A PROJECT REPORT ON

"GenAI in Blender: AI driven prompt based image generation."

(SUBMITTED TO MIT SCHOOL OF COMPUTING, LONI, PUNE IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE)

BACHELOR OF TECHNOLOGY
(Computer Science & Engineering)

BY

Candidate Name:	Enrollment No.:
Harkeerat Dhunda	**MITU21BTCS0221**
Vilakshan	**MITU21BTCS0718**

Under the guidance of

Prof. Sneha Singha

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

MIT SCHOOL OF COMPUTING
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
MIT ART, DESIGN AND TECHNOLOGY UNIVERSITY,
RAJBAUG CAMPUS, LONI-KALBHOR, PUNE 412201

CERTIFICATE

This is to certify that the project report entitled

“GenAI in Blender: AI driven prompt based image generation”

Submitted by

Name of the Candidate:	PRN/ Enrollment No:
Harkeerat Dhunda,	**MITU21BTCS0221**
Vilakshan	**MITU21BTCS0718**

is a bonafide work carried out by them under the supervision of Prof./Dr/ Vrushali Kondhalkar and it is submitted towards the partial fulfillment of the requirement of MIT ADT university, Pune for the award of the degree of Bachelor of Technology (Computer Science and Engineering)

Prof. Sneha Singha	**Dr. Shradha Phansalkar**
Guide	Head of Department

DECLARATION

We, the team members

Name	Enrollment No
Harkeerat Dhunda	**MITU21BTCS0221**
Vilakshan	**MITU21BTCS0718**

Hereby declare that the project work incorporated in the present project entitled is original work. This work (in part or in full) has not been submitted to any University for the award or a Degree or a Diploma. We have properly acknowledged the material collected from secondary sources wherever required. We solely own the responsibility for the originality of the entire content.

Date: 29/04/2024

Name & Signature of the Team Members

Member 1: _____

Member 2: _____

Name and Signature of Guide

Seal/Stamp of the College

Place: Pune

Date: 02/11/23

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
MIT SCHOOL OF COMPUTING,
RAJBAUG, LONI KALBHOR,
PUNE – 412201

EXAMINER'S APPROVAL CERTIFICATE

The project report entitled "**GenAI in Blender: AI driven prompt based image generation**" submitted by Harkeerat Dhunda (MITU21BTCS0221**)**, Vilakshan (MITU21BTCS0718) in partial fulfillment for the award of the degree of Bachelor of Technology (Computer Science & Engineering) during the academic year 2021-22, of MIT-ADT University, MIT School OF COMPUTING, Pune, is hereby approved.

Examiners:

1.

2.

ACKNOWLEDGEMENT

It gives us great pleasure in presenting the project report on 'GenAI in Blender: AI driven prompt based image generation.'

We would like to take this opportunity to thank our internal guide Prof. Sneha Sngha for giving us all the help and guidance we needed. We are really grateful for her support. Her valuable suggestions were very helpful.

We are also grateful to Dr. Ganesh Pathak (Head of Computer Science & Engineering) for indispensable support, suggestions.

Name, Enrollment No.

Harkeerat Dhunda, MITU21BTCS0221

Vilakshan MITU21BTCS0718

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.

Table of Contents

Chapter 1: INTRODUCTION

1.1 Introduction

Now more than ever, with the possibilities that are happening in creativity and the huge technological growth, there arises the need to boost the capabilities of Blender with the most recent state-of-the-art AI. At the end of the day, our core objective is to have a very clear link between human creativity and the nuanced capabilities of the AI that usher in the design revolution.

At the heart of our endeavor lies the utilization of Generative Adversarial Networks (GANs), a subcategory of machine learning models that can create life-like, varied material. Our program, using models like StyleGANs and StabilityAI, is meant to empower creators to produce unique, eye-catching designs with easy instructions.

Key innovations in machine learning include algorithmic generative artificial intelligence algorithms that create unique data samples similar to their training examples across many forms of media: from text and photos to movies and 3D models, examples abound in art, engineering, and imaging. Reference [1] shows Generative AI growth has seen scales in its abilities increase proportionately, such that a program like OpenAI's ChatGPT amassed over a hundred million users in a very short period after its debut.

Democratization of these tools has lowered the user barrier, and text-guided image generation is now available to people of all technicalities. So, in effect, what this has done is that platforms like Dream Studio by Stability AI or OpenAI's DALL-E give users intuitive front ends.

Our effort is to leverage the closeness of human creativity with AI ingenuity, break the restraints of creation, and open a new age in artistic expression.

1.2 Scope:

The scope of this project encompasses the integration of state-of-the-art AI technologies, specifically Generative Adversarial Networks (GANs), into the Blender software platform. The project aims to enhance Blender's capabilities by leveraging AI algorithms such as StyleGANs and StabilityAI to facilitate the generation of lifelike and diverse content, including images, textures, and 3D models.

Key components of the project's scope include:

1. Research and Development: Conducting thorough research on the latest advancements in AI, particularly in the field of generative artificial intelligence, to identify suitable algorithms and methodologies for integration into Blender.
2. Algorithm Selection and Implementation: Selecting appropriate AI models, such as StyleGANs and StabilityAI, and implementing them within the Blender environment to enable the generation of high-quality, visually appealing content.
3. User Interface Design: Designing intuitive user interfaces that enable creators of varying technical backgrounds to interact seamlessly with the AI-driven features within Blender. This includes developing user-friendly tools for inputting instructions and parameters for content generation.
4. Training and Optimization: Training the AI models using diverse datasets to ensure their ability to generate diverse and realistic content across different styles and genres. Additionally, optimizing the performance of the AI algorithms to enhance speed and efficiency during content generation.
5. Testing and Validation: Conducting rigorous testing and validation procedures to assess the accuracy, robustness, and usability of the AI-integrated features within Blender. This involves soliciting feedback from beta testers and incorporating necessary adjustments based on user input.
6. Documentation and Training: Providing comprehensive documentation and training materials to guide users in utilizing the AI-driven features effectively within Blender. This includes tutorials, user manuals, and troubleshooting guides to support users at various skill levels.

7. Community Engagement: Engaging with the Blender community to gather insights, address concerns, and foster collaboration in advancing the integration of AI technologies within the platform. This involves soliciting feedback through forums, social media channels, and developer communities.

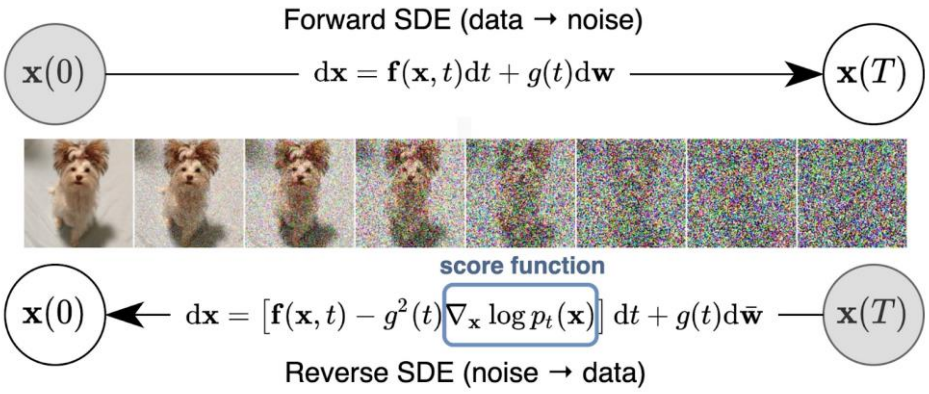
The project's scope is aimed at empowering creators to unleash their creativity and enhance their productivity through the seamless integration of AI-driven tools within the Blender software ecosystem. By expanding Blender's capabilities with cutting-edge AI technologies, the project endeavors to catalyze innovation and redefine the boundaries of artistic expression in the digital realm.

Chapter 2: CONCEPTS AND METHODS

2.1 Introduction to Diffusion Models

In the progressive aspect, diffusion models prove most effective in denoising Gaussian noise data. Denoiser fitting training of D under this framework is performed, and the approach can be easily applied to a wide range of formalisms for diffusion models. Denoiser training is performed using score matching to minimize the difference between the denoised output and the clean image. Realistic samples are obtained by numerically integrating an SDE backward from t=1 to t=0.

Figure 1: Diffusion Model (Source: Reference [2])



2.2 Principles of Diffusion Distillation

The main aim of the diffusion distillation techniques remains to simplify the training of diffusion models by distilling the complex model into the simple one. It involves training the simple model in such a way that it imitates the behavior of the complex model, mostly by matching the output distribution of denoised images. There are numerous techniques put forth to achieve efficient model distillation; for example, Progressive Distillation and Consistency Distillation.

2.3 Leveraging Latent Diffusion Models

Reference [3] shows that latent diffusion models are a subclass of diffusion models that work in a low-dimensional latent space. The use of a latent space lowers the memory requirement of LDMs and scales up to enormous model sizes and resolutions. One of the recent advancements with the LDMs—the MMDiT family—has shown competitive results in text-to-image synthesis.

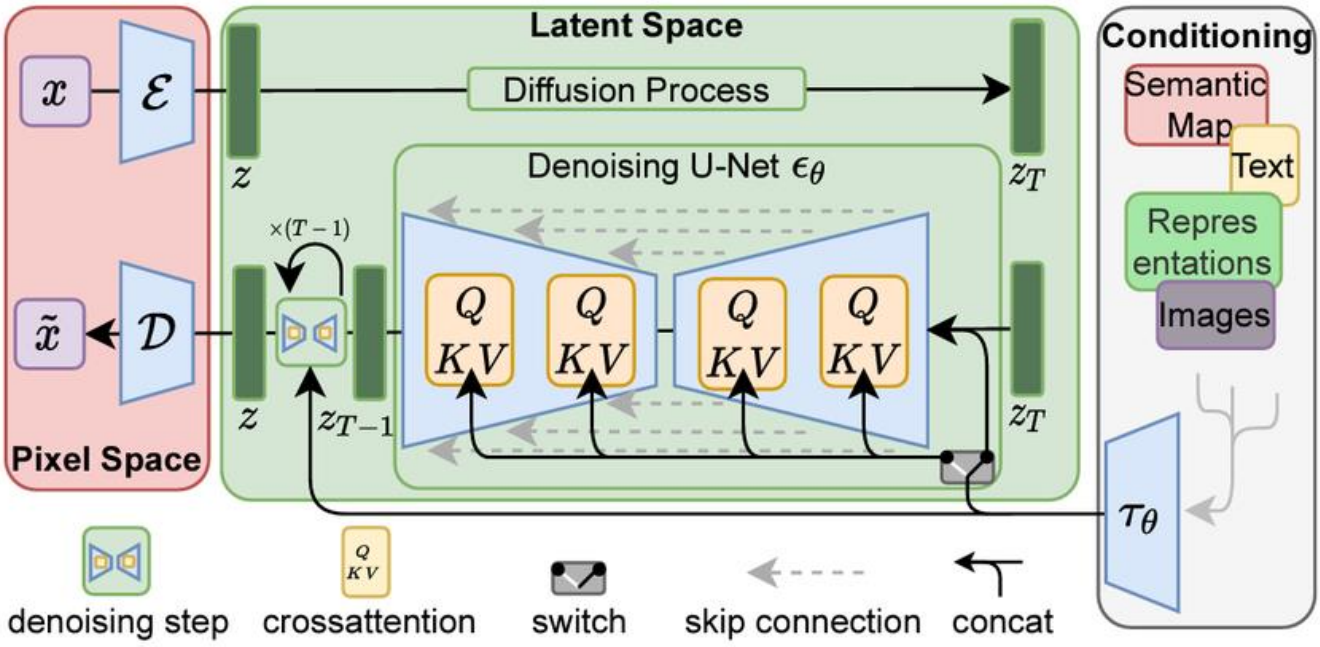


Figure 2: Latent Diffusion Model (Source: Reference [3])

2.4 Application in StabilityAI API

The described approach, containing the diffusion models and distillation techniques, underlies the image generation capabilities of the StabilityAI API. StabilityAI leverages these principles to enable a user-empowered environment equipped with state-of-the-art tools to create high-quality images across numerous domains.

2.5 Approaches to Implementation

Different strategies can be considered when it comes to integrating techniques for image generation like Stable Diffusion within Blender. One way to do it would be to integrate the image-generation models, such as Stable Diffusion, directly into Blender, and use pre-trained weights from the model. Then get the image from the model through inference of the model. The second approach would be to develop a custom interface in Blender for interactive guidance of the image-generating process from textual prompts or other sources of user input. The third approach would write a separate application with an interface to Blender.

2.6 Selection of Approach

After careful consideration, the decision was made to develop a custom interface within Blender for image generation using techniques like Stable Diffusion. This approach was chosen for its potential to provide users with a seamless and intuitive environment for generating images directly within Blender, utilizing the power of state-of-the-art image generation models.

2.7 Utilization of Blender Python API

The Blender Python API has added a facility for using the power of Hugging Face models in Blender, by making use of the 3D rendering environment and cutting-edge natural language processing. Being an open-source software, the links being developed between the two worlds of 3D graphics and deep learning are pretty useful, thus providing new paths for creative expression and exploration.

2.8 Creation of the Preferences Panel

This panel is the central point in Blender from which the user can access and change every setting related to the workflow for AI-based image creation. It has iteratively considered user feedback and design in such a way that it is optimized for an intuitive and user-friendly way to easily adjust parameters and options according to individual needs.

2.9 Functionality of the Preferences Panel

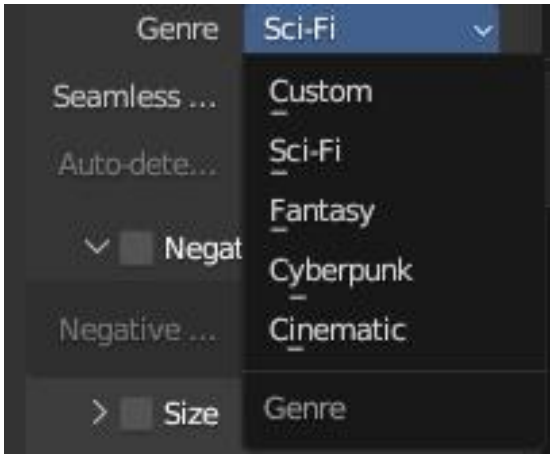


Figure 3: Genre Options in Preferences Panel

The preferences panel is a pivotal element in our project, crucial for user experience and permitting customization inside Blender. It is an easy and useful user-friendly interface accessible to any user. It allows advanced functionality through such a panel, allowing one to choose AI models and even integrate language prompts for users to be able to elaborate on their creativity. It rethinks image creation with amalgamated characteristics. Its features are summarized below.

- Model Selection: One can choose from a variety of Generative AI models, ranging from pre-built to custom.
- HuggingFace Prompt: Integration with Hugging Face allows users to utilize powerful language models for generating image prompts.
- Custom Prompt: Integration with the Hugging Face enables users to generate image prompts using powerful language models.
- Genre Selection: Users are allowed to provide genres of the images they wish, which may include things like sci-fi, fantasy, cyberpunk, and environment types like landscapes, concept arts, and photography.
- Integration: The panel integrates perfectly with the Blender interface to ensure a seamless user experience and smooth workflow.
- Auto-detection: Automatically detect parameters and settings, such as image size and source image, to enable easy rendering of images by users.

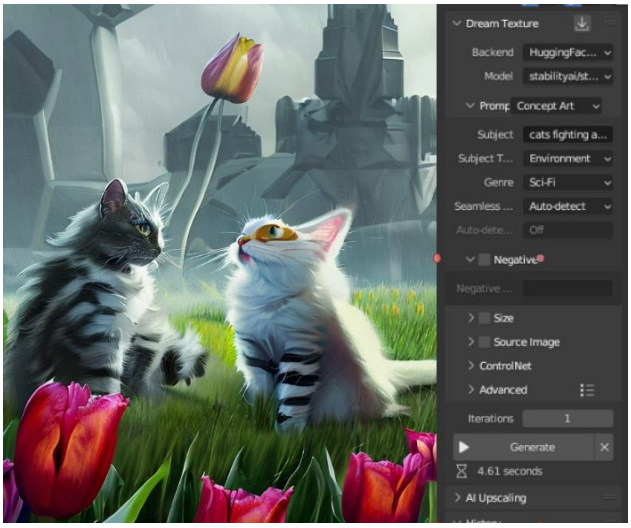


Figure 4: Preferences Panel with an example of cats as subjects in an environment

Chapter 3: LITERATURE SURVEY

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

Table 3.1: Literature Sure

Chapter 4: RESULTS

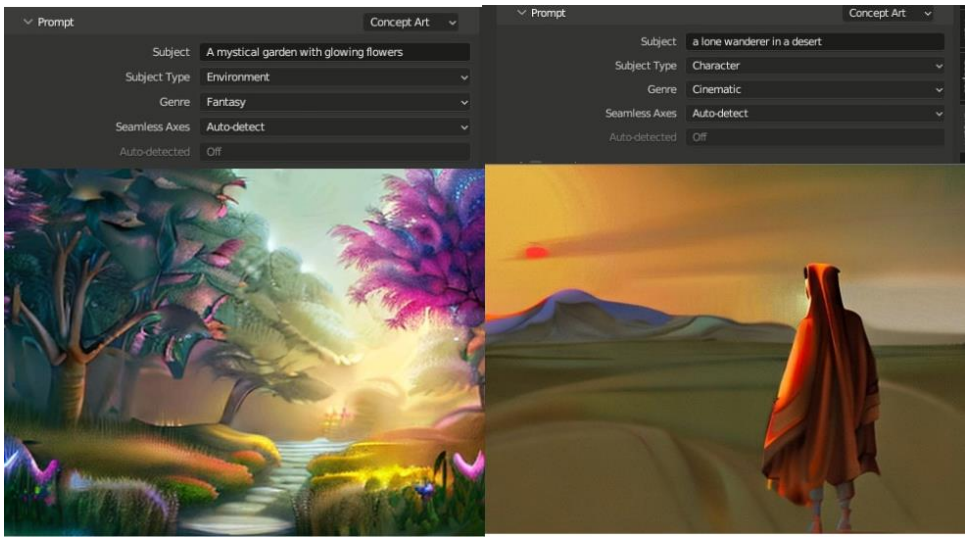


Figure 5: Final results in Blender of the integration with two prompts: “A mystical garden with glowing flowers” and “A lone wanderer in the desert.”

4.1 Visual Quality:

High visual quality involves sharp details of the images with vivid colors and realistic textures. The glowing flowers in the mystic garden and the lone wanderer in the desert are both clear, and in concordance with the input description. It shows fine-grained detail, where the AI model detects complex patterns on the glowing flowers and a large expanse in the desert landscape.

4.2 Coherence and Consistency:

The created images are very consistent and coherent with the provided description, so the AI model is working in the way that it actually understands and properly processes the provided textual prompts. Glowing flowers are integrated into the scene of the magical garden around them, which creates a feel of surreal enchantment.

4.3 Artistic Interpretation:

The palette, the lighting effects, and equally the composition of the scenes bring the emotion in such a way that the audience gets invited into the world filled with wonder and curiosity.

Chapter 5: CONCLUSION AND FUTURE WORK

Herein, we integrated the state-of-the-art AI technologies within the Blender software platform, particularly Generative Adversarial Networks (GANs). We aimed at empowering creators in their content creation process by enhancing Blender's capabilities, allowing the creation of lifelike and diverse content via intuitive interfaces and cutting-edge algorithms.

Using approaches such as Stable Diffusion and Latent Diffusion Models (LDM), we demonstrate the type of disruptive power this could have over the leverage of creativity in Blender. Creators will now be able to leverage AI-powered capabilities to produce quality images with ease, beyond the commonly experienced limitations of traditional resources, and open up new avenues in artistic development.

The Preferences Panel is embedded within Blender as the main development hub from which one can customize the workflow, choose an AI model, and integrate their choice of language prompts as well as the option for choosing the image genres. This is an intuitive interface for democratizing the whole process of image creation for every creator, at any level, while delivering loads of advanced functionality for those that need it.

Going forward, there are more explorations and many enhancements that must be brought in this sphere. To follow the latest AI breakthroughs and include new techniques within the Blender ecosystem, continuous research and development would be extremely important. Moreover, inclusion of a larger repertoire of AI models and improving their integration into Blender would enhance its creative prowess.

Equally important will be user feedback and the iterative design of the user experience, which keeps Blender intuitive and easily accessible. Collaboration and interaction with the Blender community will be of importance in future developments to address the emerging needs and challenges.

Furthermore, efforts are underway to optimize AI algorithms in order to continue improving the performance and efficiency of Blender, such that operation even with large datasets and complex models will remain smooth and responsive. This may involve the use of parallel processing techniques, model compression, and hardware acceleration to make the best utilization of computational resources.

This ultimately continues the work of democratizing AI-powered creativity and bringing inclusivity to the Blender community. Certainly, well-documented tutorials with in-depth support resources will empower a big user base to explore creative potentials towards the greater vibrancy of Blender-powered artistry. In sum, the integration of AI technologies into Blender is a major leap ahead in the march towards further development in digital art and creative expression. Thus, riding on the synergy between the human and the AI creative ingenuity, new realms of possibility will be unlocked, and a new era of artistic innovation will be born.

BIBLIOGRAPHY

[1] Ali, S., Ravi, P., Williams, R., DiPaola, D., & Breazeal, C. (2024). Constructing Dreams Using Generative AI. Proceedings of the AAAI Conference on Artificial Intelligence, 38(21), 23268-23275. https://doi.org/10.1609/aaai.v38i21.30374

[2] Yang, L., Zhang, Z., Song, Y., Hong, S., Xu, R., Zhao, Y., Zhang, W., Cui, B., & Yang, M. (2022). Diffusion Models: A Comprehensive Survey of Methods and Applications. ArXiv. /abs/2209.00796

[3] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2021). High-Resolution Image Synthesis with Latent Diffusion Models. ArXiv. /abs/2112.10752