Archives available at journals.mriindia.com

International Journal on Advanced Computer Engineering and Communication Technology

ISSN: 2278-5140
Volume 14 Issue 01, 2025

Advancements in AI for Text-to-3D Model Generation: A Comparative Study of Meshy.ai and LumaLabs.ai Genie

Mrigyisha Sawant
U.G. Student, Department of Artificial Intelligence and Data Science, Dr. D Y Patil College of Engineering & Innovation, Varale, Talegaon, Pune, India

Peer Review Information

Submission: 21 Feb 2025
Revision: 25 March 2025
Acceptance: 30 April 2025

Keywords

AI-Driven 3D Modeling
Text-To-3D Model
Geometric Modeling

Abstract

This paper examines the effectiveness of AI-driven 3D model generation by comparing two prominent text-to-3D platforms. This study evaluates how well these models translate textual prompts into detailed and structurally systematic 3D outputs across various object categories, including organic forms, geometric structures, and humanoid characters. Each generated model is assessed based on key factors such as generation speed, shape accuracy, textural accuracy, mesh quality, and realism. The comparative analysis highlights the strengths and limitations of these AI models, particularly in handling fine textures, maintaining geometric consistency, and adapting to both simple and complex prompts provided to them. The study provides insights into the current capabilities and potential improvements needed for AI-based 3D model generation in creative and engineering applications.

INTRODUCTION

The rapid advancements in Artificial Intelligence have significantly transformed content generation across multiple domains, such as text, images, and 3D models. While AI-driven image generation has shown remarkable progress in translating textual prompts into detailed and creative visual representations, AI-powered 3D model generation remains an evolving field. These models make use of deep learning, computer vision, and neural rendering techniques to construct three-dimensional representations from textual prompts, enabling new possibilities in digital content creation, gaming, virtual reality, and design automation while reducing the manual efforts needed and allowing for a diverse user base in 3D modelling.
Recent developments in text-to-3D generation have led to the emergence of platforms such as Meshy.ai and LumaLabs.ai Genie, which offer automated 3D asset creation from simple text inputs. These tools promise to bridge the gap between conceptualization and 3D modeling, allowing users to generate complex, textured, and structured 3D assets without extensive manual modeling. However, while these AI-powered platforms have shown promise, their effectiveness in generating high-fidelity and structurally coherent 3D models varies depending on the nature of the input prompt, computational trade-offs, and the underlying algorithms.
This paper presents a comparative study of the platforms mentioned, evaluating their performance across different categories of 3D model generation. The analysis examines factors such as generation speed, accuracy, texture quality, and mesh integrity based on structured text prompts ranging from simple geometric shapes to complex, stylized human-like figures. By assessing these factors, this research aims to provide insights into the strengths and limitations of current AI-driven 3D model generators.

The remainder of this paper is organized as follows: Section II provides an overview of AI-driven 3D model generation techniques. Section III outlines the methodology used for the comparative analysis. Section IV presents the experimental results and discussions and Section V provides the conclusion to this comparative study.

LITERATURE SURVEY

Text-to-3D generation has evolved significantly, with many methods focusing on improving realism, efficiency, and control over generated shapes. Many early approaches relied on direct optimization techniques, while newer methods integrate diffusion models and neural representations.

Poole et al. [1] introduced DreamFusion, which uses a 2D diffusion model to optimize 3D objects, generating more realistic and detailed results. Lin et al. [2] improved upon this with Magic3D, refining resolution and structure using a two-stage process. Other works, like Michel et al. [3], focused on neural stylization, where Text2Mesh allows fine-tuning textures based on provided prompts.

Recent studies have also explored bench-marking and evaluation. Zhang et al. [4] and Su et al. [5] introduced structured ways to compare different methods, but in this study, we opt for a more practical approach—direct visual inspection and qualitative assessment.

Despite advancements, challenges remain in fine-grained control, mesh quality, and efficiency. This work explores how well existing models generate 3D shapes based on simple visual comparison.

METHODOLOGY

Text-to-3D generation platforms, Meshy and Lumalabs Genie, are evaluated through controlled experiments to compare their capabilities in generating high-quality 3D models based on model descriptions. This study analyzes the strengths, weaknesses, and practical applications.

Three categories of prompts are designed to assess different aspects of 3D model generation:

1. Simple Object: A well-defined shape with texture (“An orange kitten”).
2. Geometric Structure: A spatially arranged shape (“Three cubes stacked on top of each other, each slightly rotated”).
3. Human-like Model: A complex organic form requiring anatomical accuracy (“Anime girl with long black hair wearing a school uniform”).

These prompts evaluate the ability of each platform to handle the different levels of detail and structure.

Each text prompt undergoes a two-stage workflow for both platforms. Meshy outputs four untextured mesh models, allowing manual texture application in a secondary process. Lumalabs Genie generates four fully textured models by default, with options for high-resolution refinement for each. For the high resolution refinement, Meshy enables manual selection of a model for high-resolution texturing while Lumalabs Genie permits up-scaling to enhance texture detail.

The generated models are assessed based on qualitative criteria:

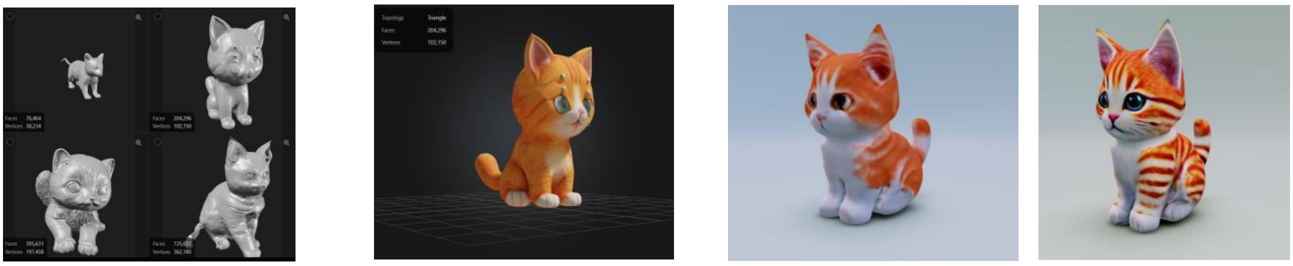
1. Generation Speed: Time taken to produce low and high-resolution outputs.
2. Accuracy: The degree of correspondence between the generated model and input text.
3. Texture Quality: The sharpness, detail, and realism of applied textures.
4. Mesh Quality: Structural accuracy, polygon density, and smoothness.
5. Realism: Evaluated for organic models, particularly human-like figures.

This evaluation provides a practical comparison, focusing on usability rather than pure benchmarking.

EXPERIMENTAL RESULTS 1. Simple Object: An Orange Kitten

The evaluation of a simple organic model, an orange kitten, highlights distinct differences in speed, accuracy, texture quality, and mesh structure between Meshy and Lumalabs Genie. Lumalabs Genie demonstrated a faster initial generation speed for low-resolution models, while Meshy excelled in high-resolution refinement and texturing.

In terms of accuracy, Lumalabs Genie produced a more recognizable and well-defined kitten shape, particularly in both low- and high-resolution outputs. However, Meshy provided superior base texture quality, delivering finer details and better shading consistency. When highresolution refinement was applied, Lumalabs Genie surpassed Meshy in texture quality. Meshwise, Meshy generated a more structured and detailed model with improved topology. Figure 1 illustrates the outputs generated from the prompt "An orange kitten" across both platforms.



(a) (b) (c)
(d)

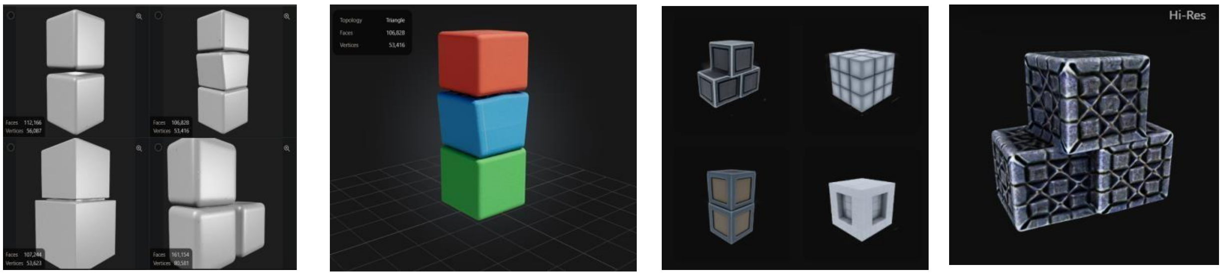
Fig. 1. Orange Kitten Model Generation. Prompt: "An orange kitten" (a) Untextured orange kitten model generated by Meshy. (b) High-resolution textured orange kitten model generated by Meshy. (c) Orange kitten model generated by Lumalabs Genie. (d) High-resolution textured orange kitten model generated by Lumalabs Genie.

2.

Geometric Structure: Three Stacked Cubes with slight rotations.

The geometric prompt, featuring three stacked cubes with slight rotations, tested each platform's ability to generate precise, structured forms. Lumalabs Genie was faster in producing the initial low-resolution model, while Meshy performed better in high-resolution texturing and refinement.

In terms of accuracy, Meshy significantly outperformed Lumalabs Genie, producing a more geometrically precise arrangement. Additionally, Meshy provided higher texture quality, ensuring greater consistency across surfaces. The structural integrity of the mesh was also superior in Meshy's output, with well-defined edges and improved polygon distribution. Figure 2 illustrates the outputs generated from the prompt "Three cubes stacked on top of each other, each slightly rotated" across both platforms



(a) (b) (c)
(d)

Fig. 2 Cube Model Generation. Prompt: "Three cubes stacked on top of each other, each slightly rotated" (a) Untextured cube model generated by Meshy. (b) High-resolution textured cube model generated by Meshy. (c) Cube model generated by Lumalabs Genie. (d) High-resolution textured cube model generated by Lumalabs Genie.

3. Human-like Model: Anime Girl with Long Black Hair

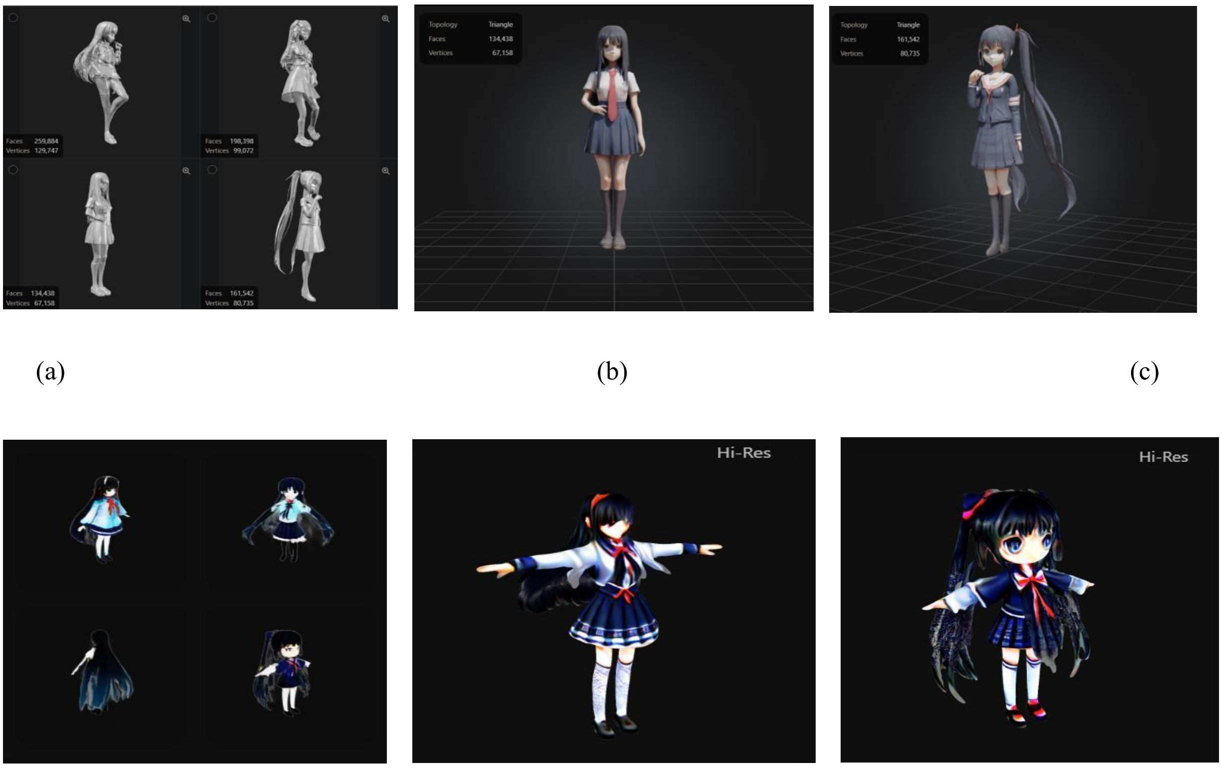
The most complex evaluation involved generating an anime-style human model, tested with two variations:

1. Long hair (loose)
2. Long hair in two ponytails

Lumalabs Genie was faster in generating the initial low-resolution model with simple textures, but Meshy performed significantly better in high-resolution refinement and texturing. Meshy vastly outperformed Lumalabs Genie in accuracy, delivering a significantly more anatomically correct and proportional human figure. Texture quality was also notably higher in Meshy's outputs, with enhanced depth, better lighting effects, and realistic shading. Similarly, Meshy excelled in mesh quality, particularly in rendering detailed features such as hair strands, clothing folds, and facial structures.

Realism was another major distinguishing factor—Meshy provided a more lifelike and expressive representation, capturing nuanced details. Additionally, Meshy demonstrated greater pose creativity, producing more dynamic and natural postures.

Figure 2 illustrates the outputs generated from the prompt "Anime girl with long black hair wearing a school uniform" across both platforms.



(a) (b) (c)

(d) (e) (f)

Fig. 3 Anime Girl Model Generation. Prompt: "Anime girl with long black hair wearing a school uniform" (a) Untextured anime girl model with long black hair generated by Meshy. (b) Highresolution textured anime girl model with long black hair generated by Meshy. (c) High-resolution textured anime girl model with two ponytails generated by Meshy. (d) Low-resolution textured anime girl model with long black hair generated by Lumalabs Genie. (e) High-resolution textured anime girl model with long black hair generated by Lumalabs Genie. (f) High-resolution textured anime girl model with two ponytails generated by Lumalabs Genie.

CONCLUSION

Recent advancements in AI-driven text-to-3D model generation have led to the development of platforms like Meshy.ai and LumaLabs.ai Genie, each offering distinct capabilities. Meshy excels in texture quality, mesh detail, and realism, making it particularly well-suited for complex organic models. Its high-resolution refinement further enhances model accuracy and visual appeal. On the other hand, LumaLabs Genie stands out for its speed and accessibility, providing a free, efficient solution for generating simpler, stylized models and low-resolution outputs, making it an attractive option for rapid prototyping.

Despite their strengths, both platforms have limitations. Meshy requires paid credits for extended use, while LumaLabs Genie, though free, has a longer processing times and offers less detailed textures compared to Meshy. However, both are evolving, with ongoing improvements in handling complex organic shapes and textures. Future advancements could introduce faster processing speeds, greater control over generation settings, and enhanced texture mapping, further pushing the boundaries of AI-driven 3D model creation.

References

B. Poole, A. Jain, J. T. Barron, and B. Mildenhall, "DreamFusion: Text-to-3D using 2D diffusion," 2022, arXiv:2209.14988.

C.-H. Lin, J. Gao, L. Tang, T. Takikawa, X. Zeng, X. Huang, K. Kreis, S. Fidler, M.-Y. Liu, and T.-Y. Lin, "Magic3D: High-resolution text-to-3D content creation," 2023, arXiv:2211.10440.

O. Michel, R. Bar-On, R. Liu, S. Benaim, and R. Hanocka, "Text2Mesh: Text-driven neural stylization for meshes," 2021, arXiv:2112.03221.

Y. Zhang, B. Cui, Q. Yang, Z. Li, and Y. Xu, "Benchmarking and learning multi-dimensional quality evaluator for text-to-3D generation," 2024, arXiv:2412.11170.

S. Su, X. Cai, L. Gao, P. Zeng, Q. Du, M. Li, H. Shen, and J. Song, "GT23D-Bench: A comprehensive general text-to-3D generation benchmark," 2024, arXiv:2412.09997.

Y. He, Y. Bai, M. Lin, W. Zhao, Y. Hu, J. Sheng, R. Yi, J. Li, and Y. Liu, "T^{3}Bench: Benchmarking current progress in text-to-3D generation," arXiv:2310.02977, 2024.

C. Li, C. Zhang, J. Cho, A. Waghwase, L.-H. Lee, F. Rameau, Y. Yang, S.-H. Bae, and C. S. Hong, "Generative AI meets 3D: A survey on text-to-3D in AIGC era," 2024, arXiv:2305.06131.

O. Michel, R. Bar-On, R. Liu, S. Benaim, and R. Hanocka, "Text2Mesh: Text-driven neural stylization for meshes," 2021, arXiv:2112.03221.

A. Chen, R. Yi and L. Ma, "Text-Guided 3D Object Generation via Disentangled Shape and Appearance Score Distillation Sampling," 2023 16th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Taizhou, China, 2023, pp. 1-6, doi: 10.1109/CISP-BMEI60920.2023.10373352.

Meshy AI, "Creating stunning 3D models with AI," 2024. [Online]. Available: https://www.meshy.ai/.[Accessed: 16-Feb-2025].

Lumalabs, "Dream Machine," 2024. [Online]. Available: https://lumalabs.ai/dreammachine.[Accessed: 01-Mar-2025].