

Image Generation using stable diffusion & Comfy UI

A Project Report

submitted in partial fulfillment of the requirements

of

AICTE Internship on AI: Transformative Learning

with

TechSaksham – A joint CSR initiative of Microsoft & SAP

by

MABBU KUSHAL SAI, kushalsaimabbu@gmail.com

Under the Guidance of

Jay Rathod

ACKNOWLEDGEMENT

We would like to take this opportunity to express our deep sense of gratitude to all individuals who helped us directly or indirectly during this thesis work.

Firstly, we would like to thank my supervisor, Jay Rathod, for being a great mentor and the best adviser I could ever have. His advice, encouragement and the critics are a source of innovative ideas, inspiration and causes behind the successful completion of this project. The confidence shown in me by him was the biggest source of inspiration for me. It has been a privilege working with him for the last one year. He always helped me during my project and many other aspects related to the program. His talks and lessons not only help in project work and other activities of the program but also make me a good and responsible professional .

.

M Kushal Sai

I, Mabbu Kushal Sai, sincerely express my gratitude for the opportunity to complete my internship successfully. This experience has been highly enriching, allowing me to gain practical knowledge and hands-on expertise in various domains.

During my internship, I had the opportunity to learn new concepts, enhance my technical skills, and explore real-world applications. In particular, I gained a brief yet insightful understanding of **Image Generation using Stable Diffusion and Comfy UI**, which has expanded my knowledge in AI-driven creative tools.

I would like to extend my heartfelt thanks to my mentors, trainers, and the entire team for their guidance and support throughout this journey. Their valuable insights and encouragement have greatly contributed to my professional growth.

This internship has been a rewarding experience, and I look forward to applying the skills I have acquired in future endeavours.

Mabbu Kushal Sai

ABSTRACT

The rapid advancement of artificial intelligence in image generation has led to the development of powerful models like **Stable Diffusion**, which can create high-quality, realistic images from textual descriptions.

This project, "**Image Generation using Stable Diffusion & Comfy UI**," explores the capabilities of this AI-driven approach and provides a user-friendly interface for seamless image synthesis. The primary objective is to understand and implement Stable Diffusion for generating images based on user input while utilizing **Comfy UI**, a modular and intuitive interface, to simplify the workflow.

This project addresses the need for an accessible and flexible tool that allows users to generate, modify, and fine-tune AI-generated images without requiring extensive technical expertise.

The methodology involves understanding the architecture of Stable Diffusion, implementing Comfy UI for a node-based workflow, experimenting with various prompts and model parameters, and evaluating the generated images based on clarity, accuracy, and artistic quality. Key findings demonstrate that **Stable Diffusion, combined with Comfy UI, offers a powerful yet user-friendly solution for AI-driven image generation**, enabling the creation of diverse and high-quality images with customizable styles and resolutions. The project highlights the efficiency of a node-based workflow, allowing users to experiment dynamically.

In conclusion, this work provides valuable insights into AI-based image generation and its applications. By leveraging Stable Diffusion and Comfy UI, users can create visually appealing content effortlessly. Future work can explore further optimizations, fine-tuning models, and integrating additional AI-driven enhancements for improved control and output precision.

TABLE OF CONTENT

Abstract	I
Chapter 1. Introduction	1
1.1 Problem Statement	1
1.2 Motivation	1
1.3 Objectives	2
1.4 Scope of the Project	2
Chapter 2. Literature Survey	3
Chapter 3. Proposed Methodology	
Chapter 4. Implementation and Results	
Chapter 5. Discussion and Conclusion	
References	

LIST OF FIGURES

Figure No.	Figure Caption	Page No.
Figure 1	BOTTEL WITH PURPLE COLOUR GALARY	5
Figure 2	A GIRL WITH BLACK ROSE	5
Figure 3	A BOY CROSS THE ROAD WITH HIS GRANDFATHER	5
Figure 4		
Figure 5		
Figure 6		
Figure 7		
Figure 8		
Figure 9		

LIST OF TABLES

[illegible]

CHAPTER 1

Introduction

1.1 Problem Statement:

Generating high-quality images from text is complex and requires artistic skills. Existing AI models like **Stable Diffusion** solve this but are difficult for non-technical users. This project integrates **Comfy UI**, a user-friendly interface, making AI-driven image creation accessible without coding expertise, enhancing creativity and efficiency across various fields.

1.2 Motivation:

AI-driven tools are transforming digital content creation, but many remain inaccessible to non-experts. **Stable Diffusion** offers immense potential, yet its complexity limits widespread adoption. By integrating **Comfy UI**, this project makes image generation easier for artists, designers, and hobbyists, enabling applications in game design, advertising, media production, and personalized content.

1.3 Objective:

This project aims to **simplify AI-powered image generation using Stable Diffusion and Comfy UI**. Key objectives include implementing a user-friendly workflow, enabling seamless text-to-image conversion, optimizing model parameters for enhanced image quality, and exploring practical applications across creative industries, ultimately making AI-driven design more accessible and efficient.

1.4 Scope of the Project:

The project focuses on **text-to-image generation using Stable Diffusion with Comfy UI** for an intuitive workflow. It explores different parameters, styles, and resolutions. Limitations include hardware requirements, model bias, and dependency on pre-trained data. Future scope includes advanced fine-tuning, real-time editing, and expanding AI-generated content applications across industries.

CHAPTER 2

Literature Survey

2.1 Review relevant literature or previous work in this domain.

AI-based image generation has evolved significantly, with models like GANs (Generative Adversarial Networks) and VAEs (Variational Autoencoders) leading early developments. More recently, Diffusion Models, such as Stable Diffusion, DALL·E, and Mid Journey, have demonstrated superior image quality and control. Research highlights the effectiveness of latent diffusion models in generating highly detailed and diverse images from textual descriptions.

2.2 Mentioning existing models, techniques, or methodologies related to the problem.

Key techniques in AI image generation include GANs, VAEs, and Diffusion Models. Stable Diffusion, based on Latent Diffusion Models (LDMs), is one of the most efficient models, allowing text-to-image synthesis with fine-tuned controls. Tools like Dream Studio, AUTOMATIC1111, and Comfy UI provide different interfaces, with Comfy UI offering a modular, node-based approach

2.3 Highlight the gaps or limitations in existing solutions and how your project will address them.

Despite advancements, many AI image-generation tools are complex, requiring coding knowledge or advanced configurations. Existing solutions lack user-friendly workflows for beginners. This project bridges the gap by integrating Comfy UI, simplifying model interaction, and making AI-driven image creation accessible to non-technical users while retaining flexibility for advanced users.

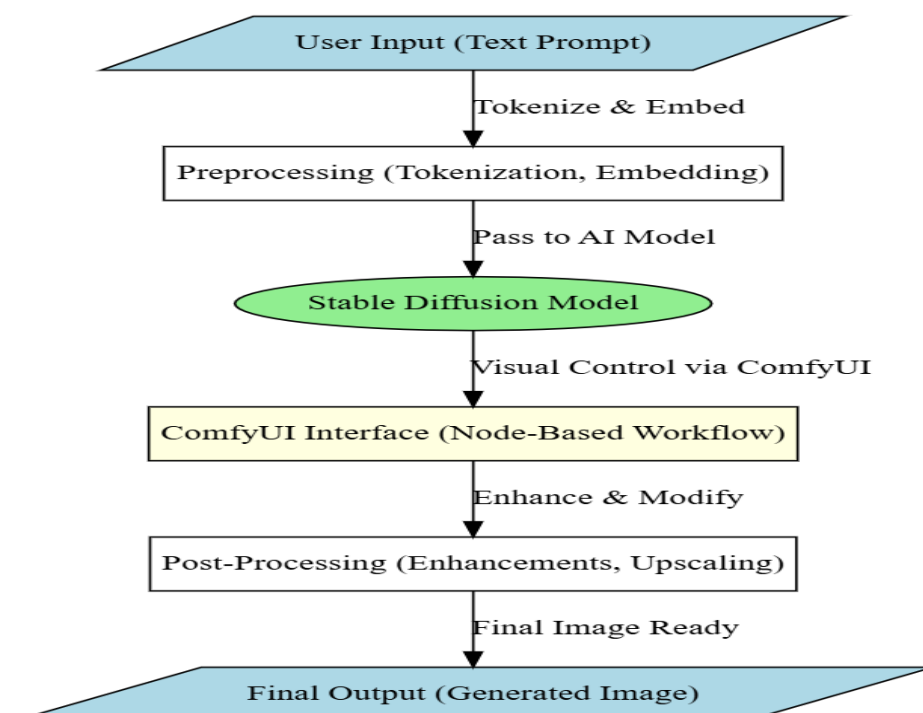
CHAPTER 3

Proposed Methodology

3.1 System Design

The proposed system follows a **text-to-image generation pipeline** using **Stable Diffusion** and **Comfy UI**. Below is the high-level architecture of the system:

1. **User Input:** The user provides a text prompt describing the desired image.
2. **Preprocessing:** The text input is tokenized and processed to be compatible with the Stable Diffusion model.
3. **Stable Diffusion Model:** The model generates an image based on the processed text input.
4. **Comfy UI Interface:** Provides a node-based visual workflow, allowing users to adjust parameters like image resolution, sampling method, and model selection.
5. **Post-processing:** Enhancements such as upscaling, filtering, or style adjustments are applied.
6. **Final Output:** The generated image is displayed, ready for download or further modifications.



3.2 Requirement Specification

Mention the tools and technologies required to implement the solution.

3.2.1 Hardware Requirements:

3.2.1.1 Processor: Minimum Intel i5/Ryzen 5, Recommended Intel i7/Ryzen 7 or higher

3.2.1.2 RAM: Minimum 8GB, Recommended 16GB+

3.2.1.3 GPU: NVIDIA RTX 3060 (Minimum), Recommended RTX 3080/4090 for faster processing

3.2.1.4 Storage: Minimum 50GB free space, SSD recommended for faster operations

3.2.2 Software Requirements:

3.2.2.1 Operating System: Windows 10/11, Linux (Ubuntu preferred), or macOS

3.2.2.2 Stable Diffusion Model: Pre-trained models like Stable Diffusion 1.5, 2.1, or SDXL

3.2.2.3 Comfy UI: Installed as the front-end workflow interface

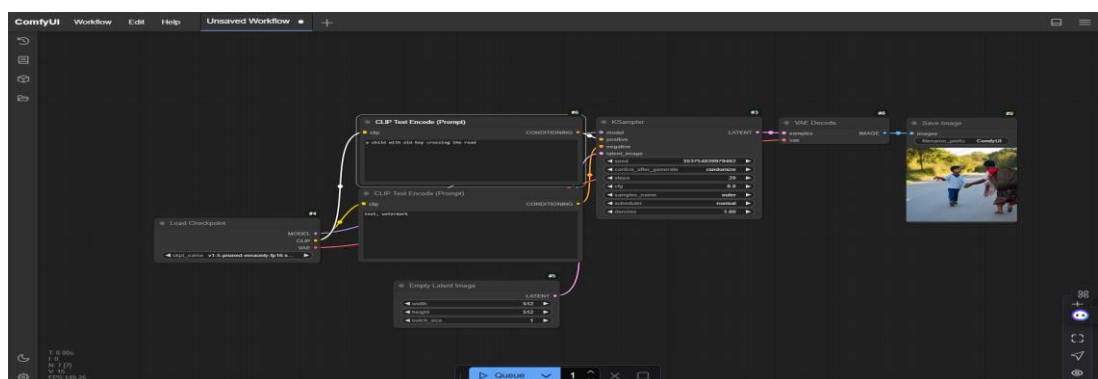
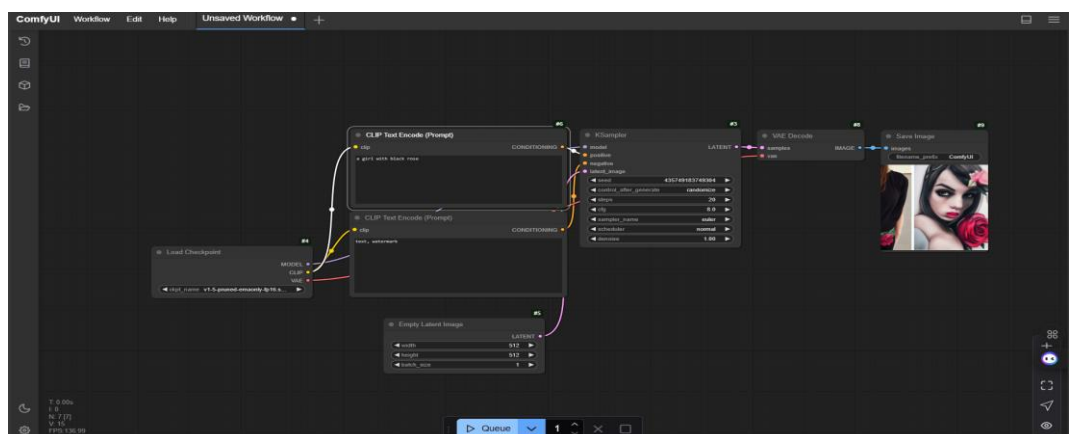
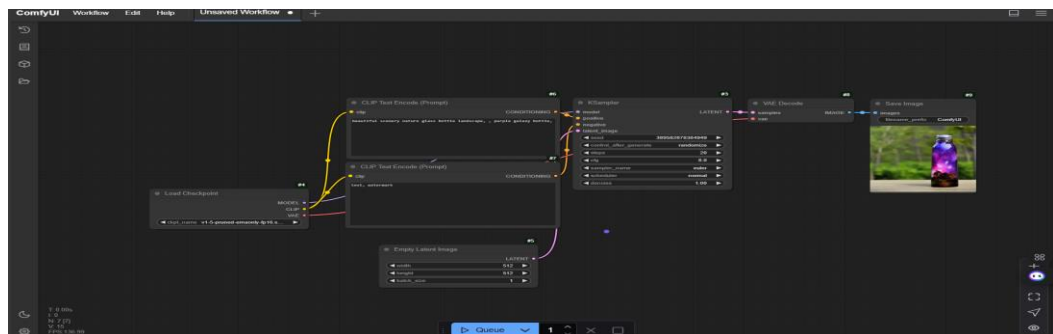
3.2.2.4 Python 3.10+: Required for executing AI models and dependencies

3.2.2.5 Additional Libraries: torch, diffusers, transformers, OpenCV, NumPy

CHAPTER 4

Implementation and Result

4.1 Snap Shots of Result:



4.2 GitHub Link for Code:

<https://github.com/Mabbukushalsai/AICTE-Image-Generation-using-stable-diffusion-Comfy-UI-/tree/main>

CHAPTER 5

Discussion and Conclusion

5.1 Future Work:

Future improvements can focus on **enhancing image quality, reducing generation time, and improving user control** over the output. Fine-tuning the **Stable Diffusion model** with custom datasets can generate more domain-specific images. **Real-time editing** features in Comfy UI, such as interactive brush-based modifications, can provide greater flexibility. Additionally, integrating **GAN-based upscaling models** can improve image resolution while maintaining details. Expanding support for **mobile and cloud-based implementations** can make AI-powered image generation more accessible.

5.2 Conclusion:

This project successfully demonstrates the **integration of Stable Diffusion with Comfy UI** to simplify AI-powered image generation. By providing a **user-friendly, node-based workflow**, it enhances accessibility for both beginners and advanced users. The project contributes to **making AI-generated images more customizable and efficient**, with applications across various creative industries. Future advancements can further refine usability, performance, and real-world adaptability.

REFERENCES

- [1]. Gal, Rinon, Adi Haviv, Yuval Alaluf, Amit H. Bermano, Daniel Cohen-Or, and Gal Chechik. 2024. “ComfyGen: Prompt-Adaptive Workflows for Text-to-Image Generation.” *arXiv [cs.CV]*. <https://doi.org/10.48550/ARXIV.2410.01731>.
- [2]. Sobania, Dominik, Martin Briesch, and Franz Rothlauf. 2024. “ComfyGI: Automatic Improvement of Image Generation Workflows.” *arXiv [cs.CV]*. <https://doi.org/10.48550/ARXIV.2411.14193>.
- [3]. Li, Tianxin. 2023. “Innovative Approach to Children’s Desk Lamp Using Stable Diffusion Technology.” In *Frontiers in Artificial Intelligence and Applications*. Frontiers in Artificial Intelligence and Applications. IOS Press.
- [4]. Rubio-Madrigal, Celia, Adarsh Jamadandi, and Rebekka Burkholz. 2025. “GNNs Getting ComFy: Community and Feature Similarity Guided Rewiring.” *arXiv [cs.LG]*. <https://doi.org/10.48550/ARXIV.2502.04891>.