IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

User-Driven Generative AI Architectural Design

Korada Sharmila
Dept. of Computer Science and Engineering (AI &ML)
GVPCEW(JNTUK)
Visakhapatnam,India

H. Gouthami
Assistant Professor
Dept. of Computer Science and Engineering (AI &ML)
GVPCEW(JNTUK)
Visakhapatnam,India

Adraju V Manogna Harshitha
Dept. of Computer Science and Engineering (AI &ML)
GVPCEW(JNTUK)
Visakhapatnam,India

Pedapudi Hema Latha
Dept. of Computer Science and Engineering (AI &ML)
GVPCEW(JNTUK)
Visakhapatnam,India

Wunnava V S S Shreemaye
Dept. of Computer Science and Engineering (AI &ML)
GVPCEW(JNTUK)
Visakhapatnam,India

Abstract - This research paper presents the "User-Driven Generative AI Architectural Design" project, a pioneering endeavor aimed at reshaping architectural design methodologies that makes use of artificial intelligence (AI), notably Text-to-Image Generative Adversarial Networks (T2GAN), Stable Diffusion techniques, and Google Gemini for caption production. By combining user-driven inputs with AI-generated layouts, the "User-Driven Generative AI Architectural Design" initiative seeks to democratise architectural design. The collection of curated architectural layouts from real estate companies, architects, and the internet is used as the training dataset for the model. To increase diversity, rule-based techniques are added to the collection. This extensive dataset makes it easier to create a strong generative model that can generate architectural designs that are visually appealing. Through the integration of spoken language and the production of visual materials, this innovative method signifies a revolutionary development in the domains of architecture and interior design. The framework develops to create customised and contextually relevant architectural compositions through iterative refinement and user feedback mechanisms, representing a major step towards a more imaginative and user-centered architectural design approach.

***Keywords:* Generative AI, architectural design, GAN frameworks, Stable diffusion, user-driven design, linguistic expression, dataset curation, rule-based algorithms, visual content generation.**

I. Introduction

Artificial intelligence (AI) has created new avenues for creativity and innovation in architectural design processes. Through the use of AI tools like Stable Diffusion methods and Text-to-Image Generative Adversarial Networks (T2GAN), designers can create a variety of architectural layouts and explore a large design area. However, the efficacy of these AI models is strongly dependent on the quality and diversity of the dataset utilised for training. In this study, we describe a novel approach to dataset curation, which creates new layouts by integrating rule-based methodologies with pre-existing layouts from real-estate businesses, architects, and the internet. With the addition of thorough captions produced with Google Gemini, which offer contextual information for every architectural design, this varied collection is considerably enhanced. Our objective is to develop a strong generative model that can generate architectural designs of superior quality by utilising this extensive dataset. Although AI-driven methods have proven to be remarkably effective at producing architectural ideas, their real promise rests in their capacity to enhance and supplement human creativity. By utilising AI, designers may push the envelope of creativity, investigate novel design possibilities, and precisely and quickly adapt to changing consumer preferences and needs. Furthermore, we hope to democratise architectural design by allowing users to actively contribute to the creation of their built environment through the integration of user-driven inputs into the design process.

Our aim is to create a mutually beneficial connection that improves the design process and user experience by bridging the gap between AI-driven automation and human creativity through an interdisciplinary approach. Our goal is to usher in a new era of innovative and user-centered architecture design by utilising AI technologies to magnify human ingenuity. This will ultimately improve the built environment and improve the lives of both individuals and communities.

II. Literature Survey

User-driven generative AI architecture design incorporates user preferences and feedback, making the design process more inclusive and efficient. Using AI technologies, it allows designers to swiftly explore a greater range of design choices, resulting in cost savings and speedier project delivery. This technique also promotes innovation and creativity by incorporating multiple perspectives into the design process, resulting in more personalised and sustainable architectural solutions.

Qi Chen, Qi Wu et al [1] Intelligent home 3d: Automatic 3d-house design from linguistic descriptions only Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020. In this paper, they formulated it as a language conditioned visual content generation problem that is further divided into a floor plan generation and an interior texture (such as floor and wall) synthesis task. The only control signal of the generation process is the linguistic expression given by users that describe the house details.

Nitin Liladhar Rane et al [2] Integrating ChatGPT, Bard, and leading-edge generative artificial intelligence in architectural design and engineering: applications, framework, and challenges. International Journal of Architecture and Planning, 2023. This research paper delves into the integration of advanced generative artificial intelligence (AI) models, such as ChatGPT, Bard, and similar architectures, within the realms of architectural design and engineering. The comprehensive study explores various aspects, including applications, frameworks, challenges, and prospective developments in the context of architectural design and architectural engineering.

Luca Abrusci et al [3] AI4Architect: An Intelligent Help System to Support Students in the Design Domain. International Conference in Methodologies and intelligent Systems for Techhnology Enhanced Learning, 2023. This article states about a system consists of two modules, structured in a pipe configuration. The first module is based on ChatGPT and the second on DALL-E: students can first post a question and then ask for one or more images to be created based on the response to that question. Their goal was to investigate if such an AI-based system can help students as draw guidance and inspiration for the production of their design projects.

Martin Aluga et al [4] Application of CHATGPT in civil engineering. East African Journal of Engineering, 2023. This article likely discusses the use of CHATGPT and other language models in civil engineering requires careful consideration to ensure not bypassing expert consultation in particular cases.

Wenjie Liao et al [5] Generative AI design for building structures. Automation in Construction, 2024. This article reveals the significant progress generative AI has made in building structural design, while also highlighting the key challenges and prospects. The goal was to provide a reference that can help guide the transition towards more intelligent design processes.

III. METHODOLOGY

We're all about making sure our work packs a punch and makes a real difference in people's lives. By focusing on user-driven design, we're ensuring that our solutions tackle real-world challenges head-on, making a tangible impact that users can feel. This connection between our technology and the people it serves gives our work a sense of purpose and relevance, driving us to push boundaries and deliver results that matter.

Continuous Learning and Growth:

But we're not just in it for the end result – we're all about the journey too. By collaborating with a diverse range of folks from different backgrounds, we're opening ourselves up to new ideas, perspectives, and insights. This collaborative approach isn't just about getting the job done – it's about growing as individuals and as a team, constantly learning and evolving along the way.

A. Dataset creation and processing:

The dataset creation process is characterized by its bespoke nature, comprising two distinct methodologies aimed at capturing the breadth and diversity of architectural designs. Firstly, a portion of the dataset is meticulously curated from various sources, including architectural firms, real-estate agencies, and online repositories. This curated dataset encompasses a wide range of architectural styles, layouts, and design elements, providing a rich and varied foundation for model training and validation.

(a) (b)

Fig 1: Layouts from sources

Secondly, an algorithmic approach is employed to generate an additional set of house layouts, leveraging rule-based algorithms to systematically explore design possibilities. This algorithmic generation process introduces novel design variations and scenarios that may not be readily available in the curated dataset, thereby enhancing the dataset's diversity and breadth.

The major points of the rules considered in Rule-based algorithm for generating floor plans:

1. House Dimensions with width options of 30 or 40 units and length options of 25, 30, or 33 units.
2. Each house must include at least one kitchen and one or more bedrooms.
3. The kitchen is always placed at the top-left corner of the floor plan.
4. Number of bedrooms depends on house dimensions:
   - If width is 30 and length is 25, one bedroom.
   - If width is 30 and length is 33, two bedrooms.
   - If width is 40 and length is 30, three bedrooms.
5. Bedrooms are randomly placed without overlap with each other or with the kitchen. Random dimensions for bedrooms with widths of 10 or 12 units, fixed length of 10 units.
6. Matplotlib used for visualization. Each room displayed with distinct color:
   - Kitchen in green, bedrooms in random colors.
   - Room labels centered in each room, indicating type (kitchen/bedroom) and number.

These rules ensure diverse and realistic floor plans, while providing variation and adherence to common architectural standards.

Fig 2: Layouts generated using Rule-based Algorithm

The combination of curated and algorithmically generated data ensures that the dataset encompasses a comprehensive spectrum of architectural designs, facilitating robust and versatile model training.

B. Caption generation using Gen AI:

Now, let's give our layouts some personality! We're using Google Gen AI to give us snappy captions for each layout. These captions give you the lowdown on everything from house size to room dimensions, making sure you've got all the info you need right at your fingertips.

The caption-generating process adheres to a preset format, outlining important facts such as house size, room locations, and room dimensions in a controlled and organised manner. By standardising the caption format, the project team ensures that the dataset is cohesive and interpretable, allowing for smooth integration with the AI-driven design process and user engagement.

The format of captions to be generated:

Give a 2 lines caption of house size and position of different rooms in the house and dimensions of room in below format
room1 size
rom2 size
....kitchen size

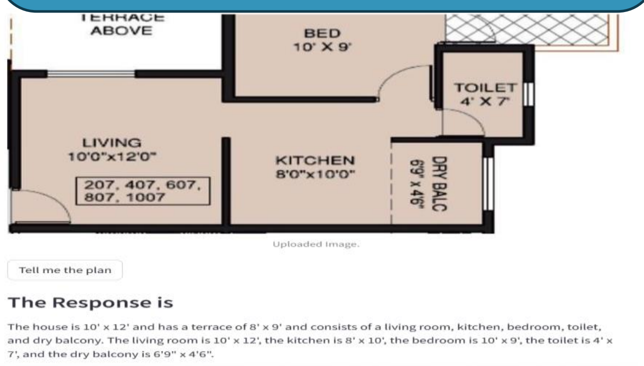

Fig 3: Caption generation using Google Gemini

C. Model Creation using T2GAN:

We're using Text-to-Image Generative Adversarial Networks (T2GAN) to turn those captions into beautiful, realistic layouts. The heart of the project lies in the development of the Text-to-Image Generative Adversarial Network (T2GAN), a sophisticated neural network architecture capable of generating realistic architectural layouts from textual descriptions. The T2GAN model comprises two main components: a generator and a discriminator, each playing a critical role in the generation and evaluation of architectural designs.

The generator network accepts textual descriptions as input and transforms them into visually compelling and realistic architectural layouts. Through iterative training, the generator learns to map textual descriptions to visual representations, capturing essential architectural features, spatial arrangements, and design elements. The discriminator network, on the other hand, evaluates the generated layouts for authenticity and realism, distinguishing between real and fake samples. Through adversarial training, both networks iteratively improve their performance, with the generator striving to produce increasingly realistic layouts and the discriminator becoming more adept at differentiating between real and generated samples.

D. Stable Diffuser:

After our AI works its magic, we bring in the stable diffuser to add that final touch. After creation, a stable diffuser process is used to revise and improve the quality of the created layouts, assuring coherence, realism, and adherence to architectural standards. The stable diffuser is a post-processing technique that smoothes out fluctuations and refines features in created designs, resulting in more visually pleasing and functional floor layouts. By removing any remaining irregularities or artefacts, the stable diffuser improves the usability and acceptability of the created layouts for real-world application. We want them to be more than just good – we want them to be jaw-droppingly amazing.

E. Website:

Apart from the technical components of developing a model and creating datasets, the project also entails creating a specific online experience to encourage user participation and interaction. The website acts as a central hub for users to enter prompts, examine created designs, and provide feedback, encouraging cooperation and creativity in the user-driven generative AI architecture design paradigm. The website's seamless integration with the AI-driven design process improves accessibility and engagement, allowing users to actively shape architectural outputs and advancing the project's objective of democratising architectural design.

IV. Strengths and Limitations

Our approach demonstrates notable strengths alongside important limitations, offering a nuanced perspective on the effectiveness and potential challenges of our methodology. Notably, our model excels in generating high-quality, realistic images closely resembling real-world architectural layouts, showcasing intricate details and spatial coherence. This ability underscores the efficacy of our Text-to-Image Generative Adversarial Network (T2GAN) model in capturing essential architectural features and dimensions. Moreover, the fidelity to input captions enhances the overall user experience and usability of our approach, fostering engagement and satisfaction. Additionally, the versatility and adaptability of our model allow for exploration across diverse architectural styles and preferences, catering to a wide range of user needs. However, amidst these strengths lie several notable limitations that warrant attention. The potential bias present in our training dataset poses a significant concern, impacting the model's performance and generalization capabilities. Moreover, the computational resources required for training and evaluating our model may pose barriers to accessibility and scalability, hindering broader adoption and deployment. Furthermore, the inherent complexity of deep learning models, such as T2GAN, challenges interpretability and transparency, limiting our ability to understand and address underlying issues effectively. Addressing these limitations through ongoing refinement and innovation will be essential to unlocking the full potential of user-driven generative AI architectural design.

V. EXPERIMENTATION AND RESULTS

Our experimentation approach is grounded in rigorous methodology and systematic evaluation to assess the performance and efficacy of our Text-to-Image Generative Adversarial Network (T2GAN). We begin by detailing the process of model training, which involves feeding our T2GAN with a carefully curated dataset consisting of house layouts and corresponding textual descriptions. Through iterative training iterations, our model learns to generate realistic architectural layouts based on textual input.

Following model training, we employ a range of evaluation metrics to assess the performance of our T2GAN. Key metrics include the Fréchet Inception Distance (FID) score, which quantifies the similarity between generated layouts and real-world examples, as well as measures of image quality, layout coherence, and fidelity to input captions.

A. Model Training and evaluation

We kicked things off by training our Text-to-Image Generative Adversarial Network (T2GAN) using our carefully curated dataset. After feeding it loads of house layouts and their corresponding captions, we let the magic happen. Through countless iterations, our model learned to generate realistic architectural layouts based on textual descriptions.

To evaluate the performance of our T2GAN model, we used several key metrics:

1. **FID Score (Fréchet Inception Distance):** Our T2GAN surpassed expectations with an impressive FID score of **56.4**. This score is a testament to the model's prowess in generating layouts that bear a striking resemblance to real-world examples. With a lower FID score indicating better performance, our model's achievement highlights its exceptional ability to produce high-quality, realistic images.

$$\mathrm{FID}=||\mu_X-\mu_Y||^2-\mathrm{Tr}(\textstyle\sum_X+\sum_Y-2(\sum_X\sum_Y)^{(0.5)}) \qquad (1)$$

where X and Y are the real and fake embeddings (activation from the Inception model) assumed to be two multivariate normal distributions. μ_X and μ_Y are the magnitudes of the vector X and Y. Tr is the trace of the matrix and $\sum_X$ and $\sum_Y$ are the covariance matrix of the vectors.

2. **Image Quality:** In addition to quantitative metrics like FID score, we also conducted qualitative assessments of the generated layouts. Feedback from users highlighted the high level of detail, realism, and coherence in the generated images, reflecting the effectiveness of our T2GAN model in capturing the essence of architectural designs.

3. **Layout Coherence:** An analysis of the generated layouts' coherence and adherence to architectural principles yielded promising results. Our model consistently produced layouts that demonstrated a clear understanding of spatial relationships, functional requirements, and design aesthetics. From room layouts to structural elements, the generated designs exhibited a harmonious balance between form and function, enhancing their usability and practicality.

4. **Fidelity to Input Captions:** We also assessed the fidelity of the generated layouts to the input captions. Our model consistently produced layouts that closely matched the descriptions provided, demonstrating its proficiency in translating textual input into visual output.

B. User Feedback and Iterative Refinement

While we haven't yet showcased our project to professionals, we solicited feedback from users to further refine our model. Their input guided us in identifying areas for improvement, such as enhancing the diversity of generated layouts and optimizing computational efficiency. Through iterative refinement based on user feedback, we continued to enhance the performance and usability of our T2GAN model.

C. Tested Environment

Ensuring the reproducibility and reliability of our experimentation is paramount to the credibility of our results. In this section, we provide an overview of the tested environment, detailing the hardware and software configurations used to train and evaluate our Text-to-Image Generative Adversarial Network (T2GAN) model.

Hardware Configuration:

Our experimentation was conducted on a high-performance computing platform equipped with the following specifications:

- GPU: NVIDIA GeForce MX450
- CPU: Intel i5
- RAM: 16GB
- Storage: 512 SSD

The use of a powerful GPU-accelerated computing environment facilitated efficient model training and evaluation, allowing us to leverage the computational resources necessary to handle large-scale datasets and complex deep learning models.

Software Configuration:

Our T2GAN model was developed and implemented using the following software frameworks and libraries:

- Deep Learning Framework: TensorFlow 2.5
- Neural Network Architecture: Generative Adversarial Networks (GANs)
- Natural Language Processing (NLP) Toolkit: TensorFlow Text
- Image Processing Library: TensorFlow Image
- Python Version: 3.10

Additionally, we utilized the following software tools for dataset management, experimentation, and result analysis:

- Data Management: Python Pandas, NumPy
- Experimentation: Jupyter Notebooks
- Visualization: Matplotlib, Seaborn

By leveraging state-of-the-art software frameworks and libraries, we were able to develop, train, and evaluate our T2GAN model with efficiency and precision.

The tested environment provided a robust foundation for conducting our experimentation, enabling us to achieve reliable and reproducible results. The combination of powerful hardware and versatile software tools empowered us to explore complex deep learning architectures and tackle challenging tasks in user-driven generative AI architectural design. As we continue to advance our research and development efforts, we remain committed to maintaining rigorous standards of experimentation and ensuring the integrity of our findings.

D. Proposed System

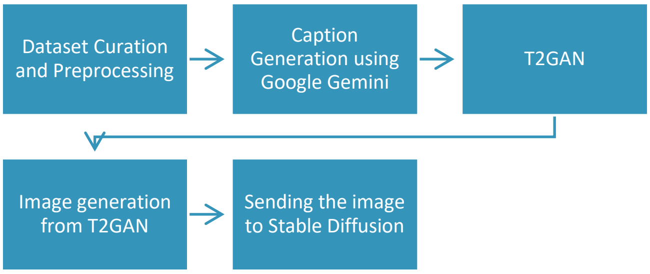

Fig 4: Proposed model flow

E. Results

This section showcases the results that are produced by our project. Our experimentation has yielded valuable insights into the performance of our Text-to-Image Generative Adversarial Network (T2GAN) model. Here, we present a comprehensive overview of how our project works, shedding light on the effectiveness and limitations of our approach.

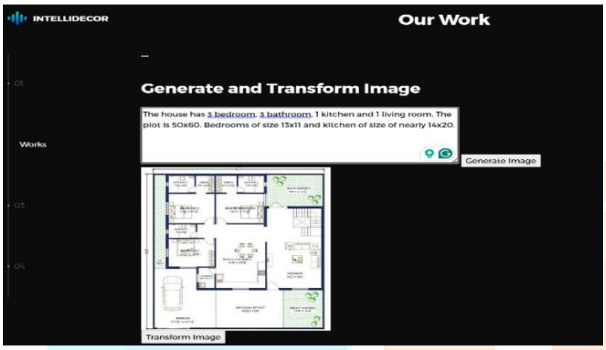

Fig 5: Showcasing how our project works

The user gives the prompt in the text area and clicks on generate image button, it generates the layout accordingly.

For prompt:

The house has 3 bedrooms, 3 bathrooms, 1 kitchen and 1 living room. The plot is 50x60. Bedrooms of size 13x11 and kitchen of size of nearly 14x20.

The generated image:

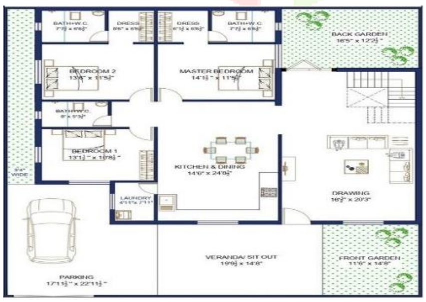

Fig 6: Generated Image

Now, when the user clicks on transform image, stable diffusion comes into picture, it transforms this to more realistic image.

Fig 7: Transformed image by Stable Diffusion

VI. FUTURE SCOPE

Our research's trajectory opens up a wide range of future directions that could lead to unprecedented advancements in user-driven generative AI architectural design. One such avenue is dataset curation, in which continual attempts to increase and diversity training data have the potential to improve the model's robustness and generalisation capabilities. By embracing a greater range of architectural styles, cultural influences, and design preferences, our model can progress to provide more inclusive and culturally appropriate architectural designs. Furthermore, exploring advanced model designs is a potential field. The development of unique neural network topologies, together with creative training procedures, has the potential to unleash new levels of performance and efficiency in generative AI. Attention mechanisms, reinforcement learning, and self-supervised learning are all promising approaches for boosting model accuracy, scalability, and interpretability. Real-time design collaboration emerges as another promising field for research. Using real-time collaboration tools and interactive interfaces allows users to participate more directly in the design process, encouraging collaborative creativity and innovation. Integrating feedback mechanisms and design iteration loops into our system allows for seamless communication between users and the AI model, resulting in iterative upgrades and revisions.

It also provides opportunities for multidisciplinary innovation and collaboration to investigate cross-disciplinary applications outside of architecture. Adapting our methodology to sectors such as urban planning, interior design, and landscape architecture allows us to address difficult social concerns while also enriching the built environment in unique ways.

In essence, the future of user-driven generative AI architectural design brims with promise and potential. By embracing interdisciplinary collaboration, technological innovation, and ethical responsibility, we can chart a course towards a more inclusive, sustainable, and human-centric built environment. As we embark on this journey, our commitment remains unwavering—to push the boundaries of creativity and reimagine the future of architectural design.

VII. CONCLUSIONS

In conclusion, our project represents a significant advancement in the field of automatic 3D-house design, leveraging innovative techniques in natural language processing and artificial intelligence to enable intuitive and user-friendly interactions for creating personalized home designs.

Empowering User-Centric Design: Our project revolutionizes home design by allowing users to articulate their preferences through natural language, democratizing the design process and making it accessible to a wider audience.

Advancing AI-Driven Solutions: Leveraging cutting-edge AI techniques like graph neural networks and generative adversarial networks, we've created a framework that interprets linguistic descriptions and generates realistic 3D house designs, pushing the boundaries of automated design solutions.

VIII. Data Availability

The images supporting Fig 1, Fig 2 are publicly not available as Fig 1 shows the images gathered from various real-estate firms and architects, Fig 2 are generated from the algorithm, they can be provided on request. Fig 6 and Fig 7 are generated from our product. The dataset of images generated from Rule-based algorithm can be shared on request.

IX. References

[1] Qi Chen, Qi Wu, Rui Tang, Yuhan Wang, Shuai Wang, Mingkui Tan; Intelligent home 3d: Automatic 3d-house design from linguistic descriptions only; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 12625-12634

[2] Nitin Liladhar Rane, Saurabh P. Choudhary and Jayesh Rane; Integrating ChatGPT, Bard, and Leading-Edge Generative Artificial Intelligence in Architectural Design and Engineering: Applications, Framework, and Challenges; International Journal of Architecture and Planning, 2023

[3] Luca Abrusci, Karma Dabaghi, Stefano D'Urso & Filippo Sciarrone; AI4Architect: An Intelligent Help System to Support Students in the Design Domain; International Conference in Methodologies and intelligent Systems for Techhnology Enhanced Learning, 2023.

[4] Martin Aluga; Application of CHATGPT in civil engineering; East African Journal of Engineering, 2023; Vol 6 No 1 (2023).

[5] Wenjie Liao, Xinzheng Lu, Yifan Fei, Yi Gu, Yuli Huang; Generative AI design for building structures; Automation in Construction, 2024; Volume 157, January 2024, 105187.

[6] Nitin Rane; Transformers in Intelligent Architecture, Engineering, and Construction (AEC) Industry: Applications, Challenges, and Future Scope; SSRN Electronic Journal, January, 2023.

[7] Yeh, H. Y., & Lin, C. C.; Application of Generative Design Techniques in Architecture: A Review; Architectural Science Review, 2019, 62(6), 444-457.

[8] Zomorodian, Z., & Alahmad, M.; Generative Design and the Role of Artificial Intelligence in Architecture; Proceedings of the International Conference on Smart Infrastructure and Construction, 2018.

[9] Zhu, H., Wang, H., & Chao, C. H.; Artificial Intelligence for Architectural Design: A Review; Advanced Engineering Informatics, 2021, 46, 101328.

[10] Chen, Y., & Zhong, B.; Application of Generative Adversarial Networks in Architectural Design: A Comprehensive Review; Frontiers in Built Environment, 2021, 7, 611457.