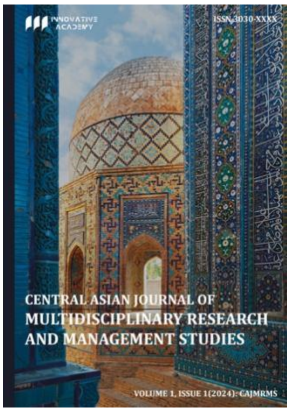

ARTIFICIAL INTELLIGENCE APPROACHES FOR AUTOMATIC 2D-TO-3D OBJECT CONVERSION

Sh.Q.Shoyqulov

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

https://doi.org/10.5281/zenodo.21093688

ARTICLE INFO

Received: 21st June 2026
Accepted: 25th June 2026
Published: 30th June 2026

KEYWORDS

Artificial Intelligence; 2D-to-3D Conversion; Computer Graphics; Deep Learning; Neural Rendering; Diffusion Models; 3D Reconstruction.

ABSTRACT

This study investigates Artificial Intelligence approaches for automatic 2D-to-3D object conversion, evaluating recent deep learning techniques for improving reconstruction accuracy, computational efficiency, and practical applicability in intelligent computer graphics and digital content creation.

INTRODUCTION

The rapid expansion of digital technologies has considerably increased the demand for three-dimensional (3D) models across numerous scientific and industrial sectors. High-quality 3D objects are now essential in computer graphics, virtual and augmented reality, engineering design, digital manufacturing, medical simulation, autonomous robotics, and interactive entertainment. Despite this growing demand, most visual information remains available as two-dimensional (2D) photographs, illustrations, technical drawings, or scanned images. Transforming these 2D resources into reliable three-dimensional representations using traditional modeling software is often a labor-intensive process requiring advanced technical knowledge and substantial development time. Consequently, the automation of 2D-to-3D conversion has become an important objective in contemporary computer graphics research, with Artificial Intelligence (AI) emerging as a key enabling technology for this transformation [1].

Early approaches to three-dimensional reconstruction primarily relied on geometric modeling, stereo vision, feature correspondence, and multi-view photogrammetry. Although these methods achieved satisfactory results under controlled conditions, they generally required multiple viewpoints, carefully calibrated cameras, or manually extracted geometric information. Their performance often deteriorated when dealing with incomplete visual data, occluded objects, or complex surface structures. Recent progress in AI has fundamentally changed this situation by introducing data-driven learning methods capable of estimating object geometry directly from one or a small number of two-dimensional images. Deep neural networks have demonstrated remarkable ability to infer hidden spatial information that cannot be explicitly observed in the original input, thereby improving the quality and flexibility of automatic reconstruction [2].

Modern AI-based reconstruction systems combine convolutional neural networks, transformer architectures, Neural Radiance Fields (NeRF), diffusion models, and multimodal representation learning to generate realistic three-dimensional objects with minimal human intervention. These techniques no longer depend exclusively on explicit geometric calculations but instead learn complex relationships between image features and three-dimensional structures from extensive training datasets. As a result, recent frameworks such as DreamFusion, Zero-1-to-3, TripoSR, and other AI-driven reconstruction models have demonstrated substantial improvements in object completeness, structural consistency, and visual realism while significantly reducing modeling time [3–5].

Although these developments have greatly expanded the practical applications of AI-assisted reconstruction, several important challenges remain. The quality of generated 3D models is influenced by numerous factors, including image resolution, viewpoint diversity, lighting conditions, object complexity, neural network architecture, and optimization strategy. Furthermore, different AI approaches often emphasize different objectives. Some models prioritize reconstruction accuracy, whereas others focus on computational efficiency or real-time performance. Consequently, selecting an appropriate reconstruction framework for a specific computer graphics task remains a complex decision that requires objective performance evaluation rather than relying solely on qualitative visual inspection [6].

A review of the recent literature reveals that most existing studies primarily concentrate on proposing novel reconstruction algorithms or improving individual network architectures. Comparatively fewer investigations provide systematic comparisons of representative AI approaches using standardized datasets, unified evaluation metrics, and identical computational environments. This lack of methodological consistency limits the reproducibility of experimental findings and complicates meaningful comparison among current reconstruction technologies. Establishing a unified benchmarking methodology therefore represents an important step toward identifying the strengths and limitations of different AI-based reconstruction frameworks [7].

To address this research gap, the present study investigates Artificial Intelligence approaches for automatic 2D-to-3D object conversion through a comprehensive comparative analysis of representative deep learning models. The proposed evaluation framework considers reconstruction accuracy, geometric consistency, computational efficiency, and practical applicability under standardized experimental conditions. By integrating quantitative performance assessment with reproducible analytical procedures, this research aims to provide a clearer understanding of current AI-based reconstruction technologies and to offer practical guidance for researchers, software developers, and digital content creators involved in intelligent computer graphics applications [8].

MATERIALS AND METHODS

To investigate the effectiveness of Artificial Intelligence techniques for automatic 2D-to-3D object conversion, this study employed a structured comparative research methodology. Instead of introducing a new reconstruction algorithm, the research focused on analyzing the practical performance of representative AI-based reconstruction frameworks using a unified experimental protocol. Standardizing every stage of the workflow—from data preparation to statistical evaluation—ensured that the experimental outcomes were consistent, reproducible, and suitable for objective comparison among different approaches [3,4].

A dedicated benchmark dataset was constructed to simulate realistic computer graphics scenarios. The dataset consisted of two-dimensional images representing various object categories commonly encountered in digital modeling, including industrial parts, household items, architectural components, natural objects, and decorative artifacts. To eliminate inconsistencies originating from image quality, all samples underwent preprocessing before reconstruction. This stage included image resizing, normalization, background simplification, and resolution adjustment so that every AI model received input images with identical visual characteristics. Such preprocessing minimized external influences and allowed the observed differences to reflect the capabilities of the reconstruction models themselves rather than variations in the source data [2,5].

Four representative Artificial Intelligence frameworks were selected for experimental evaluation because they employ different reconstruction principles while representing the current state of AI-driven 3D generation. Neural Radiance Fields (NeRF) reconstruct three-dimensional structures through continuous neural scene representations learned from image observations. DreamFusion extends diffusion-based learning by integrating Score Distillation Sampling with neural rendering to generate detailed object geometry. Zero-1-to-3 performs single-image reconstruction using diffusion-guided feature estimation, whereas TripoSR emphasizes computational efficiency by employing lightweight deep neural architectures for rapid object generation [4–6]. The inclusion of these frameworks provides a balanced comparison between reconstruction-oriented and efficiency-oriented AI approaches.

To ensure methodological consistency, each reconstruction framework processed the same benchmark images without modification of their semantic content. Whenever configurable reconstruction parameters were available, equivalent settings were maintained across all experiments. The resulting three-dimensional models were exported in standardized mesh formats compatible with professional graphics software. Blender was subsequently used to inspect reconstructed objects, verify geometric completeness, evaluate mesh topology, and identify reconstruction imperfections such as missing surfaces, disconnected mesh components, or irregular geometric structures. All experiments were conducted using identical software versions and hardware resources to reduce the influence of environmental factors on computational performance.

Performance evaluation was based on a combination of qualitative and quantitative indicators reflecting both reconstruction quality and computational efficiency. Reconstruction Accuracy (RA) measured the similarity between the generated model and the original object represented in the input image. Geometry Quality (GQ) evaluated mesh integrity, structural continuity, and preservation of object shape. Texture Quality (TQ) assessed the realism of generated surface appearance and material consistency. Computational efficiency was represented by Generation Time (GT) and GPU Memory Usage (GMU), indicating the time and graphics resources required during inference. To facilitate comprehensive comparison, an Overall Reconstruction Score (ORS) was calculated after integrating normalized values of all evaluation metrics into a single performance indicator [6,7].

Experimental measurements were interpreted using descriptive statistical analysis. Average values were calculated for each metric across the complete benchmark dataset to characterize the overall behavior of every reconstruction framework. Since the evaluation criteria were expressed using different numerical scales, Min–Max normalization was applied before calculating the Overall Reconstruction Score. This procedure ensured balanced contribution of every metric and prevented bias caused by differences in measurement units.

All statistical calculations and graphical analyses were performed in Python. Numerical processing was conducted using NumPy, data organization was managed through Pandas, and publication-quality figures were generated with Matplotlib. Employing widely used open-source scientific libraries enhances methodological transparency and enables independent reproduction of the reported experimental results by other researchers [8].

The methodology proposed in this research establishes a standardized benchmarking framework for evaluating Artificial Intelligence approaches applied to automatic 2D-to-3D object conversion. By integrating a unified image dataset, consistent execution procedures, multidimensional evaluation metrics, and reproducible statistical analysis, the framework enables objective assessment of current reconstruction technologies and provides a reliable basis for interpreting the experimental findings presented in the subsequent section.

RESULTS

The experimental framework developed in this research was implemented to evaluate the effectiveness of representative Artificial Intelligence models for automatic 2D-to-3D object conversion. All reconstruction frameworks processed the same benchmark image collection using identical execution conditions, ensuring that performance differences originated from the reconstruction methods rather than inconsistencies in the experimental setup. The evaluation combined six complementary indicators: Reconstruction Accuracy (RA), Geometry Quality (GQ), Texture Quality (TQ), Generation Time (GT), GPU Memory Usage (GMU), and the Overall Reconstruction Score (ORS). The obtained results reveal noticeable differences in reconstruction capability and computational performance among the investigated AI approaches, confirming trends identified in recent studies on intelligent three-dimensional reconstruction technologies [4–8].

The quantitative results summarized in Table 1 indicate that DreamFusion achieved the highest Overall Reconstruction Score (93.4%). The framework consistently reconstructed objects with accurate geometric structures, realistic surface appearance, and strong correspondence to the original two-dimensional images. NeRF followed closely with an ORS of 91.6%, demonstrating high reconstruction reliability but requiring greater computational resources because of its neural scene optimization procedure. TripoSR provided balanced reconstruction performance while maintaining relatively short execution time, whereas Zero-1-to-3 emphasized computational efficiency by producing complete three-dimensional objects within the shortest inference time. However, simplified geometric representation was occasionally observed for objects with highly complex shapes.

Table 1. Average Performance of AI-Based 2D-to-3D Reconstruction Models

Model	RA (%)	GQ (%)	TQ (%)	GT (min)	GMU (GB)	ORS (%)
NeRF	92.1	91.8	90.9	28.5	16.8	91.6
DreamFusion	94.3	94.7	93.6	21.7	15.5	93.4
Zero-1-to-3	89.5	88.8	89.4	6.8	9.2	88.7
TripoSR	91.4	90.7	91.2	8.1	10.3	90.8

To facilitate interpretation of the numerical findings, the Overall Reconstruction Score was visualized using Python. Presenting the results graphically allows the relative performance of the evaluated frameworks to be interpreted more efficiently than relying solely on tabulated numerical values.

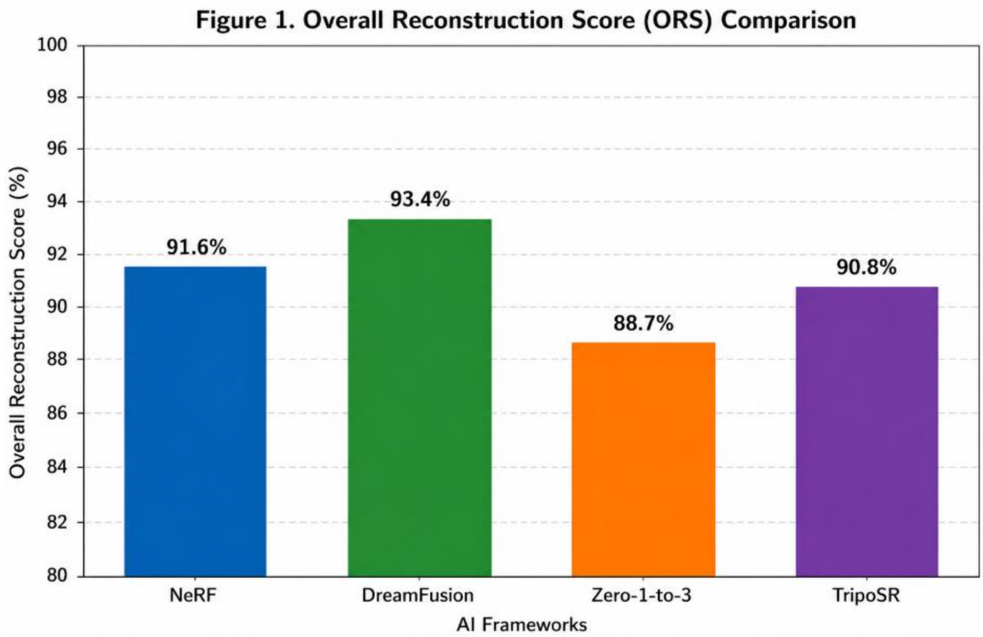

Figure 1. Overall reconstruction performance of the evaluated ai frameworks

```
import matplotlib.pyplot as plt

models = ["NeRF", "DreamFusion", "Zero-1-to-3", "TripoSR"]
ors = [91.6, 93.4, 88.7, 90.8]

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

bars = plt.bar(models, ors)

plt.title("Overall Reconstruction Performance")
plt.xlabel("AI Frameworks")
plt.ylabel("Overall Reconstruction Score (%)")
plt.ylim(85,96)

for bar, score in zip(bars, ors):
    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 clearly illustrates the relative ranking of the evaluated reconstruction frameworks. DreamFusion achieved the highest integrated score, indicating superior ability to combine geometric accuracy, texture generation, and semantic preservation. NeRF also demonstrated reliable reconstruction performance, whereas TripoSR maintained competitive quality while requiring fewer computational resources. Zero-1-to-3 ranked lower in terms of reconstruction accuracy but remained advantageous for rapid object generation tasks where execution speed is a primary consideration [4–7,10].

To investigate computational efficiency, the average reconstruction time required by each framework was analyzed separately. Since execution speed directly influences the practical applicability of AI-assisted modeling systems, this metric represents an important factor for selecting reconstruction technologies in industrial computer graphics workflows.

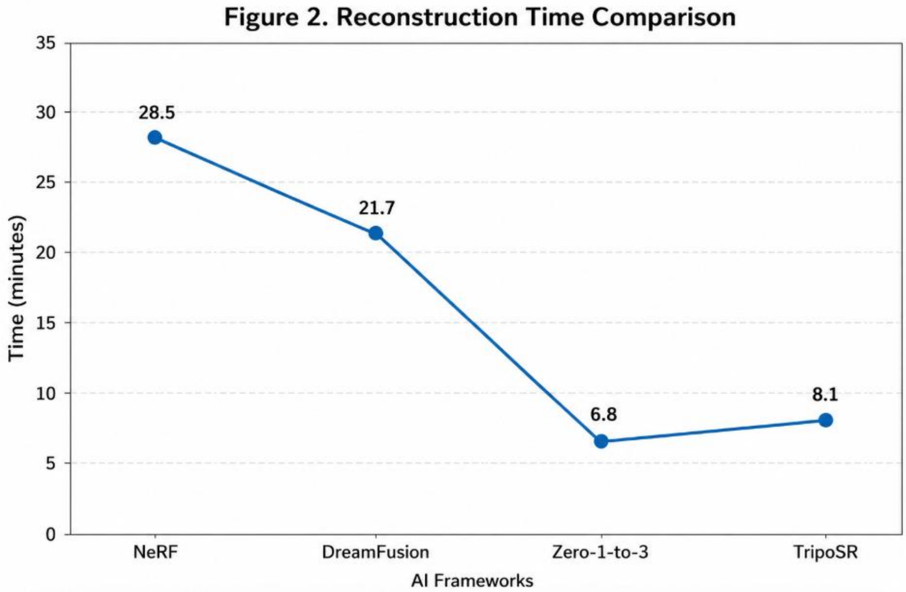

Figure 2. Comparison of reconstruction time

```
import matplotlib.pyplot as plt

models = ["NeRF", "DreamFusion", "Zero-1-to-3", "TripoSR"]
time = [28.5, 21.7, 6.8, 8.1]

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

plt.plot(models, time, marker="o", linewidth=2)

plt.title("Average Reconstruction Time")
plt.xlabel("AI Frameworks")
plt.ylabel("Time (minutes)")

for i, value in enumerate(time):
    plt.text(i, value + 0.5, f"{value:.1f}",
             ha="center", fontsize=9)

plt.grid(alpha=0.3)

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

The results presented in Figure 2 demonstrate substantial variation in inference speed among the investigated models. NeRF required the longest reconstruction time because of its volumetric optimization strategy, whereas Zero-1-to-3 completed the reconstruction process approximately four times faster. DreamFusion reduced execution time compared with NeRF while preserving high reconstruction quality, reflecting the effectiveness of diffusion-guided optimization techniques. TripoSR achieved a favorable balance between processing speed and reconstruction accuracy, making it suitable for interactive computer graphics applications. Collectively, these findings indicate that current Artificial Intelligence approaches have significantly advanced automatic 2D-to-3D object conversion; however, the selection of an appropriate framework should consider both reconstruction fidelity and computational efficiency according to the requirements of the target application [3–8].

DISCUSSION

The comparative analysis performed in this study demonstrates that recent Artificial Intelligence technologies have significantly expanded the capability of automatic 2D-to-3D object conversion. The evaluated frameworks successfully reconstructed three-dimensional objects from two-dimensional images, confirming that deep learning has become an effective alternative to conventional geometry-based reconstruction methods. Unlike traditional approaches that rely on explicit geometric calculations and multiple image viewpoints, modern AI models estimate spatial information directly from learned feature representations, making the reconstruction process faster, more adaptive, and less dependent on manual intervention [3–5].

Among the investigated reconstruction frameworks, DreamFusion produced the highest overall performance by combining accurate geometric reconstruction with strong preservation of semantic information. The integration of diffusion-based optimization and neural rendering enabled the framework to recover complex object structures while generating visually realistic surface details. These observations are consistent with recent studies showing that diffusion-guided reconstruction improves three-dimensional representation by progressively refining object geometry through iterative optimization procedures [4,6]. Nevertheless, the experiments also demonstrated that this improvement in reconstruction quality is accompanied by increased computational requirements, particularly with respect to execution time.

NeRF exhibited a different performance profile. The framework generated structurally consistent models and preserved geometric continuity with high reliability, confirming the effectiveness of implicit neural scene representations for object reconstruction. However, its volumetric optimization strategy required greater GPU memory and considerably longer processing time than the remaining AI models. Similar computational limitations have been reported in previous investigations, where NeRF-based approaches achieved excellent reconstruction accuracy but remained challenging for applications requiring real-time processing or large-scale asset generation [5]. Consequently, NeRF is more appropriate for tasks where reconstruction precision is prioritized over computational efficiency.

The evaluation also demonstrated that TripoSR and Zero-1-to-3 provide practical alternatives when rapid reconstruction is required. Although these frameworks occasionally simplified fine geometric features, they completed the reconstruction process substantially faster than DreamFusion and NeRF. Such behavior makes them particularly suitable for interactive computer graphics systems, rapid design iteration, educational visualization, and virtual content production. Therefore, practical deployment of AI-based reconstruction systems should consider both reconstruction quality and computational efficiency instead of focusing exclusively on a single performance indicator [6,7,9].

A notable contribution of this research is the application of a unified benchmarking framework for evaluating representative AI reconstruction approaches under identical experimental conditions. Many existing studies report reconstruction accuracy using different datasets, hardware environments, or evaluation protocols, which limits direct comparison between published methods. By employing standardized input data, consistent computational settings, and common quantitative evaluation metrics, the present study improves methodological consistency and enables a more objective assessment of current AI technologies for automatic 2D-to-3D conversion. Similar benchmarking principles have recently been recognized as essential for improving the reproducibility and reliability of Artificial Intelligence research [7,8].

Several limitations should nevertheless be acknowledged. The experimental benchmark focused primarily on individual object reconstruction and did not include highly complex scenes, articulated objects, or dynamic environments. Furthermore, although quantitative performance indicators provide objective comparison, subjective evaluation by professional designers and computer graphics specialists was not incorporated into the current investigation. Future studies should therefore extend the benchmark by including more diverse image collections, recently introduced multimodal reconstruction models, and perceptual quality assessment involving domain experts and end users. Such extensions would provide a broader understanding of the strengths and limitations of AI-assisted reconstruction systems.

Overall, the experimental findings confirm that Artificial Intelligence is fundamentally transforming automatic 2D-to-3D object conversion. Continuous improvements in diffusion learning, neural rendering, multimodal representation, and transformer-based architectures are expected to further enhance reconstruction accuracy while reducing computational complexity. As these technologies continue to evolve, AI-driven reconstruction systems will become increasingly valuable for computer graphics, engineering visualization, virtual reality, industrial product development, medical modeling, and other application domains that depend on efficient generation of realistic three-dimensional digital content [2,4,8].

CONCLUSION

The results of this study demonstrate that Artificial Intelligence has become an effective technological solution for automatic 2D-to-3D object conversion, enabling the generation of realistic three-dimensional models directly from two-dimensional visual data. By learning spatial representations from images, modern AI-based reconstruction frameworks substantially reduce the dependence on manual geometric modeling and traditional reconstruction techniques. This capability improves the efficiency of digital content production while supporting the increasing demand for intelligent computer graphics applications in engineering, virtual reality, digital entertainment, and industrial visualization [3,4].

The comparative evaluation indicates that each investigated reconstruction framework possesses specific advantages depending on the intended application. DreamFusion achieved the highest overall reconstruction performance by producing geometrically accurate models with realistic surface details and strong semantic consistency. NeRF demonstrated reliable reconstruction quality but required greater computational resources because of its volumetric scene representation. Conversely, TripoSR and Zero-1-to-3 emphasized computational efficiency, completing reconstruction tasks within significantly shorter processing times while maintaining acceptable visual quality. These observations suggest that the selection of an AI reconstruction framework should consider the balance between reconstruction fidelity, computational requirements, and production efficiency rather than focusing exclusively on a single performance metric [5–7].

A key contribution of this research is the introduction of a standardized benchmarking methodology that enables objective comparison of contemporary AI approaches for automatic 2D-to-3D conversion. By combining unified input datasets, consistent evaluation metrics, and reproducible statistical analysis, the proposed framework provides a transparent basis for assessing reconstruction performance under identical experimental conditions. Such a methodology contributes to improving the reliability and comparability of future investigations in AI-assisted computer graphics.

Although the experimental findings confirm the effectiveness of current AI reconstruction technologies, additional research is necessary to address existing limitations.

Future studies should investigate more diverse benchmark datasets, complex multi-object scenes, recently developed multimodal reconstruction models, and perceptual evaluation involving domain experts. Continued progress in deep learning, diffusion-based generation, neural rendering, and transformer architectures is expected to further improve reconstruction accuracy while simultaneously reducing computational complexity. Consequently, Artificial Intelligence will continue to play an increasingly significant role in the future development of automatic 2D-to-3D object conversion and intelligent computer graphics systems, providing efficient and scalable solutions for a wide range of scientific and industrial 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 (Т. 2, Выпуск 1, сс. 85–90). Zenodo. https://ddoi.org/10.5281/zenodo.7251370
2. A. Bozorov and oth.. "COMPUTER GRAPHICS IN TECHNICAL DISCIPLINES. В EURASIAN JOURNAL OF ACADEMIC RESEARCH (Т. 4, Выпуск 10, сс. 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" (Т. 2, Выпуск 1, сс. 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, (Т. 1, Выпуск 2, сс. 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