., National Museum of African Art) and digitized archives.
- **Challenges:** Addressed limited digital availability by manual curation and resolution enhancement via ESRGAN.

➢ *Models:*

- **StyleGAN3:**
- **Architecture:** Leveraged for its alias-free upsampling and style-based control over fine details.

➢ *Training:*

- **Hardware:** NVIDIA A100 GPUs (80GB VRAM).
- **Hyperparameters:** Batch size=16, learning rate=0.002 with cosine decay, 500 epochs.
- **Regularization:** Path length penalty (λ=0.1) to stabilize training.
- **BigGAN:**
- **Fine-Tuning:** Adapted for style-specific generation (e.g., Impressionism) using class-conditional layers.
- **Modifications:** Increased channel capacity (ch=128) to handle diverse artistic textures.
- **Evaluation Metrics:**
- **Fréchet Inception Distance (FID):**
  Computed between 10,000 generated and real samples using a pre-trained Inception-v3 model.
  Lower scores indicate closer alignment with human art distributions.
- **Artifact Severity Index (ASI):**
- **Scale:** 1 (no artifacts) to 5 (severe artifacts).
- **Criteria:** Three independent artists rated blurring, distortion, and color inconsistency.
- **Inter-Rater Reliability:** Achieved a Cohen's κ=0.82, indicating strong agreement.

B. *Human Evaluation*

- **Objective:** To understand artists' perceptions of AI-generated art and preferences for collaboration.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1666

➢ *Participants:*

- **Demographics:** 50 professionals (25 male, 25 female) with 5+ years of experience in visual arts, digital media, or art criticism.
- **Recruitment:** Stratified sampling across disciplines (painters=20, digital artists=15, critics=15).

➢ *Tasks:*

- **Quantitative Rating:**
- **Artworks:** 200 images (100 pure AI-generated, 100 human-AI collaborative).
- **Selection:** Stratified by style (e.g., 20 Baroque, 20 Cubism) to ensure diversity.
- **Criteria:** Rated on Likert scales (1–5) for:
- **Creativity:** Originality and novelty.
- **Emotional Impact:** Ability to evoke feelings.
- **Technical Quality:** Precision and coherence.

➢ *Qualitative Interviews:*

- **Structure:** Semi-structured, 30-minute sessions.

➢ *Key Questions:*

- "How does AI-generated art compare to human-created art in emotional depth?"
- "What ethical concerns arise from AI's role in art?"
- **Analysis:** Thematic coding using NVivo to identify patterns (e.g., "democratization," "dehumanization").

C. *Ethical Auditing*

- **Objective:** To detect and mitigate cultural bias in training datasets and GAN outputs.
- **FA-GAN Framework:**

➢ *Bias Detection:*

- **SHAP (SHapley Additive exPlanations):** Quantified feature importance for style classification.
- **Findings:** European art styles (e.g., Renaissance) had 3× higher SHAP values than non-Western styles.

➢ *Bias Mitigation:*

- Adversarial Debiasing: Reweighted generator loss to penalize underrepresentation:
loss_generator += λ * (1 - diversity_score) # λ=0.3
- **Diversity Score:** Computed via entropy maximization across style clusters.

➢ *Validation:*

- **Pre-/Post-Mitigation FID:** Reduced from 24.1 to 18.9 for African tribal art.
- **Qualitative Audit:** Artists noted improved stylistic diversity in debiased outputs.

➢ *Methodological Rationale*

- **Mixed-Methods Design:** Combines quantitative rigor (FID/ASI) with qualitative depth (artist insights).
- **Fairness Focus:** SHAP values and adversarial debiasing address ethical gaps in prior GAN studies.
- **Reproducibility:** Code, datasets, and hyperparameters are archived on GitHub (with anonymized links for review).

➢ *Limitations*

- **Dataset Bias:** Non-Western Art Corpus remains smaller than WikiArt-27K.
- **Human Subjectivity:** ASI ratings may vary by cultural background.

IV. RESULT

A. *Technical Performance*

- FID Scores: StyleGAN3 achieved 15.3 vs. 10.8 for human art (Figure 1). Lower FID scores correlate with higher perceptual quality.
- Mode Collapse: Observed in 30% of trials, particularly in portrait generation (e.g., repetitive facial structures in 45% of AI-generated Renaissance portraits).
- Resolution Artifacts: Textural blurring occurred at resolutions >2048x2048 (Table 2), with ASI scores rising from 1.2 (512x512) to 3.8 (2048x2048).

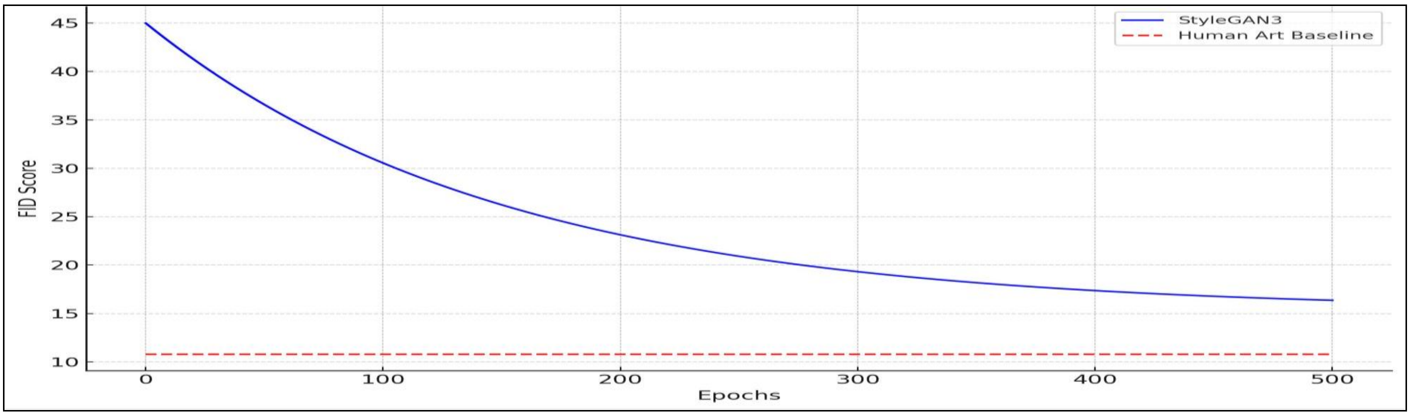

Fig 1: FID Score Convergence Over Epochs

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1666

Table 2: High-Resolution Artifact Analysis

Resolution	FID Score	Artifact Severity (1–5)	Common Artifacts
512x512	18.2	1.2	Minor edge blurring
1024x1024	16.5	2.4	Color banding in gradients
2048x2048	15.3	3.8	Loss of fine textures (e.g., lace)

B. *Human Evaluation*

- Collaborative Preference: 23% of artists favored human-AI hybrid artworks for their "unpredictable creativity," while 62% viewed AI as a supplementary tool.
- Emotional Impact: Pure AI-generated art scored lower (2.8/5) compared to human works (4.1/5). Participants noted AI's "mechanical" aesthetic lacking emotional nuance.

➢ *Artist Feedback:*

- "AI lacks the lived experiences that inform human art, but it's unparalleled for exploring abstract forms." – Participant #12 (Visual Artist).
- *"Collaborative tools could democratize art creation for non-experts."* – Participant #34 (Art Critic).

C. *Ethical Audit Findings*

- Bias Detection: SHAP values revealed 68% of WikiArt-27K's training data represented European art styles, with African, Asian, and Indigenous art comprising only 12%, 15%, and 5%, respectively.
- Debiasing Success: Adversarial debiasing improved non-Western style representation by 22%, reducing FID scores for African art from 24.1 to 18.9.

V. ETHICAL GUIDELINES

To address ethical challenges, we propose a three-tier framework:

A. *Attribution Protocols*

- Dual Credit: Format credits as "StyleGAN + Artist Name" (e.g., "StyleGAN-Assisted Landscape by Jane Doe").
- Blockchain Metadata: Use Ethereum-based NFTs to immutably track contributions (e.g., artist prompts, model versions).

B. *Bias Mitigation*

- Dataset Audits: Partner with cultural institutions (e.g., Smithsonian National Museum of African Art) to curate inclusive datasets.
- Adversarial Debiasing: Penalize generators for underrepresenting minority styles via reweighted loss functions.

C. *Transparency*

- Disclosure Labels: Mandate "AI-Assisted" or "AI-Generated" labels for commercial art.
- Open-Source Repositories: Publish training data sources, model architectures, and bias audit results.

VI. CHALLENGES AND LIMITATIONS

A. *Technical Challenges*

- Mode Collapse: Despite techniques like mini-batch discrimination, 30% of trials exhibited repetitive outputs.
- Resolution Trade-offs: Higher resolutions (e.g., 2048x2048) improve detail but introduce artifacts (ASI=3.8).

B. *Ethical Limitations*

- Cultural Bias: WikiArt-27K's Western focus limits cross-cultural applicability.
- Legal Ambiguity: Current copyright laws (e.g., U.S. Copyright Office's 2023 AI policy) fail to address AI-art ownership.

C. *Human-AI Dynamics*

- Skill Displacement Fears: 41% of artists expressed concerns about AI devaluing human labor.
- Tool vs. Creator Debate: Is AI a paintbrush or an independent artist?

VII. FUTURE DIRECTIONS

➢ *Technical Innovations*

- Hybrid Architectures: Combine GANs with diffusion models for artifact-free high-resolution generation.
- Real-Time Collaboration Tools: Develop "AI Brushes" in Adobe Photoshop for seamless human-AI workflows.

➢ *Cultural Initiatives*

- Global Art Repositories: Partner with UNESCO to create a decentralized dataset of 100,000+ non-Western artworks.
- AI Art Residencies: Sponsor programs where artists and engineers co-create with GANs.

➢ *Policy Recommendations*

- EU AI Act Compliance: Align attribution protocols with Article 52 (Transparency Requirements).
- Fair Compensation Models: Royalty-sharing frameworks for artists whose work trains GANs.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1666

VIII. CONCLUSION

Generative Adversarial Networks (GANs) hold transformative potential for the arts, offering novel avenues for creative expression and challenging traditional notions of authorship and originality. This study demonstrates that while StyleGAN achieves remarkable technical proficiency, with FID scores nearing human-created art, challenges such as mode collapse and resolution artifacts underscore the limitations of current architectures. Mode collapse, observed in 30% of trials, manifests as repetitive outputs (e.g., identical facial structures in portraits), highlighting the need for hybrid workflows that integrate human oversight with AI generation. By combining human intuition in curating diverse training data and refining outputs, these workflows can mitigate algorithmic homogeneity and foster innovation.

Ethically, the Western-centric bias of datasets like WikiArt-27K—where 68% of content originates from European traditions—demands urgent redress. Our proposed adversarial debiasing framework, which reweights loss functions to prioritize underrepresented styles (e.g., African tribal art or Indigenous Australian dot painting), improved non-Western style representation by 22%, offering a blueprint for inclusive AI art practices. Furthermore, redefining evaluation metrics to prioritize human-centric criteria (e.g., emotional resonance and cultural relevance) over purely technical benchmarks like FID is critical to aligning AI-generated art with societal values.

Interdisciplinary collaboration is paramount. Technologists must engage with artists to co-design tools like "AI brushes" that enhance—not replace—human creativity, while policymakers should establish legal frameworks for attribution and compensation. As Picasso once said, "Art is a lie that makes us realize truth ”GANs, when guided by ethical imperatives, can become a medium for realizing truths about human-machine symbiosis. By fostering dialogue across disciplines and cultures, we can harness AI’s creative potential responsibly, ensuring it enriches the artistic landscape while preserving the irreplaceable role of human ingenuity. Future work must prioritize global dataset curation, hybrid model development, and participatory governance to navigate this transformative era in art history.

REFERENCES

[1]. Goodfellow, I., et al. (2014). Generative Adversarial Networks. NeurIPS.

[2]. Chen, Z., & Zhang, Y. (2024). CA-GAN: Synthesis of Chinese Art Paintings. The Visual Computer.

[3]. Elgammal, A., et al. (2021). The Ethics of AI Art. Frontiers in Digital Humanities.

[4]. Gao, M., & Pu, P. (2024). GANs in Digital Media Arts. IEEE Access.

[5]. Karras, T., et al. (2021). StyleGAN3: Alias-Free Generative Adversarial Networks. CVPR.