© JAN 2026 | IRE Journals | Volume 9 Issue 7 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I7-1713234

AI-Powered 2D to 3D transformation for Interactive Learning

MYTHRI[1], G PAVANI[2], AISHWARYA[3], BHAGYA K[4], GAGANA V[5]
[1, 2, 3, 4, 5] *Assistant Professor, Department of CSE (Data Science), RNS Institute of Technology, Bengaluru, India*

Abstract- ***Today, AI and NLP are changing how students learn and how teaching materials are created.However, traditional 2D learning materials often lack the interactive and immersive qualities necessary for mastering complex subjects ,such as science and engineering .The present work describes the system that applies different technologies in order for learning to be more interesting and easier to access. Our proposed solution works in uploading 2D images. These images are subjected to processing in turning them into simple 3D models through AI and computer-vision methods. In enhancing the engagement for learning, the integrated NLP-powered chatbot provides information concerning descriptive details of the 3D model in text and audio formats and a 3D outline texture for drawing a diagram. Our system should be of an easy nature to use for anyone and be able to support interactive learning in many users, which can significantly revolutionize classroom teaching by moving beyond passive content consumption to active, immersive engagement.***

Keywords- ***Artificial Intelligence AI, Natural Language Processing NLP, 2D to 3D conversion, Computer-vision, Deep-learning, Interactive Learning, Educational Technology, Chatbot, Text-to-Speech TTS.***

I. INTRODUCTION

Artificial Intelligence Transforming Education with AI and NLP: Immersive Learning System In the fast-moving pace of 21st-century education, old methods of teaching are not enough to captivate the futuristic minds of modern students by relying only on flat images and long notes.Students now use phones, apps and rich digital content daily so they expect similar interactive learning, find it challenging to engage deeply with abstract or complex subjects when presented in a flat,uninspiring manner. Topics such as human anatomy,science are difficult to grasp in their full three-dimensional reality from a two-dimensional image. This often leads to rote memorization rather than true understanding, this usually makes students memorize without properly understanding the topic, hindering critical thinking and the real-world application of knowledge.

This project proposes an innovative educational system designed to connect the students by leveraging the power of Artificial Intelligence (AI), Machine Learning(ML), and Natural Language Processing (NLP).Our objective is to achieve an interactive and engaging learning platform that can turn mere viewing into an exploration experience for the student.

For instance, if a person wants to learn biology it would be quite difficult to understand concepts like the functionality of the human heart ,structure and functions of a brain if they were represented on a two-dimensional diagram. Similar problems can be seen in learning mechanical engineering concepts including the functions of motors and machinery. It would be very hard for the visualization and kinetic learners.

Moreover, there is no scope for interaction and feedback in a 2D image. As soon as a student analyzes an image on a textbook page or on a slide, there is no chance to analyze it again either through rotation, animation or giving real-time information.This one-way method of understanding may lessen curiosity and independent learning. At the same time, young people have digital experiences such as 3D gaming applications. The gap between students digital use and their learning in school is growing rapidly.Traditional educational systems mainly work on 2-D images from textbooks and classroom materials to explain complex ideas, visual concepts and information. These flat representations have less quality or quantity for subjects such as biology, engineering and medicine where a true understanding requires three-dimensional visualization. It is difficult for students to grasp and understand 2D diagrams into 3D structures, which results in repeatedly revising the concepts or byheart

rather than conceptual clarity. Moreover, the technique is not interactive, leading to a learners becoming disinterested in the topic.

This struggle to visualize and interact with complex concepts such as deep learning reduces the student understanding in the concept and increases the gap between modern digital technologies and classroom teaching methods.Particularly, this problem is for students in resource limited institutions where physical 3D models or lab kits are not accessible. Basically,handling this challenge allows for better understanding of the concept, curiosity and making easy access to quality education.

II. LITERATURE SURVEY

The integration of AI with NLP has opened a whole new frontier in educational technology. Many research works have contributed essential theoretical and practical aspects that aim to build intelligent, available and immersive learning platforms. This literature review highlights the close relevance with the current project: conversion of 2D images into 3D interactive models, outlining textured models and using AI chatbots.

A. Traditional Methods of 2D-to-3D Conversion

Early research in 2D-to-3D image reconstruction made considerable use of traditional computer vision techniques and mathematical modeling. These approaches focused on depth estimation, mesh creation, and rendering using rule-based algorithms and handcrafted features. While such methods provided foundational insights, they often struggled with efficiency, accuracy and scalability when applied to complex educational diagrams. Mamand et al. (2024) reviewed both traditional and modern AI-based approaches, comparing algorithmic efficiency, accuracy and computational complexity offered a detailed guide for developers. Their work is particularly relevant for our project selecting similar deep learning methods and tools for converting educational diagrams into accurate and usable 3D models.

B. Machine Learning model - MiDaS

Birkl, Wofk, and Müller presented in 2023 MiDaS v3.1-A Model Zoo for Robust Monocular Relative Depth Estimation (Intel Labs, 2023). Their work puts much focus on generating relative depth maps from single images, enhancing geometric coherence and semantic preservation across diverse datasets. With the power of deep learning architectures trained on large-scale data, MiDaS enables strong generalization both indoors and outdoors. MiDaS is applied in this project to provide a robust basis in depth-aware 2D-to-3D transformations, enabling better reconstruction of the 3D model and allowing the WebAR chatbot to present more visually consistent educational content.

C. AI-Based Text-to Speech coversion

IJCRT 2025 introduces an text-to-speech that incorporates NLP, which enables the automation of daily functions like opening websites and sending messages. The study illustrates the flexibility of Python TTS frameworks for multi-functional systems. This supports the project's vision to extend the use of TTS beyond static narration, thus enabling interactive educational chatbots that respond dynamically to learner queries about 3D models.

D. Chatbot Using Flask Framework

Another major contribution is made by Anonymous (2025), Designing a Chatbot Application Using the Flask Framework and Rule Based Algorithm. But it highlights the importance of using Flask for handling chat bot routing and communications. Although it is based on a rule-based algorithm and not on machine learning, the structure and design at the backend clearly define functionality for handling user requests and giving respective outputs. The discussion on scaling capabilities with chat bots using Flask clearly fits into the project objective of developing a text-to-speech enabled chat bot. It clearly confirms and reinforces that it is a correct decision to use Flask as a backend platform enabling smooth communications with the user interface for narrating 3D models.

III. MOTIVATION AND OBJECTIVE

Our motivation stems from the idea that education should be inclusive, interactive and enjoyable for all

learners. Many scientific studies and educational research have shown that students learn better when they can interact with the material, receive instant feedback and explore content in multiple ways(text, audio, visual, and tactile).

The Key reasons that inspired us to conduct this project include the following:

1.Improved Understanding Through Visualization: Seeing a 3D model helps learners understand spatial relationships and functionality better than flat images do.

2.Practicing Sketching through 3D textures: The outline texture helps students to draw the part easily.

3.Voice-and Text-Based Learning: Adding a chatbot helps guide users. It answers questions, provides definitions, and explains functions.Even visually impaired students can benefit from speech output.The main aim of this project was to improve students learning using modern technology. The objectives are divided into the following core parts:

1.Technical Objectives: To develop an AI-based system using OpenCV and deep learning to convert two-dimensional (2D) images into three-dimensional (3D) models.

2.To use PyTorch3D, Blender is used to refine and manipulate 3D models from image data.

3.To integrate AR.js or 8thWall to create WebAR scenes that run smoothly on any browser-supported mobile device.

4.To build a lightweight, NLP-powered chatbot using Rasa or GPT-based APIs that can describe a 3D model in both text and voice.

5.To download the image along with summary for better reference .

IV. PROPOSED METHODOLOGY

Our project aims to learn between static 2D learning materials and immersive,3D interactive experiences. The proposed methodology outlines a multistage process leverages advanced AI and web technologies to achieve.

1. 2D to 3D Conversion

This foundational stage prioritizes on transforming a 2D image into a digital 3D representation..

•Depth Estimation: Following segmentation, AI-based deep learning models are used to infer depth information from the 2D image, generating a 3D representation.

•3D Model Generation and Refinement: We will use the estimated depth data to construct initial the 3D models. PyTorch3D is instrumental because of its capabilities in differential rendering, allowing us to iteratively refine the generated 3D models for improved accuracy and visual quality.

2. NLP-Based Chatbot and Audio Processing

An intelligent chatbot enhances the learning experience by providing interactive information and accessibility features.

•Chatbot Development: We will develop a chatbot powered by sophisticated NLP models, potentially utilizing frameworks such as GPT or Rasa. This chatbot can understand and respond to user queries related to the displayed 3D model.

•Text-to-speech (TTS) Conversion: To ensure better accessibility and engagement, we will integrate text-to-speech (TTS) functionality, allowing the chatbot to communicate information verbally.

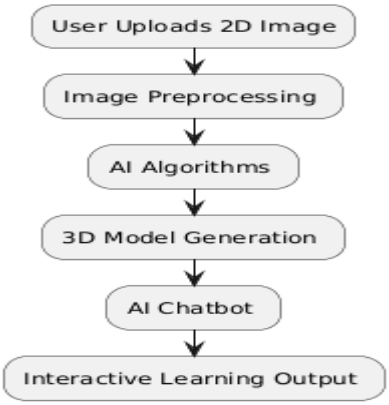
•Dynamic Information Retrieval: The chatbot will extract and synthesize relevant information from predefined datasets and research materials to provide accurate and dynamic responses tailored to the 3D model's subject.

5. Testing and Optimization

•Performance Testing: Rigorous performance testing will be conducted to ensure the real-time rendering of 3D models and smooth user interaction .

Chatbot Optimization: The chatbot's response time and accuracy will be continuously optimized through iterative testing and model refinement.

© JAN 2026 | IRE Journals | Volume 9 Issue 7 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I7-1713234



A. Block Diagram

1.User Uploads 2D Image:The process begins with a student or educator uploading a diagram,or illustration.

2. Image Preprocessing:Using OpenCV, the image was resized, segmented and analyzed for edges to prepare it for AI-based interpretation.

3. AI Algorithms: MiDaS or DPT estimates depth from a 2D image to infer a 3D structure.

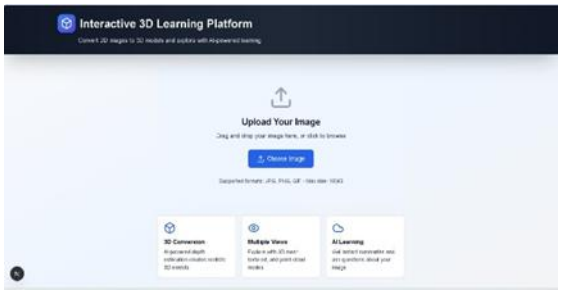
4. 3D Model Generation: PyTorch3D creates a mesh from depth data.Blender is used for texture mapping, mesh optimization, and exporting the model in formats such as glTF or OBJ.

5.AI Chatbot:A voice-enabled chatbot powered by NLP and TTS/ASR provides descriptive explanations of the 3D model in text and speech.It answers questions and supports interactive dialogues.

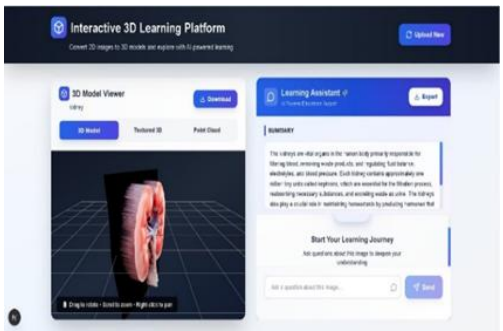
6. Interactive Learning Output:The final result is a 3D conversion that students can examine in real time.Includes features such as rotation and zoom for active learning and self-assessment.

IV. RESULT AND DISSCUSSION

The proposed Interactive 3D Learning Platform demonstrated effective integration of AI-based 3D reconstruction, natural language processing, and text-to-speech technology for enhanced educational engagement. The project worked with multiple 2D images of human organs and academic diagrams, and the result shows that it was capable of generating three different types of 3D images raw mesh images, reconstructed images with textures and images from point cloud data. The images were very accurate with high resolution and illumination. It can be seen that images with better resolution and brightness were more precise. It can be seen that images with better resolution and brightness were more precise. It aligns with previous research work because images with clear boundaries were more precisely detected using methods based on depth estimation. The NLP chatbot also added value to learning as it was capable of generating image-based explanations and answers various queries properly. It also enabled a step-by-step tutorial which would be more useful for school and university-going students. The tts functionality also made it easy for people with reading problems as they could learn via audios. The 3D viewer enabled functionalities such as rotation, zooming, and changes in mesh, point cloud, and texture mode which would be more useful for school and university-going students. The platform enabled generating reports on various PDF formats with images and 3D outputs. It would be more useful for academic purposes. From all these facts and findings, it can be obtained that it changes 2D images into 3D images and also enables intelligent conversations. Below images will demonstrate outputs.

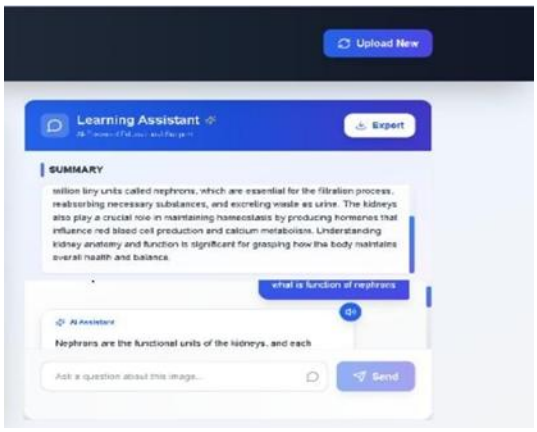


A.User Interface to upload a 2D Image



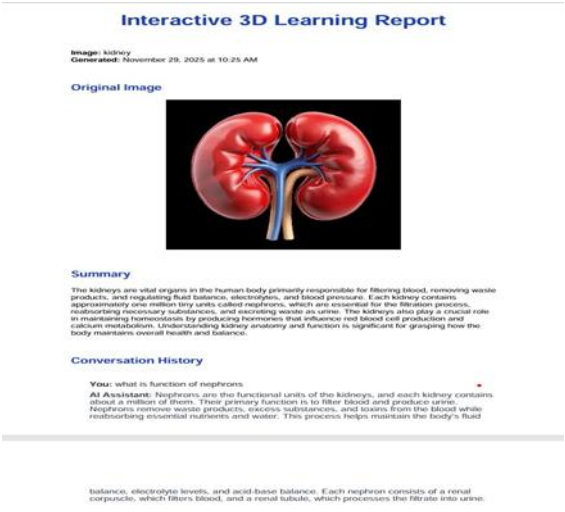
B. 3D model created with AI-supported summarization

© JAN 2026 | IRE Journals | Volume 9 Issue 7 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I7-1713234



C. Learning interface with text-to-speech

D. Textured 3D Image



E. Downloadable report with 2D image

V. CONCLUSION

This research presents an integrated AI–based learning system that transforms traditional 2D educational learning into interactive 3D experiences. By combining computer-vision techniques for 3D reconstruction, point-cloud generation and texture mapping with NLP-driven conversational assistance and text-to-speech capabilities. The proposed solution demonstrates a significant improvement in digital learning engagement. The system effectively bridges the gap between static textbook diagrams and dynamic visual understanding by enabling learners to visualize complex structures from multiple perspectives, interact with the model and obtain real-time explanations through an intelligent chatbot.

The inclusion of a 3D outline extraction module further enhances the pedagogical value by helping students practice diagram sketching with correct proportions. Experimental observations show that the system is scalable, accessible, and capable of delivering high-quality visual and textual outputs suitable for diverse educational environments. Overall, the proposed multi-technology framework highlights of technologies like AI-enhanced education, offering a foundation for future advancements such as higher-accuracy 3D reconstruction and learning modules.

VI. FUTURE SCOPE

This project has potential for future development and expansion. One important direction is accuracy of 3D reconstruction by adopting advanced deep learning models for depth estimation, mesh refinement and texture mapping by improving geometric consistency and visual fidelity. The platform can also be extended with domain-specific learning modules for students in subjects such as medicine, engineering or chemistry ensuring that learners receive contextualized and specialized support. Another promising thing is the integration of immersive technologies such as Augmented Reality (AR) and Virtual Reality (VR), which would allow students to interact with 3D models in real-world or fully virtual environments. Another is currently this project is designed to run locally and is not a cloud-based program.Future integration could incorporate cloud server deployment to enable scalability, remote accessibility, and collaborative learning across multiple devices. This would allow easy access to interactive 3D models by large groups of learners, making the system more versatile where it can be implemented in

institutions.Additionally,optimizing cross-platform usability for mobile, desktop and tablet devices, as well as integrating the system with Learning Management Systems (LMS), would further enhance its educational value by supporting structured lesson plans, assessments, and progress tracking. Finally, accessibility features can be expanded through improved text-to-speech, multilingual chatbot capabilities, and haptic feedback, ensuring inclusivity for visually impaired learners and diverse linguistic backgrounds.

REFERENCES

[1] S. S. Mamand and A. I. Abdulla, "2D to 3D Image Conversion Algorithms" ITM Web of Conferences, vol. 64, pp. 01010-01010, 2024.

[2] Y. Guo, H. Wang, Q. Hu, H. Liu, L. Liu, and M. Bennamoun, "Deep Learning for 3D Point Clouds: A Survey" IEEE Trans. Pattern Anal. Mach. Intell., vol. 43, no. 12, pp. 4338–4364, Dec. 2021.

[3] R. Birkl, D. Wofk, and M. Müller, "MiDaS v3.1 – A Model Zoo for Robust Monocular Relative Depth Estimation" Intel Labs Tech. Rep., 2023.

[4] A. Maidamshetti and R. Kanth, "A Web-Based Chatbot with Advanced Voice Control and Document Interaction Using Flask and JavaScript" International Journal for Research in Applied Science and Engineering Technology (IJRASET), vol. 13, no. VIII, Aug. 2025. doi: 10.22214/ijraset.2025.73852.

[5] P. H. S. Prasantha, "Image Text to Speech Conversion in Desired Language" International Journal of Advanced Research in Computer Science, vol. 14, no. 3, pp. 45–52, 2023.

[6] Grenze International Journal of Engineering and Technology ,"Voice Mate using Pyttsx3 and Speech Recognition" vol. 11, no. 2, pp. 109–115, 2025.

[7] S. R. Kumar and A. Patel, "AI Based Voice Assistant" International Journal of Creative Research Thoughts (IJCRT), vol. 13, no. 4, pp. 5627–5635, 2025.

[8]A. Musa, S. Danbature, M. J. Isa, and A. Ali, "Designing a Chatbot Application Using the Flask Framework and Rule-Based Algorithm" Jurnal Teknoinfo, vol. 7, no. 1, pp. 11–16, 2023, doi: 10.47233/jteksis.v7i1.1820

[9] R. Leroy, P. Trouvé-Peloux, F. Champagnat, B. Le Saux, and M. Carvalho, "Pix2Point: Learning Outdoor 3D Using Sparse Point Clouds and Optimal Transport," arXiv, Jul. 30, 2021.