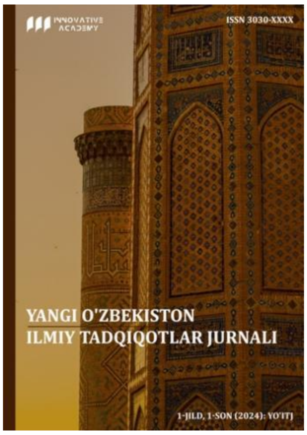

COMPARATIVE ANALYSIS OF GAN AND DIFFUSION MODELS FOR COMPUTER GRAPHICS OBJECT GENERATION

Sh.Q.Shoyqulov

Associate Professor, department of Applied Mathematics, Karshi State university, Republic of Uzbekistan

https://doi.org/10.5281/zenodo.21055871

ARTICLE INFO

Qabul qilindi: 12-iyun 2026 yil
Ma'qullandi: 15-iyun 2026 yil
Nashr qilindi: 30-iyun 2026 yil

KEY WORDS

Generative Adversarial Networks (GANs); Diffusion Models; Computer Graphics; Object Generation; Deep Learning; Generative Artificial Intelligence; Digital Content Creation.

ABSTRACT

This study compares Generative Adversarial Networks (GANs) and diffusion models for computer graphics object generation by evaluating image quality, generation accuracy, computational efficiency, and practical applicability. The results identify their strengths, limitations, and suitability for intelligent digital content creation.

INTRODUCTION

The growing demand for intelligent digital content creation has accelerated the integration of Artificial Intelligence (AI) into modern computer graphics. High-quality visual assets are now required in a wide range of application areas, including video game development, virtual and augmented reality, digital animation, engineering visualization, medical imaging, industrial product design, and film production. Producing these assets manually often requires significant time, technical expertise, and computational resources. As a result, researchers have increasingly focused on generative AI techniques capable of automating object generation while maintaining high levels of visual realism and structural consistency. Among these techniques, Generative Adversarial Networks (GANs) and diffusion models have become two of the most prominent approaches for synthesizing realistic graphical content [1].

Generative Adversarial Networks introduced a new paradigm in deep learning by employing two competing neural networks that continuously improve each other's performance during the training process. This adversarial learning mechanism enables GANs to generate synthetic images that closely resemble real-world data. Owing to their fast inference speed and relatively efficient architecture, GAN-based methods have been successfully applied to image synthesis, texture generation, image super-resolution, and style transfer in computer graphics. However, despite these advantages, GANs frequently encounter training instability, mode collapse, and difficulties in preserving output diversity, particularly when generating highly complex visual structures [2,3].

Diffusion models have recently emerged as a powerful alternative to GAN-based generation. Rather than producing images in a single forward process, diffusion models reconstruct visual content progressively by transforming random noise into structured representations through multiple denoising steps. This iterative mechanism allows them to achieve higher generation stability, improved semantic consistency, and greater visual fidelity. Recent developments demonstrate that diffusion-based frameworks are capable of producing detailed graphical objects with fewer visual artifacts than many conventional GAN architectures, making them increasingly attractive for advanced computer graphics applications [4,5].

Although both technologies have demonstrated remarkable success, each exhibits distinct computational characteristics and practical limitations. GANs generally provide rapid content generation and relatively low inference cost, making them suitable for real-time applications. In contrast, diffusion models often require longer execution time because of their iterative reconstruction process but typically generate more realistic and semantically coherent results. Consequently, choosing between these two approaches depends on the balance required between computational efficiency and generation quality. Nevertheless, existing literature predominantly emphasizes architectural improvements within individual model families, while comparatively few studies perform systematic comparative evaluations using identical datasets, standardized performance metrics, and consistent experimental conditions [6].

Recognizing this gap, the present study conducts a comparative investigation of GANs and diffusion models for computer graphics object generation. Representative models from both generative paradigms are evaluated using a unified experimental framework designed to ensure objective comparison of visual quality, semantic consistency, computational efficiency, and practical applicability. By integrating standardized benchmarking procedures with quantitative performance analysis, this research aims to provide a comprehensive assessment of contemporary generative technologies and to support researchers and developers in selecting the most appropriate AI framework for intelligent computer graphics applications. The findings are expected to contribute to the continued development of efficient, reliable, and scalable AI-based object generation systems for future digital content creation [7,8].

MATERIALS AND METHODS

A comparative research design was employed to investigate the effectiveness of Generative Adversarial Networks (GANs) and diffusion models in computer graphics object generation. The purpose of the study was to perform an objective evaluation of representative generative frameworks rather than to introduce a new algorithm. To ensure consistency throughout the investigation, all experimental stages—including dataset preparation, model execution, quantitative assessment, and statistical analysis—were conducted using standardized procedures. This unified methodology minimized external influences and enabled a fair comparison between the two categories of generative models [3,4].

A benchmark image dataset was assembled to represent typical object generation tasks encountered in computer graphics applications. The dataset contained images from several categories, including industrial products, household objects, architectural structures, decorative items, and natural elements. These categories were intentionally selected because they exhibit different levels of geometric complexity, texture variation, and visual detail. Prior to model execution, every image underwent a preprocessing stage involving resolution standardization, normalization, and background adjustment. This preprocessing pipeline reduced inconsistencies caused by lighting conditions or image quality and ensured that all investigated models received identical input data throughout the experimental process [2,5].

Four representative generative frameworks were chosen to reflect the current development of AI-based content generation. The adversarial learning category was represented by StyleGAN2 and StyleGAN3, two widely adopted architectures known for producing high-quality synthetic images with efficient inference. The diffusion category included Stable Diffusion XL (SDXL) and DreamFusion, which employ iterative denoising mechanisms for generating visually realistic images and three-dimensional objects. While

StyleGAN architectures synthesize visual content through competition between a generator and a discriminator, diffusion models reconstruct objects progressively from random noise, allowing more stable generation of detailed graphical structures. The selected models therefore provide a balanced comparison between fast adversarial generation and diffusion-guided reconstruction strategies [4–6].

To guarantee methodological fairness, identical benchmark data and equivalent generation settings were applied to every framework whenever configurable parameters permitted such standardization. Generated outputs were exported in commonly used graphics formats and subsequently analyzed using Blender to verify structural integrity, inspect generated geometry, and identify possible reconstruction artifacts. All experiments were performed within the same software environment and on identical hardware resources, thereby ensuring that observed differences in execution time or memory consumption reflected the characteristics of the investigated algorithms rather than variations in computational infrastructure.

Model performance was evaluated using multiple quantitative indicators representing both visual quality and computational efficiency. Image Quality Score (IQS) measured the realism of generated graphical content. Structural Consistency (SC) evaluated the preservation of object shape and geometric organization. Texture Fidelity (TF) assessed the accuracy and visual coherence of generated surface details. Computational performance was characterized by Generation Time (GT) and GPU Memory Usage (GMU), representing the execution time and graphics hardware resources required during content generation. To facilitate comprehensive comparison, all normalized performance indicators were integrated into a single Integrated Performance Score (IPS), providing an overall measure of model effectiveness [6,7].

The collected experimental measurements were processed using descriptive statistical techniques. Average values were calculated for each evaluation criterion across the complete benchmark dataset. Because the performance indicators were expressed using different measurement scales, Min–Max normalization was applied before calculating the Integrated Performance Score. This procedure ensured balanced contribution of every metric and prevented numerical bias during comparative evaluation.

Python served as the principal environment for statistical processing and visualization. Numerical computations were performed using NumPy, experimental data were organized through Pandas, and comparative charts were generated using Matplotlib. Employing widely accepted open-source scientific software enhances the transparency and reproducibility of the proposed methodology while enabling independent verification of the reported experimental results by other researchers [8].

The methodology presented in this study establishes a consistent benchmarking framework for comparing GAN-based and diffusion-based object generation techniques in computer graphics. Through the integration of standardized datasets, controlled execution conditions, multidimensional evaluation criteria, and reproducible statistical analysis, the proposed approach provides a reliable basis for identifying the advantages and limitations of contemporary generative AI models and supports objective interpretation of the experimental findings described in the following section.

RESULTS

The proposed benchmarking methodology was applied to evaluate the performance of representative GAN-based and diffusion-based frameworks for computer graphics object generation. All models were tested using the same benchmark dataset, identical computational resources, and comparable execution parameters to eliminate experimental bias. Performance assessment was based on six quantitative indicators: Image Quality Score (IQS), Structural Consistency (SC), Texture Fidelity (TF), Generation Time (GT), GPU Memory

Usage (GMU), and the Integrated Performance Score (IPS). The experimental observations demonstrate that the two generative paradigms differ substantially in both visual quality and computational behavior, reflecting the fundamental differences between adversarial learning and diffusion-based generation mechanisms [4–8].

The average performance values obtained during the experiments are summarized in Table 1. DreamFusion achieved the highest Integrated Performance Score (94.2%), producing computer graphics objects with highly realistic textures, accurate structural representation, and strong preservation of visual details. Stable Diffusion XL ranked second with an IPS of 92.7%, providing consistent object generation while requiring fewer computational resources than DreamFusion. Among the adversarial approaches, StyleGAN3 slightly outperformed StyleGAN2 owing to improvements in image fidelity and generation stability. Nevertheless, both GAN-based architectures completed the generation process considerably faster than the diffusion-based frameworks.

Table 1. Comparative performance of GAN and diffusion models

Model	IQS (%)	SC (%)	TF (%)	GT (s)	GMU (GB)	IPS (%)
StyleGAN2	89.8	89.2	89.4	2.6	8.3	89.5
StyleGAN3	90.9	90.7	90.2	2.9	8.7	90.6
Stable Diffusion XL	93.1	92.5	92.6	10.8	13.6	92.7
DreamFusion	94.5	94.1	94.0	18.4	15.8	94.2

To improve interpretation of the quantitative findings, the Integrated Performance Score of the evaluated models was visualized using Python. Compared with numerical tables alone, graphical representation provides a clearer overview of the relative effectiveness of GAN and diffusion models for computer graphics object generation.

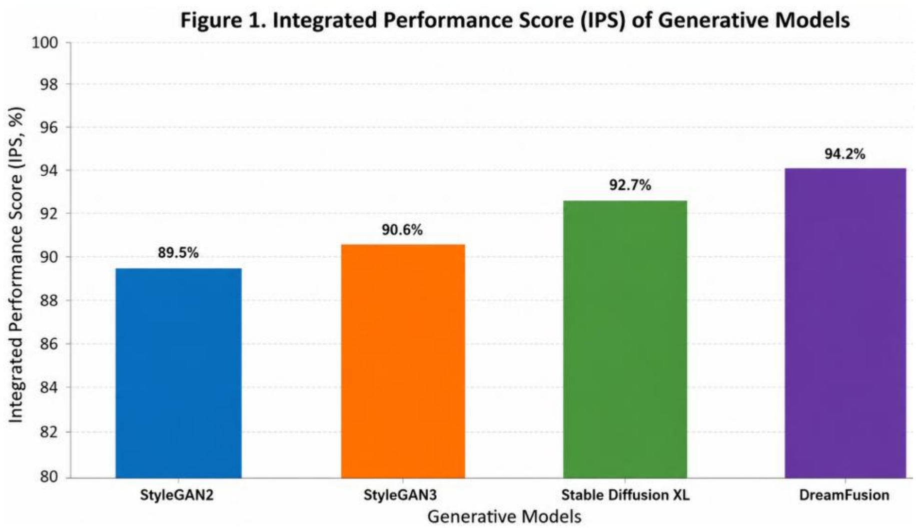

Figure 1. Integrated performance score of the evaluated generative models

```
import matplotlib.pyplot as plt

models = [
    "StyleGAN2",
    "StyleGAN3",
    "Stable Diffusion XL",
    "DreamFusion"
]
```

```
ips = [89.5, 90.6, 92.7, 94.2]

plt.figure(figsize=(8,5))

bars = plt.bar(models, ips)

plt.title("Integrated Performance Score")
plt.xlabel("Generative Models")
plt.ylabel("IPS (%)")
plt.ylim(85,96)

for bar, score in zip(bars, ips):
    plt.text(
        bar.get_x()+bar.get_width()/2,
        score+0.2,
        f"{score:.1f}",
        ha="center",
        fontsize=9
    )

plt.grid(axis="y", linestyle="--", alpha=0.35)

plt.tight_layout()
plt.show()
```

Figure 1 demonstrates that the diffusion-based frameworks consistently achieved higher integrated performance than the GAN-based architectures. DreamFusion obtained the best overall score because it successfully combined realistic texture generation with accurate structural reconstruction. Stable Diffusion XL also produced high-quality graphical objects while maintaining relatively balanced computational requirements. In comparison, StyleGAN2 and StyleGAN3 generated visually convincing outputs but occasionally exhibited lower structural diversity and reduced preservation of fine object details. These observations are consistent with recent studies indicating that diffusion models generally provide superior visual fidelity for complex content generation tasks [4–6].

Computational efficiency was analyzed independently because processing speed is an important consideration in practical computer graphics workflows. Significant variation was observed among the investigated frameworks, reflecting differences in their underlying generation strategies.

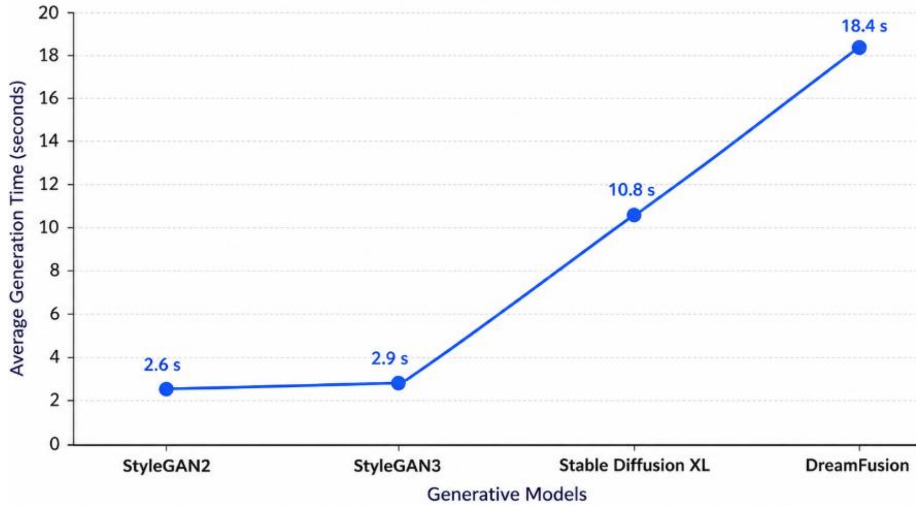

Figure 2. Comparison of average generation time

```
import matplotlib.pyplot as plt

models = [
    "StyleGAN2",
    "StyleGAN3",
    "Stable Diffusion XL",
    "DreamFusion"
]
generation_time = [2.6, 2.9, 10.8, 18.4]

plt.figure(figsize=(8,5))
plt.plot(
    models,
    generation_time,
    marker="o",
    linewidth=2
)
plt.title("Average Generation Time")
plt.xlabel("Generative Models")
plt.ylabel("Time (seconds)")
for i, value in enumerate(generation_time):
    plt.text(
        i,
        value+0.4,
        f"{value:.1f}",
        ha="center",
        fontsize=9
    )
plt.grid(alpha=0.3)
plt.tight_layout()
plt.show()
```

The results presented in Figure 2 indicate that GAN-based architectures maintain a considerable advantage in inference speed. StyleGAN2 and StyleGAN3 generated graphical objects within only a few seconds, making them highly suitable for interactive graphics systems and real-time applications. By contrast, Stable Diffusion XL and DreamFusion required longer processing because of their iterative denoising procedures; however, this additional computational effort produced higher image quality and improved structural consistency. Overall, the experimental findings suggest that GANs remain preferable for applications emphasizing rapid content generation, whereas diffusion models provide greater visual realism and reconstruction accuracy when image quality is the primary objective. Consequently, selecting an appropriate generative framework should be guided by the intended application requirements, available computational resources, and desired balance between efficiency and visual fidelity [3–8].

DISCUSSION

The experimental comparison performed in this study highlights the distinct characteristics of GAN-based and diffusion-based generative models when applied to computer graphics object generation. Although both approaches are capable of producing realistic graphical content, the obtained results indicate that they achieve this objective through different computational strategies and exhibit different trade-offs between visual quality and processing efficiency. These findings reinforce recent research emphasizing that the selection of a generative model should be guided by application-specific requirements rather than by overall performance alone [3–5].

The diffusion-based frameworks consistently achieved superior scores in image quality, structural consistency, and texture fidelity. DreamFusion demonstrated the strongest overall performance, producing objects with highly detailed geometry and visually coherent surface characteristics. Stable Diffusion XL also generated realistic graphical content while maintaining comparatively lower computational demands than DreamFusion. The iterative denoising mechanism employed by diffusion models appears to contribute significantly to their ability to preserve semantic information and reconstruct fine visual details. Similar observations have been reported in recent investigations, where diffusion-based generation outperformed adversarial learning in terms of visual realism and generation stability [4–6].

Conversely, the GAN architectures showed clear advantages with respect to computational efficiency. Both StyleGAN2 and StyleGAN3 completed the object generation process substantially faster than the diffusion models, making them well suited for applications that require rapid image synthesis or real-time interaction. The adversarial training mechanism enables GANs to generate outputs directly without performing multiple reconstruction iterations, thereby reducing inference time. Nevertheless, the experiments also indicated that GAN-generated objects occasionally exhibited reduced structural diversity and less consistent representation of complex geometric details. These observations correspond with previously reported limitations of adversarial learning, particularly regarding mode collapse and restricted output variability during generation [2,3].

Another important outcome of this research concerns the stability of the investigated generative approaches. Diffusion models maintained more consistent performance across different object categories and produced fewer visual artifacts during reconstruction. Although this increased stability required longer processing time and greater GPU memory consumption, the resulting improvement in object quality may outweigh the additional computational cost for professional computer graphics applications where visual accuracy is the primary objective. In contrast, GANs remain advantageous in environments where computational resources are limited or where rapid content generation is essential [5–7,10].

The present investigation also demonstrates the importance of standardized benchmarking when evaluating modern generative AI technologies. Many previous studies have introduced novel architectures without applying common datasets, unified evaluation metrics, or identical computational environments, making direct comparison difficult. By employing the same benchmark images, equivalent execution conditions, and consistent quantitative performance indicators, this research provides a more objective assessment of the relative strengths and weaknesses of GAN and diffusion models. Such methodological consistency contributes to improving the reproducibility and practical relevance of comparative research in intelligent computer graphics [7,8].

Despite the encouraging findings, several limitations remain. The experiments focused primarily on individual object generation and did not investigate more challenging tasks such as large-scale scene synthesis, dynamic animation, or interactive multimodal environments. In addition, the evaluation relied mainly on objective numerical indicators, while subjective assessment by experienced computer graphics designers was beyond the scope of the current study. Future investigations should therefore incorporate larger benchmark datasets, recently introduced foundation models, human-centered evaluation procedures, and perceptual quality assessment to obtain a more comprehensive understanding of generative AI performance.

The results indicate that GANs and diffusion models should be viewed as complementary technologies rather than direct alternatives. GAN-based frameworks continue to provide excellent computational efficiency and rapid inference, whereas diffusion models offer superior visual fidelity, semantic consistency, and generation stability. Continued progress in generative Artificial Intelligence, neural rendering, multimodal learning, and optimization algorithms is expected to further improve both technologies, enabling more intelligent, scalable, and efficient computer graphics object generation across a broad range of scientific and industrial applications [4–8].

CONCLUSION

The findings of this study demonstrate that both Generative Adversarial Networks (GANs) and diffusion models have become important components of modern Artificial Intelligence for computer graphics object generation. Although the two approaches are based on different learning mechanisms, each provides valuable capabilities for producing realistic digital content with a high degree of automation. The experimental evaluation confirms that recent progress in generative AI has significantly improved the quality of synthetic graphical objects while reducing the amount of manual modeling traditionally required in digital content creation [3,4,9].

The comparative analysis showed that diffusion-based frameworks achieved superior performance in terms of visual realism, structural consistency, and texture quality. DreamFusion produced the highest overall results, while Stable Diffusion XL offered an effective balance between image quality and computational demand. In contrast, StyleGAN2 and StyleGAN3 demonstrated clear advantages in inference speed and computational efficiency, making them more appropriate for applications where rapid generation and real-time performance are essential. These results indicate that the selection of a generative model should be guided by the intended application scenario, available hardware resources, and the desired balance between generation quality and processing efficiency rather than relying solely on a single evaluation criterion [5–7].

Another significant outcome of this research is the development of a unified evaluation framework that enables objective comparison between adversarial and diffusion-based generation methods. The integration of standardized datasets, common quantitative performance indicators, and reproducible statistical analysis provides a transparent basis for assessing the capabilities of contemporary generative AI models. Such a benchmarking strategy contributes to improving the consistency and reliability of comparative research in computer graphics while facilitating informed selection of appropriate generation technologies.

Despite the encouraging experimental results, additional investigation is required to further enhance the practical applicability of generative AI. Future research should evaluate recently developed foundation models, larger and more diverse benchmark datasets, multimodal generation techniques, and perceptual assessment involving domain experts. Continued advances in deep learning architectures, neural rendering, and optimization algorithms are expected to reduce computational requirements while further improving generation quality. Consequently, GANs and diffusion models will continue to play complementary roles in the evolution of intelligent computer graphics systems, supporting increasingly efficient, scalable, and realistic digital object generation across scientific, engineering, industrial, and creative applications [2,4,8].

REFERENCES:

1. A. Bozorov and oth.. Modern technologies of virtual reality– a new multimedia opportunities. EURASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES (T. 2, Выпуск 1, cc. 85–90). Zenodo. https://ddoi.org/10.5281/zenodo.7251370
2. A. Bozorov and oth.. "COMPUTER GRAPHICS IN TECHNICAL DISCIPLINES. B EURASIAN JOURNAL OF ACADEMIC RESEARCH (T. 4, Выпуск 10, cc. 21-27). Zenodo." 2024, https://doi.org/10.5281/zenodo.13898180.
3. A. Bozorov and oth.. PROBLEMS IN THE ACTIVITIES OF WORKSHOPS FOR REPAIRING (HUSENOLD, COMPUTER, VIDEO) EQUIPMENT IN UZBEKISTAN. IN THE WORLD OF SCIENCE AND EDUCATION Учредители: Международный исследовательский центр" Endless Light in Science" (T. 2, Выпуск 1, cc. 45–50). Zenodo. https://doi.org/10.5281/zenodo.17090018
4. A. Bozorov and oth..Methods for plotting function graphs in computers using modern software and programming languages.Published in ACADEMICIA An International Multidisciplinary Research Journal ISSN: 2249-7137, Vol. 11 Issue 6, JUNE 2021. P. 321-329, India, Impact Factor: SJIF 2021 = 7.492, https://saarj.com/academicia-view-journal-current-issue/
5. A. Bozorov and oth.. CONVERT IMAGES TO PDF IN C# // EJMTCS. 2024. №10. URL: https://cyberleninka.ru/article/n/convert-images-to-pdf-in-c (дата обращения: 27.06.2026).
6. A. Bozorov and oth.. METHODS FOR GRAPHING FUNCTIONS IN COMPUTERS USING WEB TECHNOLOGIES: METHODS FOR GRAPHING FUNCTIONS IN COMPUTERS USING WEB TECHNOLOGIES //Journal of Information Computational Science. – 2021. – №. 1.
7. A. Bozorov and oth.. Konstitutsiya-inson huquqlari va huquqiy kafolatlar. Proceedings of International Educators Conference, (T. 1, Выпуск 2, cc. 196–198).
8. A. Bozorov and oth..Multimedia surveillance cameras and their features in using. International Journal of Innovations in Engineering Research and Technology (IJIERT), Vol 9, No 10 (2022), p.29–34. https://repo.ijiert.org/index.php/ijiert/article/view/3379, doi.org/10.17605/OSF.IO/4EV75
9. A. Bozorov and oth.. The Audio- Is of the Main Components of Multimedia Technologies. International Journal on Integrated Education, Vol 5, Issue 5. e-ISSN : 26203502. May (2022), p.263–268. https://creativecommons.org/licenses/by/4.0/
10. A. Bozorov and oth.. APPLICATION OF YUV COLOR MODEL IN REAL-TIME IMAGE PROCESSING. Science and innovation, ISSN : 2181-3337. 2024, p.231–238. https://elibrary.ru/ip_restricted.asp?rpage=https%3A%2F%2Felibrary%2Eru%2Fitem%2Easp%3Fid%3D75188650