IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# Converting Text To 3d Objects Using Generative Ai

[1]Potharaju Naga Venkata Sai Sudheer

[2]Nallagorla Jahnavi, [3]Bommanaboyina Sri Jahnavi, [4]Loya Chandana Naga Manikanta, [5]Somala Govardhann,

[1]Assistant professor,[2,3,4,5]Students,
[1,2,3,4,5]Department of Computer Science and Engineering/
Dhanekula Institute of Engineering and Technology, Ganguru, India

***Abstract:*** Text description-to-3D object generation We integrate natural language processing (NLP), long short-term memory (LSTM) networks and generative adversarial networks (GANs) to automate text description-to-accurate and aesthetically appealing three-dimensional (3D) models. NLP methods, including named entity recognition and semantic analysis, process and extract meaningful features from the text. LSTM networks analyze the sequential nature of the text and capture contextual relationships and dependencies. The processed features are used by GANs, where the generator constructs 3D models, and the discriminator ensures their accuracy and realism. This method overcomes the difficulties of manual 3D modelling by providing a scalable and efficient solution. Potential applications include virtual reality, gaming, architecture and cultural heritage preservation. The integration of NLP and Ai driven techniques shows a significant improvement in automating the generation of 3D models from text.

***Keywords*** - Natural Language processing, objects, Generative Adversarial network, model, generative AI,3D features.

## I. INTRODUCTION

The 3D model generation uses text by leveraging of Natural Language Processing (NLP), Long-short term memory (LSTM) networks, and Generative Adversarial Networks (GANs). This aims to generate an automated system to synthesize 3D models that are consistent with textual input and realistic in nature. Tokenization, named entity recognition, stop word removal and semantic analysis methods are being implemented by NLP techniques. It does by gaining and extracting significant informative features from textual descriptions using these methods. There are various contextual relationships among the sequential patterns and the text thoroughly examined for further assessment. The LSTM networks are being used for analysis as they provide remarkably good performance and the ability to remember the dependencies. These advanced methods work by putting the current input text through multiple processes which require the exact description to be well understood which provide basis for three dimensional representations. These features are then processed by GAN framework for execution. Within this framework, the generator generates three dimensional objects based on the input data. Discriminator takes a look at realism while being presented with (generated) products from the generator network. It is framed as a generative adversarial network (GAN) framework, where the system assesses the accuracy and realism of generated models.

## II. LITERATURE REVIEW

1. "**Generation High resolution 3D model from natural language by Generative Adversarial Network**", "**Kentaro Fukamizu, Masaaki Kondo, Ryuichi Sakamoto**", **(2019),** NLP techniques including tokenization, stop-word removal, named entity recognition, and semantic analysis techniques that operate together for text preprocessing inputs and extract meaningful features. This processed text is used by GAN to generate 3D models.
2. "**3D modelling by means of Artificial Intelligence**", "**Bebeshko B, Khorolska k, Kotenko N, Desiatk A, Sauanova K, Sagyndykova S, Tyshchenko D", (2021),** The outline features of 2D images recognition and the creation of 3D models, using AI and machine learning accordingly, it would make designers work much easier, graphic designers and architects would be able to only specific features instead of spending a lot of time on the actual drawing. The key purpose of this work is to define the possible limitations of CNN usage for 3D models generation, considering output resolution and generation swiftness.
3. "**Integrating Text-to-Image AI in 3D Design Workflows**", "**Vivian Liu, jo Vermeulen, George Fitzmaurice, Justin Matejka", (2022),** Text-to-image AI can generate novel images for inspiration, but their applications for 3D design workflows and how designers can build 3D models using AI-provided inspiration have not yet been explored. that designers saw great potential in 3DALL-E within their workflows and could use text-to image AI to produce reference images, prevent design fixation, and inspire design considerations.
4. "**Large Language and Text-to-3D Models for Engineering Design Optimization", "Thiago Rios, Stefan Menzel, Bernhard Sendhoff", (2023),** Generative AI for learning large neural network models with the capability to produce essays, images, music and even 3D assets from text prompts create opportunities for a manifold of disciplines. The optimizations indicate that, it is important to ensure that the designs generated from prompts are within the object class of application, i.e. diverse and novel designs need to be realistic, and, second, that more research is required to develop methods where the strength of text prompt variations and the resulting variations of the 3D designs share causal relations to some degree to improve the optimization.
5. "**Text-guided Voxel Editing of 3D Objects**" **"Etai Sella, Gal Fiebelman, Peter Hedman, Hadar Averbuch-Elor",** (**2023**), Demonstrates how voxel grids as an intermediate representation in GANs to generate 3D models. These grids Object forms undergo a process of discretization into volumetric data through this method. Such systems allow effective data storage and easy handling mechanisms. Voxel-based approaches are especially useful for tasks requiring fine-grained geometric detail in 3D objects.

6. "**Text-to-3D Shape Generation**", **"H. Lee, M. Savva & A.X. Chang", (2024)**, The computational demands of text-to-3D pipelines, cloud-based solutions have been explored for scalability. Studies show that distributed systems employing cloud GPUs and TPUs significantly reduce processing times, enabling real-time generation of 3D objects from text.

## III. METHODOLOGIES

1. **User Authentication and Management:** Users should have access to account registration functions which include secure login procedures as well as a log out mechanism. The system must include safe functions for registering account creation and securing the logout process. All user account data including email addresses, usernames and passwords needs proper storage. User authentication should be verified using secure login credentials. Passwords should be encrypted and stored securely in the system.
2. **Text Input and Processing:** Users need access to a graphical interface within the system to accomplish the task. The system requires text input systems that receive processed data. the system using Natural Language Processing (NLP) techniques. The system must split text into tokens and detect both entities and semantically analyze entered text. The system evaluates the input texts through semantic evaluation methods to extract meaningful features.
3. **Contextual Understanding:** The system needs to implement Long Short-Term Memory (LSTM) networks. This will help to analyse sequential relationships during the processing phase. contextual dependencies in the input text to ensure accurate 3D model generation.
4. **3D Model Generation:** The system needs to develop a 3D model through its processing. The application uses processed textual description as input to create the final 3D model. Generative Adversarial Networks (GANs) generated 3D model should be visually compelling, accurate, and realistic according to the description provided by the user.
5. **Model Visualization:** Users should access a 3D viewing interface through the system. Through this system users gain the ability to work with visualize generated 3D models. Users should have capabilities to perform rotation and panning along with zooming functions on the 3D model displayed. The system enables user inspection of details through different perspectives.
6. **Error Handling:** A graceful error handling function must be implemented into the system. The program must present simple feedback messages to users when errors occur the system needs to handle all errors which occur from incorrect user input as well as model failure and system malfunctions. The system will provide alerts to users regarding the incomplete text input and unsuccessful 3D model creation.
7. **Scalability and Performance:** The system needs to scale for processing a substantial amount of data requests. The system should process numerous requests for producing 3D models. and provide real-time results. The programmed system needs backend capacity optimization to fulfil its duty properly. processing heavy AI computations, especially during GAN model training and 3D generation.

## IV. PROPOSED METHOD

Converting textual descriptions into three dimensional (3D) models integrates advanced Natural Language Processing (NLP), Long Short-Term Memory (LSTM) networks, and Generative Adversarial Networks (GANs). The system receives Textual description serves as an input for this process.
For example, the user might input a sentence like “A yellow car” This input is in Web Application platform to collect such input from users. The textual input requires several preprocessing operations for its preparation process. critical for understanding the underlying features of the text. The information moves from the processing step to a subsequent stage before it reaches the model generation part. At this point, the pre-processed text Such information gets transformed into numerical data representation using advanced NLP techniques. The Word2Vec embeddings expressions into mathematical representations of vectors that capture their semantic meaning.
Additionally, implementation of LSTM networks enables understanding the sequential dependencies and context. The text features (such as shape, color and material, size) demonstrate sequential connections between them. The generator takes the encoded features to GAN to generate a 3D model.

### Natural Language Processing (NLP):

The system follows a series of methods which process and cleanse text before the analysis:

- Tokenization: The text undergoes segmentation into smaller distinct units called tokens, units, words or sub-words, for further analysis.
- Stop word Removal: Common but the processing typically omits normal functional words ("the", “and”, "a") together irrelevant words have been eliminated.
- Named Entity Recognition (NER): Operates as specific entities such as objects, materials, and other. The text provides details about significant components existing within it.
- Semantic Analysis: The relationships between various entities to provide better understand the object’s context, features, and properties.

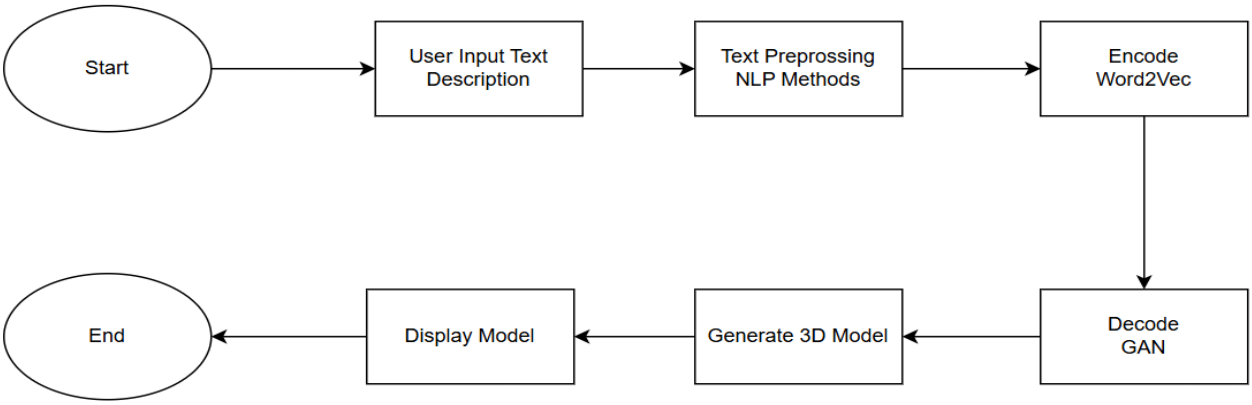
### Long Short-Term Memory (LSTM):

These networks help initiate sequential dependencies and context within the text. The LSTM application enables successful analysis of textual information in its original sequence. Text information distribution becomes possible through this tool. The model tracks context information throughout extended descriptions. The model recognizes both the single elements and their relationships across multiple text descriptions. Shape, color and material and size comprise the set of features which these models analyze.
For example, in the description “A burger” Through the LSTM, the model can interpret the sequence and learn the connection between "burger" through its training process.

### Generative Adversarial Network (GAN):

The system contains two main components:
Generator: The generator component operates on the encoded features which serve as its input and generates a 3D model. The model may be a mesh, voxel representation, or point cloud. The 3D model type depends on the requirements of the particular use case. The 3D object creation process of this generator includes the use of received input. The generated object matches all specifications established in the text input.
Discriminator: The 3D models need to be assessed for their realistic qualities generated by the generator. The discriminator evaluates generated 3D models against dataset items before delivering its assessment to the generator. The feedback from the discriminator enables the generator to enhance its performance throughout its operational period.

## V. MODEL DESCRIPTION




This model starts with user authentication to create an account to access the dashboard. After, the user can access to input their textual description of the 3D model they want to generate. The input text undergoes text preprocessing techniques of NLP methods to extract meaningful features from the text and done with semantic analysis to capture contextual relationships and dependencies. The further process is to encode Word2Vec to convert into mathematical representations of vectors that capture their semantic meaning. The processed features are used by GANs to decode the text and generate a 3D model. The process ends with downloading the model by user or displaying the model in dashboard.

## VI. IMPLEMENTATION

The Text-to-3D Model Conversion system must integrate Flask functionalities. It is a lightweight Python web framework, serves as the backbone of this project. To provide all necessary infrastructure to handle user requests. Converting textual descriptions into visually compelling 3D objects using Natural Language Processing (NLP), Long Short-Term Memory (LSTM) networks, to extract meaningful features and relationships. Further Generative Adversarial Networks (GANs) processed the text into visually compelling three-dimensional models.

## VII. RESULTS

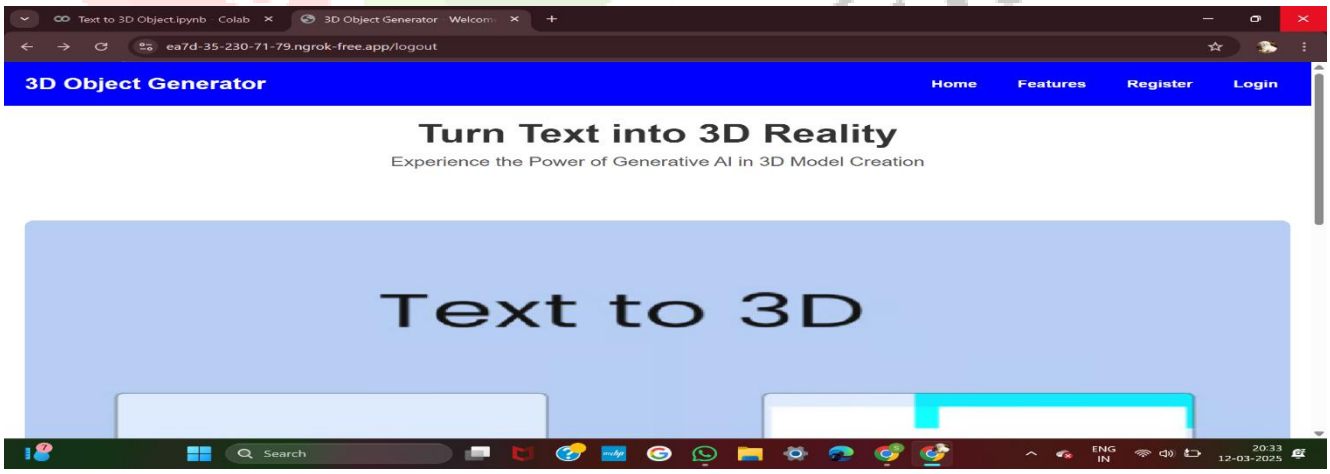



Fig 1: A web application 3D object Generator shows visual representation of process. It shows register, login, features which user can access. A box labelled Text to 3D shown with arrow pointing towards model of yellow car. This visually demonstrates how users can input text descriptions and receive 3D models as output. To access the dashboard user should create an account by register process.

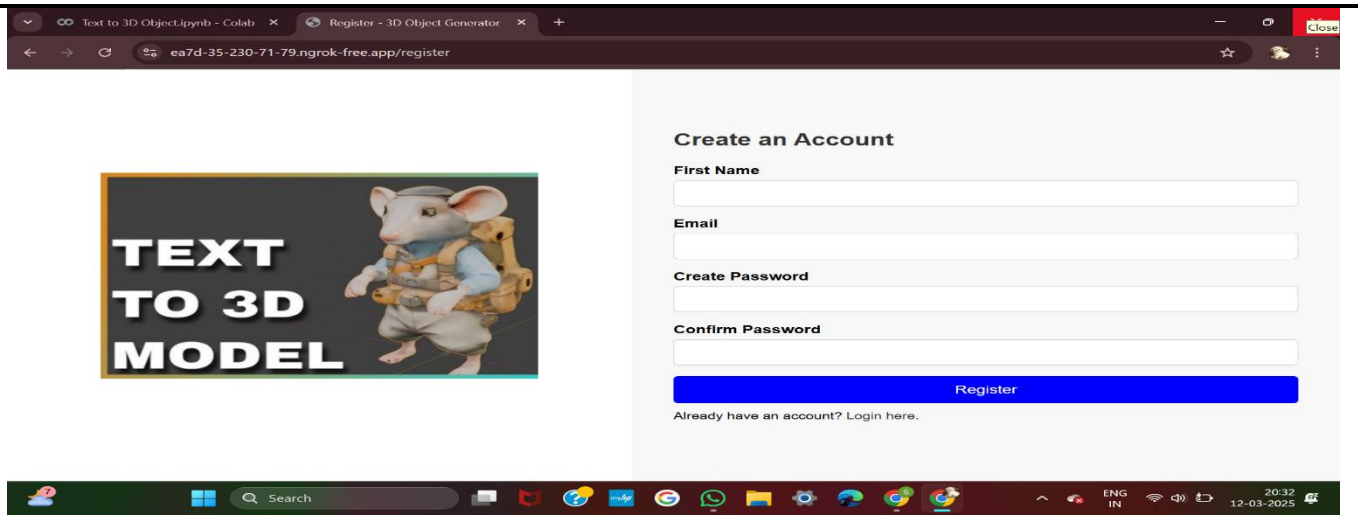



Fig 2: User should create an account with Name, Email, set up with password and get registered to access the Dashboard. There is a link below the button that reads "Already have an account? Login here." This provides an alternative option for users who already have an account with the website to directly log in instead of creating a new one.

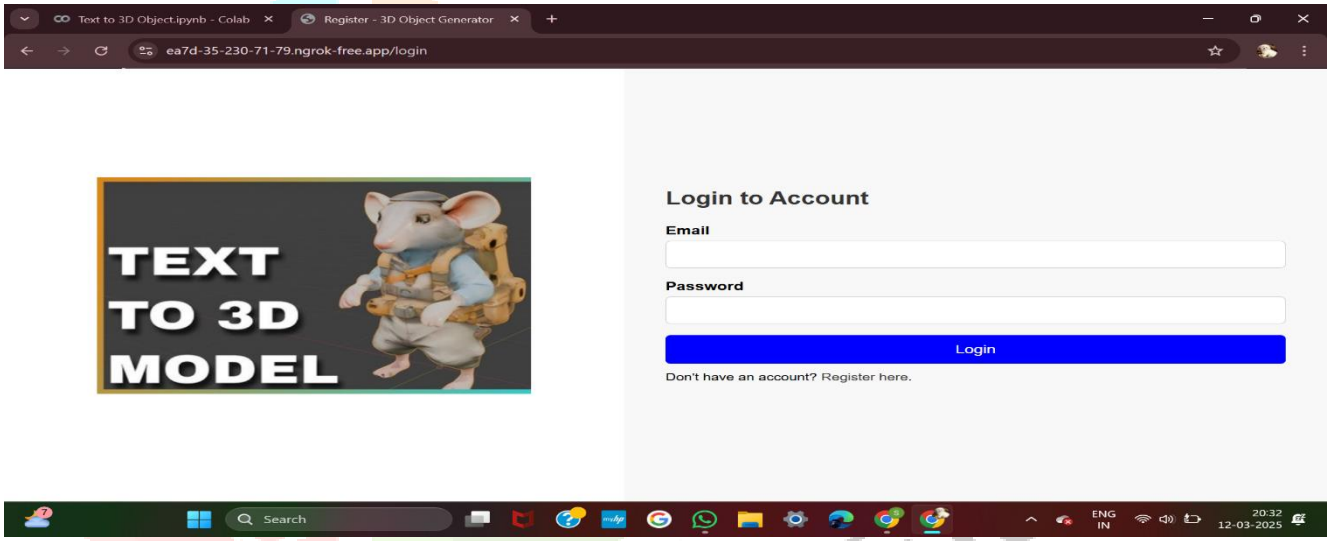



Fig 3: After successfully completion of register use have access to login with their created email and password to access their dashboard where they can give input descriptions to generate a 3D model. There is a link below the button that reads "Don't have an account? Register here." This provides if users who do not have an account yet to create one by navigating to the registration screen.

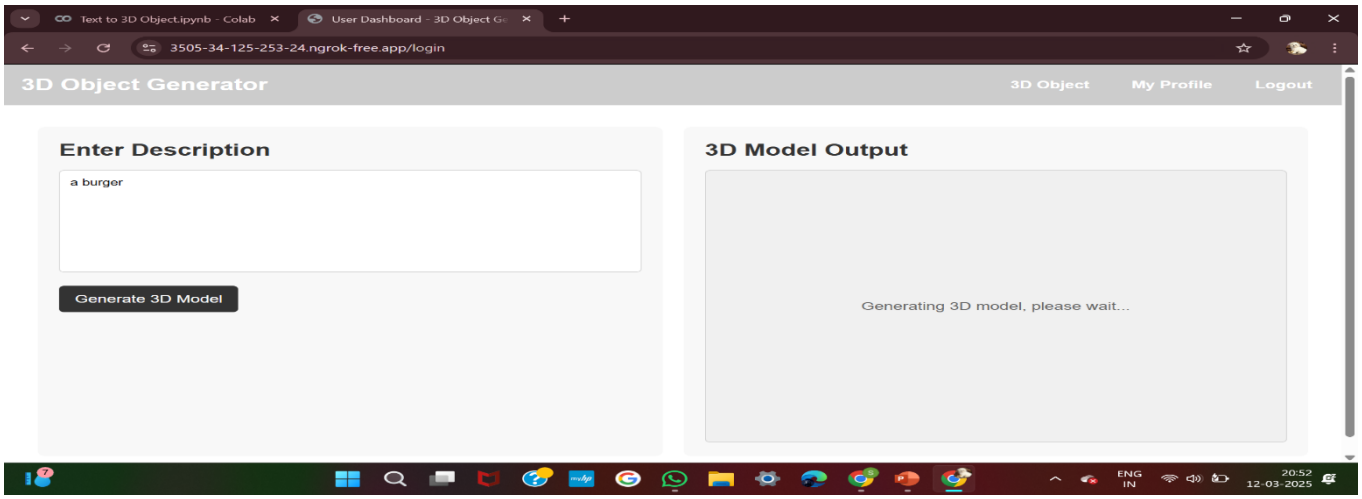



Fig 4: This shows a text input screen within a 3D model generation interface. A section titled "Enter Description" shows where a user can enter their textual description of the 3D model they want to generate.

A button labelled "Generate 3D Model". will initiate the 3D model generation process using the provided text description as input. Then the generated 3D model will be displayed at output field. input "a burger". By simply providing a textual description, the user has been able to generate a 3D representation of their desired object.

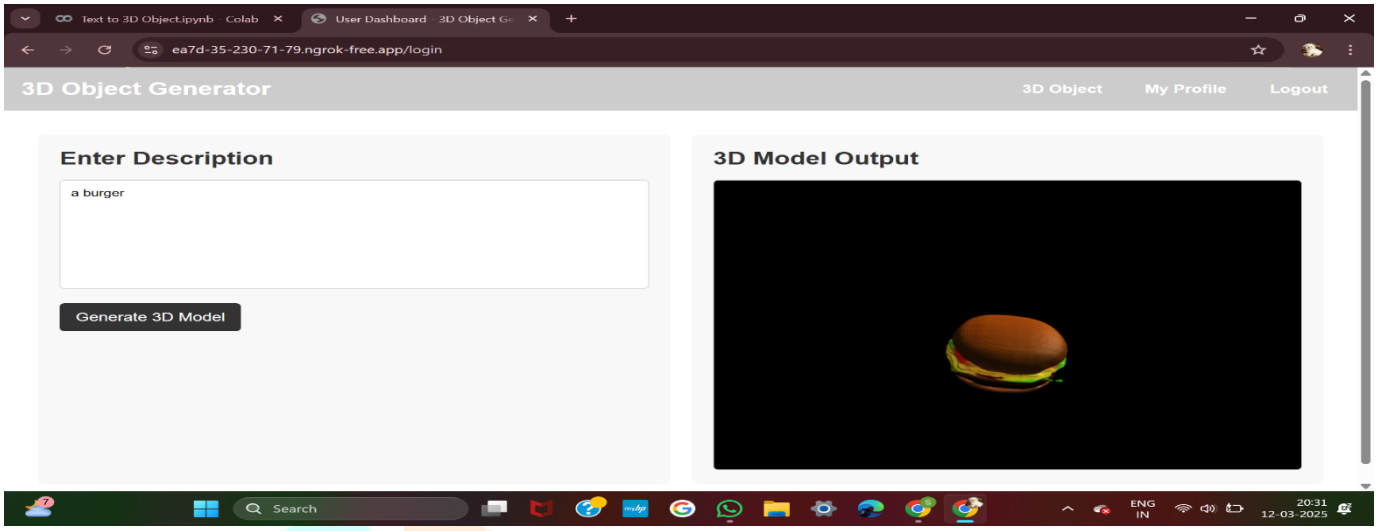



Fig 5: The "3D Model Output" section now displays the generated 3D model where the user previously input "a burger ". By simply providing a textual description, the user has been able to generate a 3D representation of their desired object.

## VIII. CONCLUSION

Text-to-3D Model Conversion System successfully integrates NLP Methods, LSTM and GAN. This method removes the difficulties that come from performing 3D modelling by hand. The system implements 3D modelling through a mechanism that provides scalable efficiency along with automated solutions for various industries, Gaming alongside virtual reality and architecture together with cultural heritage preservation composes the list of targeted industries using this solution.

## IX. ACKNOWLEDGMENT

I sincerely express my gratitude to all researchers, developers, and contributors for this study on **Converting text to 3D objects using GenAI**. I extend my appreciation to my mentors, professors, and peers for their invaluable guidance and feedback, as well as to the broader research community and open-source contributors for providing essential resources that facilitated this work. Lastly, I thank my friends for their unwavering support and encouragement throughout this journey.

## REFERENCES

**[1]** Fukamizu, K., Kondo, M., & Sakamoto, R. (2019). Generation high resolution 3d model from natural language by generative adversarial network. arXiv preprint arXiv:1901.07165.

**[2]** Bebeshko, B., Khorolska, K., Kotenko, N., Desiatko, A., Sauanova, K., Sagyndykova, S., & Tyshchenko, D. (2021). 3D modelling by means of artificial intelligence. Journal of Theoretical and Applied Information Technology, 99(6), 1296-1308.

**[3]** Liu, V., Vermeulen, J., Fitzmaurice, G., & Matejka, J. (2022). 3DALL-E: Integrating Text-to-Image AI in 3D design workflows. ArXiv. arXiv preprint arXiv:2210.11603.

**[4]** Rios, T., Menzel, S., & Sendhoff, B. (2023, December). Large language and text-to-3d models for engineering design optimization. In 2023 IEEE Symposium Series on Computational Intelligence (SSCI) (pp. 1704-1711). IEEE.

**[5]** Sella, E., Fiebelman, G., Hedman, P., & Averbuch-Elor, H. (2023). Vox-e: Text-guided voxel editing of 3d objects. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 430-440).

**[6]** Lee, H., Savva, M., & Chang, A. X. (2024, May). Text-to-3D Shape Generation. In Computer Graphics Forum (Vol. 43, No. 2, p. e15061).

**[7]** H. Singh and K. Kaur, "Utilization of Artificial Intelligence for 3D Production Process," 2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2023, pp. 33-37, doi: 10.1109/ICIRCA57980.2023.10220914.

**[8]** Amburgy, D., & Estell, J. K. (2024, March). Converting Text Into 3D Printable Braille. In 2024 ASEE North Central Section Conference.

**[9]** Lin, C. H., Gao, J., Tang, L., Takikawa, T., Zeng, X., Huang, X., ... & Lin, T. Y. (2023). Magic3d: High-resolution text-to-3d content creation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 300-309).

**[10]** Ma, Y., Liu, K., Guan, Z., Xu, X., Qian, X., & Bao, H. (2018). Background augmentation generative adversarial networks (BAGANs): Effective data generation based on GAN-augmented 3D synthesizing. Symmetry, 10(12), 734.