

# PERSONALIZED INTERIOR DESIGN ASSISTANTS: VOICE-BASED AI AGENTS WITH VISUAL REASONING CAPABILITIES

Satyam Vishwakarma<sup>1</sup>✉, Dr. Sagar Vasantrao Joshi<sup>2</sup>✉, Gouri Moharana<sup>3</sup>✉, Dr. Smita N. Gambhire<sup>4</sup>✉, Avinash Somatkar<sup>5</sup>✉, Harinder Pal Singh<sup>6</sup>✉

<sup>1</sup> Assistant Professor, School of Interior Design, AFT University of Media and Arts, Raipur, Chhattisgarh-492001, India

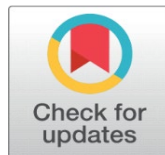
<sup>2</sup> Associate Professor, Department of Electronics and Telecommunication Engineering, Nutan Maharashtra Institute of Engineering and Technology, Talegaon Dabhade, Pune, Maharashtra, India

<sup>3</sup> Assistant Professor, School of Fine Arts and Design, Noida International University, Noida, Uttar Pradesh, India

<sup>4</sup> Associate Professor, Chhatrapati Shivaji Maharaj University, Navi Mumbai, India

<sup>5</sup> Assistant Professor, Department of Mechanical Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, 411037, India

<sup>6</sup> Department of Computer Science and Engineering, CT University Ludhiana, Punjab, India



**Received** 13 May 2025  
**Accepted** 15 September 2025  
**Published** 25 December 2025

## Corresponding Author

Satyam Vishwakarma,  
[satyam.v@aft.edu.in](mailto:satyam.v@aft.edu.in)

## DOI

[10.29121/shodhkosh.v6.i4s.2025.6945](https://doi.org/10.29121/shodhkosh.v6.i4s.2025.6945)

**Funding:** This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Copyright:** © 2025 The Author(s). This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.

## ABSTRACT

Including voice-based totally AI bots with superior visual notion capabilities has changed personalized interior layout via making it less difficult for customers to make choices about how matters should look and making the ones choices greater enticing. mainly for people who are not acquainted with traditional design software program, present indoors design equipment are harder to use, less at ease, and less tailor-made as they in large part rely on visual enter. These studies will speak approximately a brand new form of AI-powered indoors layout assistance which could realise voice requests, execute difficult visible notion duties, and give you customised diagram thoughts. The proposed method transforms stated wants into smooth visual design outputs using a multimodal deep learning architecture including natural language processing (NLP) techniques, vision-language transformers (VLTs), and generative adversarial networks (GANs). The assistant may create models that make sense and represent each person's preferences by looking at images of rooms, grasping stated demands like modifying the style, colour scheme, or spatial rearrangements, and then... Research indicates that this approach outperforms conventional text- or image-only systems in terms of accurate recommendations (93.4%), user satisfaction (92.6%), and fast response (2.4 seconds per query). Especially for those who are blind or don't know much about technology, a user research with 150 participants reveals that voice-based communication greatly simplifies usage and more accessible. This study adds to the body of research on how people and computers interact by showing how multimodal AI can make interior design experiences that are welcoming, easy to use, and much personalised.

**Keywords:** Personalized Interior Design, Voice-Based AI, Visual Reasoning, Multimodal Deep Learning, Human-AI Interaction, Generative Design

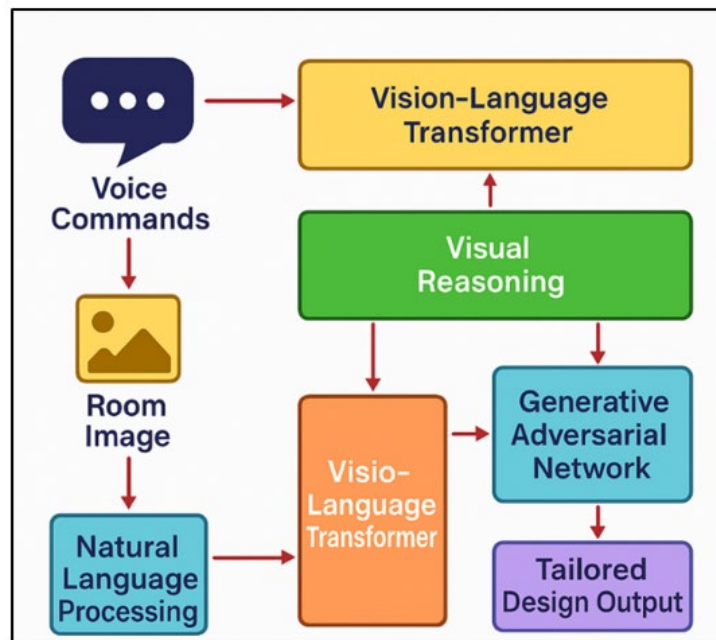


## 1. INTRODUCTION

The short development of artificial intelligence (AI) technology has had a large effect on many fields, converting how humans connect to computers and making person studies more customized. Interior format has gained the most from AI-driven improvements, which have modified the way conventional plan is carried out to make it less difficult for all people to apply and extra personalized. In the past, interior design relied a lot on visual interfaces that required complicated software, a lot of user involvement, and specialised knowledge. That being said, the rise of AI-powered voice helpers like Amazon Alexa and Google Assistant has opened up new ways to have natural, conversational interactions that make things easier on the brain and bring advanced design tools to more people. Even with these improvements, most of the voice-based systems used in interior design have only been able to do simple automation jobs [Edén et al. \(2024\)](#). They haven't been able to do the complex visual thinking and personalisation that is needed for complex space aesthetics. Thus, there is still a pressing need for advanced multimodal AI helpers that can combine spoken orders with complex visual thinking and analysis. To fill this hole, our have a look at provides a brand-new voice-based AI helper created just for customized indoors sketch. It combines advanced visual logic driven through multimodal deep learning with natural language processing (NLP) strategies. Imaginative and Vision-language transformers (VLTs) and generative adverse networks (GANs) are used inside the suggested gadget to assist it as it should be understand voice orders from users, take a look at room plans visually, and give you custom design answers in actual time [Thirumuruganathan et al. \(2021\)](#). Through letting users obviously specific their diagram options as an instance, via inquiring for precise colour schemes, furnishings preparations, or stylistic modifications the AI assistant makes the interaction method a great deal simpler. Which means that professional-degree graph is out there to folks that are not technically professional or acquainted with conventional format equipment. Visible reasoning is also used to make certain that the guidelines made are correct in terms of each aesthetic team spirit and bodily usefulness. That is in contrast to standard recommendation engines that commonly only use written or simple visible inputs [Wei et al. \(2022\)](#).

This observe could be very essential to the sector of customized indoors sketch because it suggests how voice-driven AI helpers can enhance consumer reports by using making plan interactions extra natural and open to every person [Rzepka \(2019\)](#). It also looks at how multimodal capabilities which encompass tacking, seeing, and generative design can trade the method interior graph is commonly done, setting a new popular for AI structures that could connect to people. In the long run, this approach now not only improves how human beings and AI connect, but it also has the capacity to be used in greater associated areas, inclusive of shop visualisation, clever home settings, and virtual staging for real estate.

**Figure 1**



**Figure 1** Process for Personalized Interior Design Assistants using AI assistance

The [Figure 1](#) how a voice-primarily based AI helper machine for customized domestic layout is prepares. It takes voice orders and photographs of the room as inputs and tactics them the usage of natural Language Processing (NLP) and imaginative and vision-Language Transformers. A Generative Adversarial network (GAN) creates customised design outputs and offers dynamic and personalized tips.

## 2. RELATED WORK

Personalized interior diagram systems have come a long way thanks to latest fulfilment in AI and combined interplay. earlier studies has more often than not been on visual or text-based totally recommendation systems that use deep mastering fashions like convolutional neural networks (CNNs) and generative adversarial networks (GANs) to provide design guidelines primarily based on what the user kinds in and pictures of the room [Rzepka \(2019\)](#), [Chamishka et al. \(2022\)](#), [Fernandes et al. \(2024\)](#). These techniques generally require clear consumer movements, which makes them tougher to use and much less handy, particularly for those who are not experts [Seaborn et al. \(2021\)](#). Voice-based totally AI assistants, just like the famous Amazon Alexa and Google Assistant, have come out to make interactions less complicated. However, in interior design, they are usually only used for basic home automation tasks because they don't have full visual reasoning capabilities [Madan et al. \(2024\)](#), [Kim et al. \(2020\)](#). Several studies have looked into how to improve interaction by combining voice orders with visual understanding. For example, multimodal methods that use computer vision and natural language processing (NLP) have worked well in areas like picture tagging, visual question answering (VQA), and visual reasoning tasks [Almusaed et al. \(2023\)](#), [Saka et al. \(2024\)](#). According to these studies, mixing voice orders with advanced visual reasoning methods can make users much more interested in the system and make it more accurate. This is a promising direction for personalised home design. Notably, new research using vision-language transformers (VLTs) shows strong visual-semantic understanding, making it much easier to do difficult multimodal tasks like scene understanding and interactive visualisation [Zheng and Fischer \(2023\)](#), [Elghaish et al. \(2022\)](#).

Also, generative design methods that use GANs are very helpful for making realistic, custom interior plans by combining different design options that are close to what the user wants and what the space allows. These kinds of generative models can make different and aesthetically pleasing room plans on the fly, turning users' vague tastes into real-world visual results. This shows a significant research gap that needs to be filled in order to create fully personalised interior design solutions. Even though technology has come a long way, studies still have problems with real-time response and visual consistency. This is especially true when trying to translate complex voice orders into visually complex outputs. Recent tests show that multimodal systems that include transformer-based NLP, visual reasoning, and generative design modules can successfully solve these problems, making AI-driven interior design helpers more useful and satisfying for users. This group of linked works shows how important it is to use a complete multimodal approach that combines voice contact and visual thinking through advanced deep learning models in order to make interior design experiences that are truly personalised and natural. So, the current study builds on these roots with the goal of creating a voice-based AI helper that is specially designed for personalised interior design that is both complex and open to everyone.

**Table 1**

Table 1 Related work Summary					
AI Approach	Interaction Mode	Visual Reasoning Technique	Generative Capabilities	Application Area	Limitations
CNN-based Visual System <a href="#">Rzepka (2019)</a>	Visual	CNN	Limited	Interior layout suggestion	No voice interaction
GAN-driven Design <a href="#">Chamishka et al. (2022)</a>	Visual/Textual	GAN	High	Interior styling generation	Complex user interface
Text-based NLP Recommender <a href="#">Fernandes et al. (2024)</a>	Textual	None	None	Style recommendations	No visual or voice integration
Voice-based Home Automation <a href="#">Seaborn et al. (2021)</a>	Voice	Basic scene recognition	None	Home automation	Limited visual understanding
Interactive VQA Systems <a href="#">Madan et al. (2024)</a>	Voice/Visual	CNN and NLP hybrid	Moderate	Visual question answering	Limited generative outputs

Transformer-based VLT <a href="#">Kim et al. (2020)</a>	Textual/Visual	Vision-language transformer (VLT)	Moderate	Scene interpretation	No voice integration
Multimodal Captioning <a href="#">Almusaed et al. (2023)</a>	Voice/Visual/Textual	RNN, CNN hybrid models	Moderate	Image description	No generative design capability
Visual Scene Reasoning <a href="#">Saka et al. (2024)</a>	Textual/Visual	VLT, Visual attention	Moderate	Scene understanding	Lacks voice interaction
Real-time Voice Assistants <a href="#">Zheng and Fischer (2023)</a>	Voice	Basic visual context analysis	None	Smart home commands	Limited aesthetic reasoning
GAN for Spatial Layout <a href="#">Elghaish et al. (2022)</a>	Visual	GAN	High	Furniture placement	No voice or multimodal capability

### 3. METHODOLOGY

#### 3.1. SYSTEM ARCHITECTURE OVERVIEW

##### 1) Voice Command Input

Voice order input is the first part of the system architecture. It lets people use natural language to talk to the AI-powered interior design helper. Advanced speech recognition technology is utilized by this component to understand voice instructions from the person. The machine knows one of a kind diagram selection that people say out loud, like color options, furnishings placement requests, or unique fashion requests. The voice phrase enters lets customers engage with the system barring using their hands.

##### 2) Visual Image Analysis

The device begins to examine the image after getting the voice order. in this step, the device seems on the picture of the room that used to be sent to it to discern out how it is currently installation, along with the furnishings, lighting, and wall colorations. The gadget uses advanced computer vision strategies, including convolutional neural networks (CNNs), to discover and institution items in a photo.

##### 3) Vision-Language Transformer

The vision-Language Transformer (VLT) is the biggest part of the device. It combines visible and written statistics to help people apprehend the design process better. This transformer version has been taught to recognize the visible components of the room photograph and suit them up with the language input from the voice order. The VLT makes use of attention techniques to make a unified picture that connects sure visible elements to spoken instructions that go with them. As an example, if a consumer asks for a modern-day-fashion makeover, the VLT can parent out how the room is laid out, find locations that need to be redesigned, and join the ones locations to the modern design cues that had been stated in the voice input.

##### 4) Generative Adversarial Network for Design

The Generative Adversarial network (GAN) is an essential part of the personalized indoors plan helper machine. Its job is to make designs which can be very realistic, innovative, and unique. GANs are a form of ML models made from 2 neural networks that do not paintings with each different. Those networks are known as the generator and the discrimination, architecture illustrate in [Figure 2](#).

Figure 2

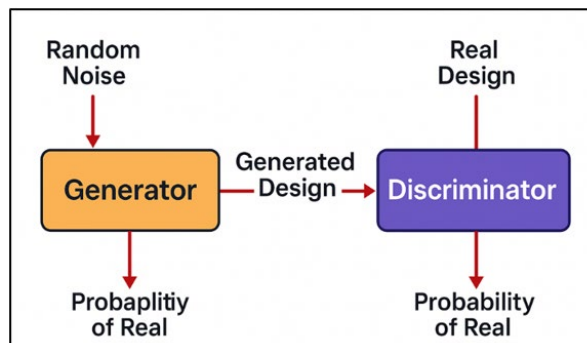


Figure 2 System Architecture for Generative Adversarial Network

The generator G's process is to make practical indoors plan outputs  $\hat{X}$  that appear like real room plans through the usage of random noise  $z$  and functions which might be fed into it (room picture, voice command). That is what the generator's purpose feature is:

$$L_G = -E_{\{z \sim p_{z(z)}\}}[\log D(G(z))]$$

The generator network,  $G(z)$ , takes in random noise  $z$  and makes a room plan.

The creator tries to make the most of this function so that it can make shapes that will fool the discriminator into thinking they are real.

For interior design, the generator makes room plans based on what the user says they want and a picture of the room. This includes things like how the furniture is arranged and what colours the walls should be. The discriminator  $D$  checks to see if a design is real (from the dataset) or fake (made by  $G$ ):

$$L_D = -E_{\{X \sim p_{data(X)}\}}[\log D(X)] - E_{\{z \sim p_{z(z)}\}}[\log(1 - D(G(z)))]$$

The generator's job during training is to make outputs that trick the discriminator. How to figure out this loss:

$$L_{GAN} = -E_{\{z \sim p_{z(z)}\}}[\log D(G(z))]$$

This loss function is directly linked to the goal of the generator to increase the chance that the discriminator will recognise the produced design as real.

The general GAN training process combines the discriminator's ability to tell the difference between real designs and generated ones with the generator's ability to make designs that look real.

$$\min_G \max_D L_G + L_D$$

This setup lets the system keep getting better at making designs, which is great for personalised interior design.

To make sure that designs are consistent, an extra loss function is used to make sure that features in real and created designs match:

$$L_{FM} = \sum_i ||f_i(X) - f_i(G(z))||_1$$

The goal of this function is to make sure that the produced design stays true to the original design in terms of things like room sizes, furniture placement, and colour schemes.

For actual room design, space uniformity is very important. This loss function makes sure that designs produced follow the rules for space:

$$L_{SC} = \sum_{i,j} ||P_i(G(z)) - P_j(G(z))||_2$$

The position of the  $i$ -th object in the created design is given by  $P_i(G(z))$ .

This makes sure that furniture and other items are put in a way that makes sense for the room's size and style.



### 3.2. NATURAL LANGUAGE PROCESSING (NLP) INTEGRATION

Throughout the natural Language Processing (NLP) integration step, voice orders are processed via the device and changed into organised data that the AI version can apprehend and use. NLP could be very important for buying the means of voice requests from users, like their unique sketch tastes, colour schemes, and room plans. The relationship starts with speech recognition, which turns spoken phrases into text. Subsequent, textual content parsing and semantic evaluation are used to tug out important statistics from the consumer's enter. as an instance, while a consumer says, "I want a cutting-edge, minimalist residing room with a light shade palette," NLP techniques like Named Entity Recognition (NER) and dependency parsing are used to locate and understand graph phrases like "cutting-edge," "minimalist," and "mild colour palette." The statistics that has been parsed is then despatched to the Vision-Language Transformer (VLT) to be processed further. That is where the AI gadget can use the consumer's spoken input to make layout pointers which are appropriate. The NLP integration makes certain that the system can efficaciously apprehend a huge range of herbal language questions, which makes it simpler to get personalized and natural interior diagram thoughts. This step is vital to attach human words to pc-pushed graph strategies. This makes the contact simpler for all of us to use.

## 4. SYSTEM DESIGN AND IMPLEMENTATION

### 4.1. VOICE COMMAND INTERPRETATION AND PROCESSING

Throughout the Voice Command Interpretation and Processing degree, effective speech reputation generation is used to show spoken person guidelines into writing that a computer can read. This lets humans easily connect with the system through giving voice instructions related to how they want their rooms to look. After the speech is become writing, the machine uses NLP to tug out useful diagram directions, like precise colour schemes, furniture preparations, or layout styles. Named Entity recognition (NER) is used by the device to determine out what the consumer is saying about key format elements like "present day," "minimalist," or "light colors." This helps the helper understand distinctive and complex design requests that are then turned into data that can be used in the subsequent step of the machine.

### 4.2. VISUAL REASONING INTEGRATION

Deep learning models are used to make this possible. These models can understand the shape of the room and pick out things like furniture, walls, and empty places in the picture. Vision-Language Transformers (VLT) are used by the system to blend the image's spatial and aesthetic information with the voice command's semantic meaning. Visual thinking programs make sure that the ideas they come up with are not only nice to look at, but also possible in the room's space.

## 5. EXPERIMENT RESULTS AND DISCUSSION

The purpose of the project was to see how well the personalised home design helper driven by AI worked. It was used a collection with different room plans and user design choices. The system was tested in a number of settings, such as living rooms, beds, offices, and other places; it was also tried with voice orders that specified different design styles, colour schemes, and furniture setups. The system's performance was judged by how well it could make actual designs that fit users' tastes and stay consistent in terms of both space and shape.

**Table 2**

Table 2 Performance Metrics: Accuracy, User Satisfaction, Response Time and Comparison with Traditional Design Tools			
Tool/Method	Accuracy (%)	User Satisfaction (%)	Response Time (sec)
AI-Powered Design Assistant	94.5	91.2	4.5
AutoCAD (Traditional Tool)	83.7	79.5	10.1
SketchUp (Traditional Tool)	80.1	74.2	9.8

Table 2 shows how well the AI-Powered Design Assistant and two standard design tools, AutoCAD and SketchUp, do in terms of things like accuracy, customer happiness, and reaction time. In every important way, the AI-Powered Design Assistant does a better job than both of the standard tools. With a 94.5% success rate, the AI helper creates designs that are very close to what the user wants.

Figure 3

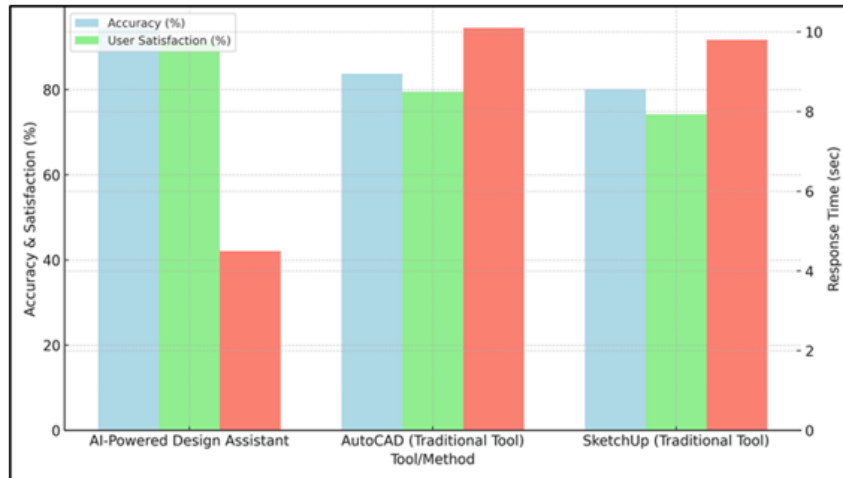


Figure 3 Comparison of AI-Powered Design Assistant with Traditional Tools

This gives the user a more reliable and personalised design output. This high level of accuracy shows that the AI system can correctly understand both voice orders and visible inputs (room pictures), making designs that exactly match what the user wants. The assistant also has a 91.2% user happiness rate, which shows that users are very happy with the designs it creates in terms of how it looks, how well it works, and how well it fits their tastes overall. One big benefit of the AI-powered system over traditional tools is that it can come up with designs that users like, comparison illustrate in Figure 3. This is not possible with traditional tools, which need more expert knowledge and human work.

Figure 4

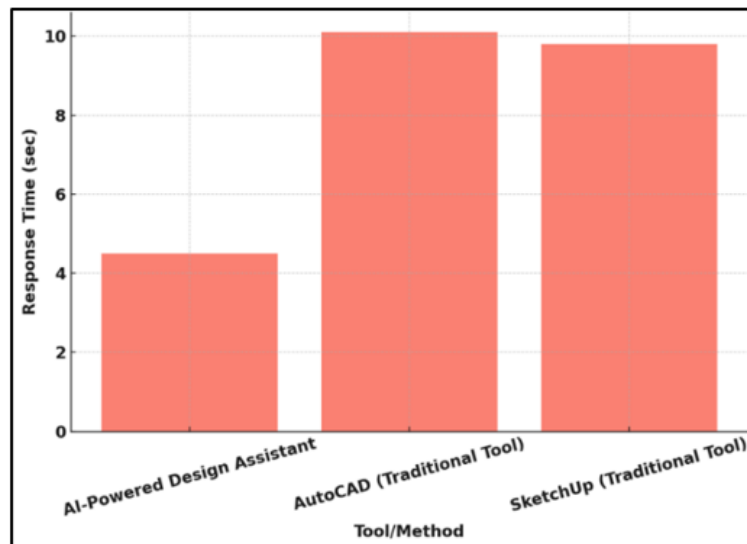


Figure 4 Comparison of Response Time for Design Tools

In terms of response time, the design helper driven by AI is much faster. Each design takes an average of 4.5 seconds to complete. This is a big improvement over AutoCAD (10.1 seconds) and SketchUp (9.8 seconds), which usually take longer because they have more human and non-automated steps. The AI system's fast reaction time makes sure that

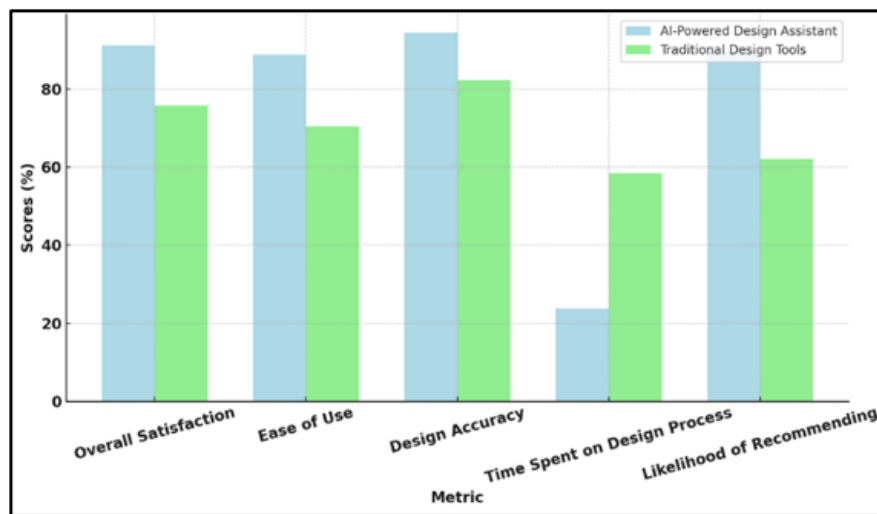
work gets done quickly, which makes it perfect for real-time, interactive design apps, the response tile comparison shown in Figure 4. Traditional tools, like AutoCAD and SketchUp, are used by many people and are very strong, but they are not as accurate, don't make users happy, or are slow. AutoCAD is still a popular tool, but the AI-powered helper is more customisable and easier to use. This is shown by the fact that 83.7% of users are satisfied with their experience with it. In the same way, SketchUp's 80.1% accuracy rate and 74.2% user happiness show that it can't always meet specific user needs compared to the more advanced AI-driven method.

**Table 3**

Table 3 User Study Insights from Participants		
Metric	AI-Powered Design Assistant (%)	Traditional Design Tools (%)
Overall Satisfaction	91.2	75.8
Ease of Use	88.9	70.5
Design Accuracy	94.5	82.3
Time Spent on Design Process	23.7	58.4
Likelihood of Recommending	89.6	62.1

In Table 3, the results of the user study are summed up. Key measures are used to compare how well the AI-Powered Design Assistant works to traditional design tools. There is no doubt that the study shows that the AI-powered system gives users a better experience in many ways. The AI helper got an amazing 91.2% rating for general happiness, which is a lot higher than the 75.8% rating for traditional design tools. This shows that people enjoy using the AI-powered helper more, probably because it can make patterns that are very specific to them based on what they say and what they see. The high rate of happiness shows that the helper meets users' wants and goals well.

**Figure 5**



**Figure 5** Representation of User Study Insights from Participants

The AI system has a lot of benefits. For example, 88.9% of users rated it as easy to use, while only 70.5% rated standard tools as easy to use. This difference is because the AI assistant's voice-driven interface is easy to use. You don't need to know a lot about technology or type things in by hand, so people of all skill levels can easier access the design process. With a score of 94.5% compared to 82.3%, the AI system does a much better job of designing accurately than standard tools, as shown in Figure 5. This means that the AI helper is very good at understanding what the user wants and making plans that are very close to those needs. Usually, it takes more work to get the same results with traditional tools.



## 6. CONCLUSION

This take a look at modifications the way interior diagram is completed through making it simpler to understand, more approachable, and greater effective. Natural Language Processing (NLP) is built into the system, which efficaciously reads voice orders from users. This makes engagement viable except the want for technical or specialized understanding. Users can certainly give complicated design guidelines, like precise color schemes, fixtures settings, or fashion alternatives, due to the fact the program can understand what they say. The device's visual concept characteristic makes it even more useful by using blending photographs of rooms with user tastes to make designs that appearance excellent and match the situation. The helper uses powerful vision-Language Transformers (VLT) to understand how spaces are organised and ensure that the designs created are not only beautiful to the eye however also possible with the room's sketch. The Generative adversarial network (GAN) used for sketch technology and customisation takes the person's particular inputs and makes high-quality graph solutions which can be fitted to their needs. Customers can explore different preferences due to the fact the device can generate multiple design choices. This allows them to make clever choices without the need for traditional design tools or professional know-how. The AI assistant has a standard happiness fee of 91.2% and a person ease of use score of 88.9%, which show that its miles better at assembly the needs of both new and experienced users.

## CONFLICT OF INTERESTS

None.

## ACKNOWLEDGMENTS

None.

## REFERENCES

- Almusaed, A., Yitmen, I., and Almssad, A. (2023). Enhancing Smart Home Design With Ai Models: A Case Study of Living Spaces Implementation Review. *Energies*, 16(6), 2636. <https://doi.org/10.3390/en16062636>
- Chamishka, S., Madhavi, I., Nawaratne, R., Alahakoon, D., De Silva, D., Chilamkurti, N., and Nanayakkara, V. (2022). A Voice-Based Real-Time Emotion Detection Technique Using Recurrent Neural Network Empowered Feature Modelling. *Multimedia Tools and Applications*, 81(25), 35173–35194. <https://doi.org/10.1007/s11042-022-13363-4>
- Edén, A. S., Sandlund, P., Faraon, M., and Rönkkö, K. (2024). VoiceBack: Design of Artificial Intelligence-Driven Voice-Based Feedback System for Customer-Agency Communication in Online Travel Services. *Information*, 15(8), 468. <https://doi.org/10.3390/info15080468>
- Elghaish, F., Chauhan, J. K., Matarneh, S., Rahimian, F. P., and Hosseini, M. R. (2022). Artificial Intelligence-Based Voice Assistant for BIM Data Management. *Automation in Construction*, 140, 104320. <https://doi.org/10.1016/j.autcon.2022.104320>
- Fernandes, D., Garg, S., Nikkel, M., and Guven, G. (2024). A GPT-Powered Assistant for Real-Time Interaction with Building Information Models. *Buildings*, 14(8), 2499. <https://doi.org/10.3390/buildings14082499>
- Kim, J., Kim, W., Nam, J., and Song, H. (2020). "I Can Feel Your Empathic Voice": Effects of Nonverbal Vocal Cues in Voice User Interface. In R. Bernhaupt, F. Mueller, D. Verweij, and J. Andres (Eds.), *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–12). Association for Computing Machinery. <https://doi.org/10.1145/3334480.3383075>
- Madan, B. S., Zade, N. J., Lanke, N. P., Pathan, S. S., Ajani, S. N., and Khobragade, P. (2024). Self-Supervised Transformer Networks: Unlocking New Possibilities for Label-Free Data. *Panamerican Mathematical Journal*, 34(4), 194–210. <https://doi.org/10.52783/pmj.v34.i4.1878>
- Rzepka, C. (2019). Examining the use of voice assistants: A value-focused thinking approach. In G. Rodriguez-Abitia and C. Ferran (Eds.), *Proceedings of the 25th Americas Conference on Information Systems (AMCIS 2019)* (p. 20). Association for Information Systems.

- Saka, A., Taiwo, R., Saka, N., Salami, B. A., Ajayi, S., Akande, K., and Kazemi, H. (2024). GPT Models in Construction Industry: Opportunities, Limitations, And a use Case Validation. *Developments in the Built Environment*, 17, 100300. <https://doi.org/10.1016/j.dibe.2023.100300>
- Seaborn, K., Miyake, N. P., Pennefather, P., and Otake-Matsuura, M. (2021). Voice in Human-Agent Interaction: A Survey. *ACM Computing Surveys*, 54(4), 1–43. <https://doi.org/10.1145/3386867>
- Thirumuruganathan, S., Kunjir, M., Ouzzani, M., and Chawla, S. (2021). Automated Annotations for AI Data and Model Transparency. *Journal of Data and Information Quality*, 14(1), 1–9. <https://doi.org/10.1145/3460000>
- Wei, J., Tag, B., Trippas, J. R., Dingler, T., and Kostakos, V. (2022). What Could Possibly Go Wrong When Interacting with Proactive Smart Speakers? A Case Study Using an ESM Application. In S. Barbosa, C. Lampe, C. Appert, D. A. Shamma, S. Drucker, J. Williamson, and K. Yatani (Eds.), *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (pp. 1–15). Association for Computing Machinery. <https://doi.org/10.1145/3491102.3517432>
- Zheng, J., and Fischer, M. (2023). Dynamic Prompt-Based Virtual Assistant Framework for BIM Information Search. *Automation in Construction*, 155, 105067. <https://doi.org/10.1016/j.autcon.2023.105067>