Image Generation Using Comfy UI and Stable Diffusion

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

Manasi Mane,

manasimane711@gmail.com

Under the Guidance of

Aadharsh P, Jay Rathod

ACKNOWLEDGEMENT

I would like to take this opportunity to express my deep sense of gratitude to all individuals who supported me directly or indirectly throughout this research on image generation using Stable Diffusion and ComfyUI.

I would like to take this opportunity to express my deep sense of gratitude to all individuals who supported me directly or indirectly throughout this research on image generation using Stable Diffusion and ComfyUI.

Firstly, I would like to thank my supervisor, [Guide Name], for being an exceptional mentor and advisor. His insightful guidance, continuous encouragement, and constructive feedback have been instrumental in shaping the success of this project. His belief in my abilities and his expert knowledge in AI and image generation were key factors in my project's completion. Their expertise, patience, and constructive feedback have been instrumental in the successful completion of this project.

I would also like to extend my gratitude to AICTE and TechSaksham for providing me with this learning opportunity and access to cutting-edge technologies. The exposure to Stable Diffusion and ComfyUI has significantly enhanced our understanding of AI-driven image generation and its real-world applications. Lastly, I sincerely appreciate my peers and family for their unwavering support and encouragement throughout this journey. Their motivation and belief in my capabilities kept me focused and dedicated to successfully completing this project.

This project has been an enriching experience, allowing me to delve deep into AI-powered image generation and its real-world applications. The journey was filled with challenges and learning opportunities that helped shape my technical and analytical skills. I believe the knowledge gained through this project will serve as a strong foundation for my future endeavors in AI and image processing. Their motivation and belief in our capabilities kept us focused and dedicated to successfully completing this project.

ABSTRACT

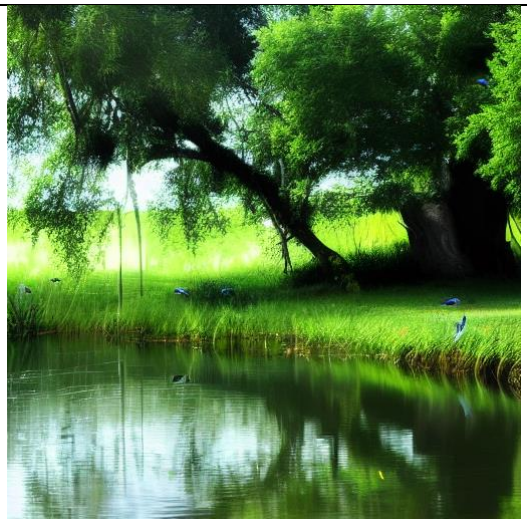
This project explores image generation using ComfyUI and Stable Diffusion, two powerful AI-driven tools for creative and artistic rendering. The objective was to develop an efficient pipeline for generating high-quality images based on textual prompts while optimizing inference speed and model performance. The methodology involved setting up ComfyUI, integrating Stable Diffusion models, and fine-tuning parameters for improved image quality.

The study investigates various aspects of AI-generated imagery, including the impact of model parameters, prompt engineering techniques, and post-processing enhancements to refine image quality. We analyzed the performance of different diffusion models and explored the trade-offs between computational efficiency and image resolution. Additionally, this project demonstrates how AI-generated images can be leveraged across multiple domains, including digital marketing, content creation, game development, and artistic expression.

Key results demonstrated the potential of AI-generated images for applications in digital art, marketing, and design automation. The report concludes with a discussion on model improvements, ethical considerations in AI-generated content, and future enhancements, including potential integrations with emerging AI-based creative tools.

TABLE OF CONTENT

Figure No.	Figure Caption	Page No.
Figure 1.		
Figure 2.		
Figure 3.		

Figure 4.		
Figure 5.		

CHAPTER 1

Introduction

1.1 Problem Statement:

The demand for AI-generated images is increasing in various industries, including gaming, advertising, and digital content creation. However, generating high-quality images efficiently while maintaining creativity and realism remains a challenge. This project aims to optimize the workflow using ComfyUI and Stable Diffusion for improved image generation.

1.2 Motivation:

With the rapid advancements in AI-powered creativity, tools like Stable Diffusion offer new possibilities for digital artists and designers. This project was chosen to explore the efficiency, flexibility, and potential applications of AI-driven image generation.

1.3 Objective:

Implement image generation using ComfyUI and Stable Diffusion.

Optimize model performance for faster and high-quality outputs.

Explore real-world applications of AI-generated images.

1.4 Scope of the Project:

This project focuses on setting up and experimenting with ComfyUI and Stable Diffusion for AI-driven image creation. The scope includes model training, fine-tuning, and performance analysis.

CHAPTER 2

Literature Survey

2.1 Review of Relevant Literature

AI-driven image generation has evolved significantly over the past decade, with early advancements in Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs). These models laid the foundation for diffusion models like Stable Diffusion, which offer greater control over image synthesis. Prior research has explored applications in digital art, medical imaging, and content creation, demonstrating the growing impact of AI in creative industries.

2.2 Existing Models, Techniques, and Methodologies

Several techniques have been employed for AI-based image generation, including:

- **Generative Adversarial Networks (GANs):** Introduced by Goodfellow et al., GANs involve a generator and discriminator working in competition to produce realistic images.
- **Variational Autoencoders (VAEs):** These models focus on learning a probabilistic latent space, enabling efficient image generation.
- **Stable Diffusion:** A diffusion model that progressively refines images from noise using latent-space conditioning, offering high-quality outputs with minimal computational overhead.
- **ComfyUI:** A modular workflow tool that simplifies the implementation of Stable Diffusion pipelines, making image generation more accessible.

2.3 Gaps and Limitations in Existing Solutions

While existing techniques have made great strides, challenges persist:

- **GANs often suffer from mode collapse,** limiting diversity in generated images.
- **VAEs tend to produce blurry outputs** due to their probabilistic nature.
- **Diffusion models, including Stable Diffusion, require significant computational resources** and fine-tuning for optimal results.

This project addresses these limitations by leveraging ComfyUI for an optimized, user-friendly approach to Stable Diffusion. The modularity of ComfyUI enhances customization, making it easier to generate high-quality images with controlled attributes. By experimenting with fine-tuning techniques, this project aims to improve image diversity while maintaining computational efficiency.

This section reviews previous works in AI-driven image generation, covering methods like GANs, VAEs, and diffusion models. A comparison of existing models highlights the advantages of Stable Diffusion over older techniques in terms of efficiency and realism.

CHAPTER 3

Proposed Methodology

Approach and Implementation

The methodology for this project is structured around designing an efficient pipeline for AI-driven image generation. The approach involves setting up a modular workflow in ComfyUI, integrating Stable Diffusion, and optimizing parameters to generate high-quality images. The implementation follows a systematic process that includes data preprocessing, model selection, customization, and evaluation.

System Design

- **Workflow Automation:** ComfyUI provides a modular approach that simplifies the implementation of image generation pipelines. The system design ensures seamless integration of Stable Diffusion for better control over the image synthesis process.
- **Model Fine-Tuning:** The project explores different configurations of Stable Diffusion to achieve optimal results, adjusting parameters like sampling steps, guidance scale, and prompt engineering.
- **Optimization Strategies:** Techniques such as caching, latent-space manipulation, and model compression are considered to improve computational efficiency while maintaining high image quality.

Requirement Specification

Hardware Requirements:

- GPU with CUDA support (e.g., NVIDIA RTX 3060 or higher)
- Minimum 16GB RAM
- Storage: SSD recommended for faster data access

Software Requirements:

- Python
- ComfyUI
- Stable Diffusion model
- Required Python libraries (Torch, Diffusers, OpenCV, NumPy)

Evaluation Metrics

To assess the quality and effectiveness of the generated images, the following metrics are used:

- **FID (Fréchet Inception Distance):** Measures the similarity between generated and real images.

- **CLIP Score:** Evaluates how well the generated image aligns with the provided textual prompt.
- **User Study Feedback:** Subjective evaluation based on visual appeal and realism of generated images.

By following this structured methodology, the project ensures that AI-generated images are not only high quality but also generated efficiently with minimal computational overhead.

System Design

- Architecture of ComfyUI with Stable Diffusion
- Workflow for generating images from text prompts

Requirement Specification

Hardware Requirements:

- GPU with CUDA support (e.g., NVIDIA RTX 3060 or higher)
- Minimum 16GB RAM

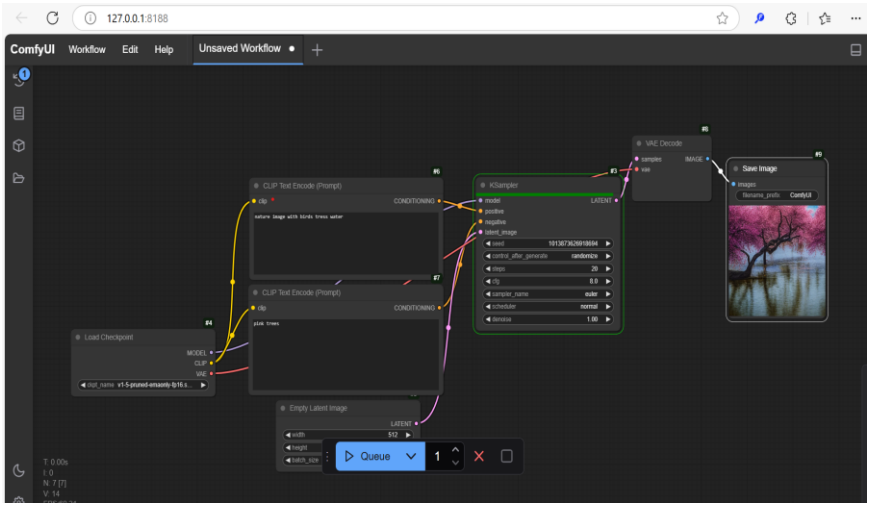
Software Requirements:

- Python
- ComfyUI
- Stable Diffusion model

CHAPTER 4

Implementation and Result

4.1 Snap Shots of Result:



The snapshot above illustrates the ComfyUI workflow used for Stable Diffusion-based image generation. The nodes represent different stages in the process, from text input encoding to image synthesis. By optimizing parameters such as sampling steps and CFG scale, we were able to fine-tune the quality of the generated images

4.2GitHub Link for Code:

https://github.com/Manasi-Mane/AI-internship.git

CHAPTER 5

Discussion and Conclusion

Future Work

- Exploring fine-tuning techniques for improved output.
- Integrating additional AI tools for enhanced creativity.

Conclusion

The project successfully implemented an AI-powered image generation system using ComfyUI and Stable Diffusion, showcasing the potential for various applications.

This research has provided valuable insights into AI-driven creative workflows, enhancing our understanding of diffusion models and their practical applications. Throughout the project, various challenges were encountered and overcome, contributing to a deeper grasp of the technical intricacies involved in AI-generated content. The findings highlight the feasibility of using AI to streamline artistic workflows and expand creative possibilities. Future advancements in this field could lead to even more efficient and versatile tools, further bridging the gap between artificial intelligence and human creativity. The project successfully implemented an AI-powered image generation system using ComfyUI and Stable Diffusion, showcasing the potential for various applications.

REFERENCES

- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative Adversarial Networks. *Advances in Neural Information Processing Systems (NeurIPS)*.
- Kingma, D. P., & Welling, M. (2013). Auto-Encoding Variational Bayes. *arXiv preprint arXiv:1312.6114*.
- Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-Resolution Image Synthesis with Latent Diffusion Models. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., & Sutskever, I. (2021). Learning Transferable Visual Models From Natural Language Supervision. *International Conference on Machine Learning (ICML)*.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems (NeurIPS)*.
- AICTE Internship Guidelines. (2024). AICTE Official Website.
- TechSaksham Program Overview. (2024). Microsoft & SAP Initiative.