MIT School of Engineering
Department of Computer Science and Engineering

Project Synopsis

Group ID: TYAI102

Project Title: Gen-AI in Blender: AI-driven prompt-based image generation

Group Members:

Enrollment Number	Roll No.	Name of student	Email Id	Contact Number
MITU21BTCS0221	2213219	Harkeerat Dhunda	harkeeratdhunda@gmail.com	8830627095
MITU21BTCS0718	2213256	Vilakshan	vilakshanjoshi7778@gmail.com	6280302674

Problem Statement:

Current limitations in Blender's integration for AI-driven image generation create a positive opportunity to enhance the user experience, enabling a smoother and more accessible incorporation of AI-generated content in 3D design projects.

Abstract:

This project aims to rethink Blender's integration for AI-powered image production by strategically incorporating powerful Generative AI technology. The major goal is to create a novel, user-friendly prompt-based add-on that works smoothly with Generative AI algorithms, overcoming existing problems and pushing the frontiers of 3D design.

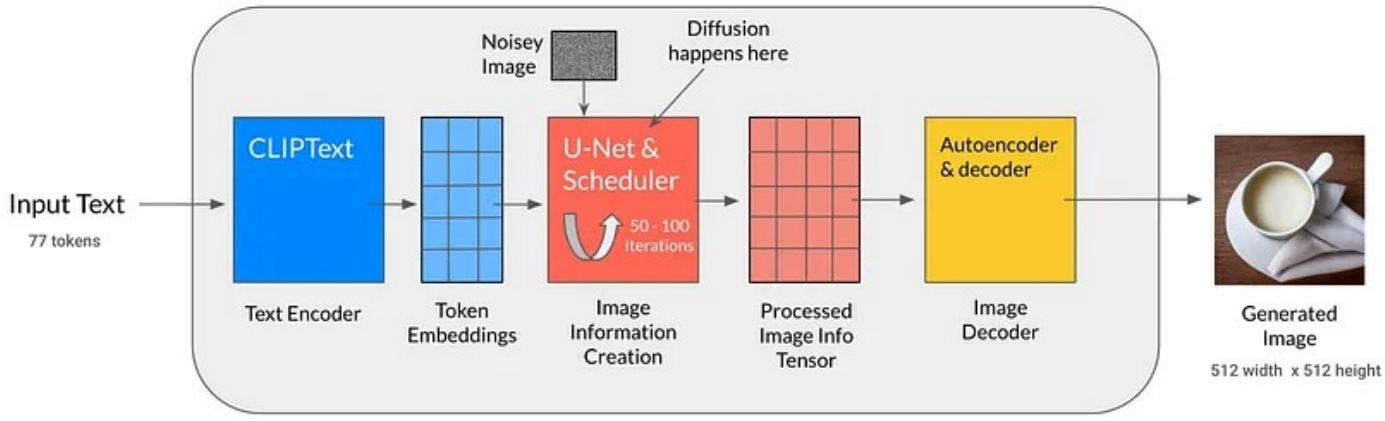
The add-on serves as a transformative conduit between human creativity and the nuanced capabilities of AI. By providing artists with a sophisticated tool, it facilitates the effortless integration and stylization of AI-generated content into their designs. The approach is not merely about addressing current hurdles but actively reshaping the landscape of 3D designs.

This innovative project involves integrating cutting-edge Generative Adversarial Networks (GANs) into the Blender environment. The heart of our technique is built around GANs, which are known for their capacity to generate realistic and diverse material. Using models such as StyleGAN, we hope to empower artists to create distinctive and visually beautiful 3D designs using simple instructions.

Literature Survey: Detail survey done

Year	Method	Results	Drawbacks	Author
2023	Utilization of Diffusion or GAN image generation methods (Stable Diffusion XL, GigaGAN).	GAN methods better for high-frequency details, with advantages in speed and precision.	Maintaining consistency across multiple images in 3D generation.	Song Bai, Jie Li
2023	Readily available generative AI on artistic practices	Focus on prompt engineering and design in the artistic process	Ready-made models might constrain the level of personalization	James Hutson, Bryan Robertson
2023	Taxonomy of Prompt Modifiers	Six types identified: subject terms, image prompts, style modifiers, quality boosters, repeating terms, and magic terms.	Limited overall working time to about 2 hours per day, depending on computation	Jonas Oppenlaender
2022	Translating Raw Descriptions into Images by Prompt-based Cross-Modal Generation	Proposed PCM-Frame leveraging CLIP and StyleGAN pre-trained models.	Challenges in handling extremely abstract or subjective terms	Yiyang Ma, Huan Yang, Bei Liu
2021	Generative Adversarial Networks for Image and Video Synthesis	. Deep Boltzmann Machines (DBMs) use MCMC sampling, but scalability may be limited. Variational AutoEncoders (VAEs) excel in learning latent representations but might produce slightly blurry images.	Evaluating and comparing GAN models is challenging due to instability.	Liu, M.-Y., Huang, X., Yu, J., Wang, T.-C., & Mallya, A.
2020	Generative Art: Between the Nodes of Neuron Networks	GANs provide an endless original making machine	Challenges arise in defining authorship and artistic value for generative art.	Bruno Caldas Vianna

Proposed System (Block Diagram):



Conclusion:

This project aims to integrate cutting-edge Generative AI, specifically GANs like StyleGAN, into Blender for prompt-based image generation in 3D designs. The goal is to overcome existing challenges and redefine the 3D design landscape, providing artists with a seamless tool to incorporate and stylize AI-generated content.

The literature survey highlights GANs' efficiency in handling high-frequency details, the importance of prompt engineering in artistic practices, and the limitations of off-the-shelf models in terms of personalization. These insights guide the project's design and development.

The proposed system integrates GANs into Blender, prioritizing user-friendly prompt-based image generation. It addresses drawbacks such as limited working time and challenges in handling abstract terms, ensuring effectiveness and usability in the creative workflow.

The annexures include essential documentation for project approval, market and financial feasibility, and relevant literature survey papers and links for further reference. This comprehensive approach ensures a well-structured and informed project execution.

Annexes:

Annexure I: Form A-Title Approval (for offline mode)

Annexure II: Form B-Market and financial feasibility (verify from guide)

Annexure III: Literature survey paper or links

1. Hutson, J., & Robertson, B. (2023). Exploring the Educational Potential of AI Generative Art in 3D Design Fundamentals: A Case Study on Prompt Engineering and Creative Workflows. Global Journal of Human-Social Science: Arts & Humanities - Psychology. https://doi.org/10.1145/3503161.3547790
2. Oppenlaender, J. (2023). A Taxonomy of Prompt Modifiers for Text-To-Image Generation. https://digitalcommons.lindenwood.edu/faculty-research-papers/485/
3. Kang, S., et al. (2023). Scaling Up GANs for Text-to-Image Synthesis. In Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). https://openaccess.thecvf.com/content/CVPR2023/html/Kang_Scaling_Up_GANs_for_Text-to-Image_Synthesis_CVPR_2023_paper.html
4. Sagescribe, A. (n.d.). Generative AI for Image Generation: SOTA, Common Methods. Medium. https://medium.com/@aisagescribe/generative-ai-for-image-generation-sota-common-methods-bea0a70c9b81
5. DeepSense.ai. (n.d.). Diffusion Models in Practice: Part 1 - The Tools of the Trade. DeepSense.ai. https://deepsense.ai/diffusion-models-in-practice-part-1-the-tools-of-the-trade/
6. Vianna, B. C. (2020). Generative Art: Between the Nodes of Neuron Networks. Artnodes. https://raco.cat/index.php/Artnodes/article/view/374003/467524