**Text-to-Image Generation using Stable Diffusion**

This repository hosts a deep learning project that generates high-quality images from natural language prompts using the “Stable Diffusion” model. The system integrates “Hugging Face Diffusers”, “PyTorch”, and a “Gradio web interface” for real-time interaction and accessibility.

**Project Overview**

Traditional generative models like GANs face limitations in diversity, resolution, and semantic alignment. This project leverages “Stable Diffusion”, a latent diffusion model, to generate diverse and photorealistic images from textual descriptions.

The model is wrapped in a clean, user-friendly Gradio interface for ease of use.

**Features**

- “Stable Diffusion” for high-quality text-to-image synthesis

- Natural language prompt support using “CLIP” encoding

- GPU-accelerated with “PyTorch + CUDA”

- Interactive web-based UI with “Gradio”

- Generates varied artistic styles: photorealism, painting, surreal, etc.

- No user data is stored — all processing is done in-memory

**Tech Stack**

- Python

- PyTorch

- Hugging Face Diffusers

- Transformers

- Gradio

- CLIP

- VAE, U-Net

Example

|  |  |
| --- | --- |
| **Input Text** | **Prompted Image** |
| A ballerina dancing |  |
| “A dog playing with ball.” |  |