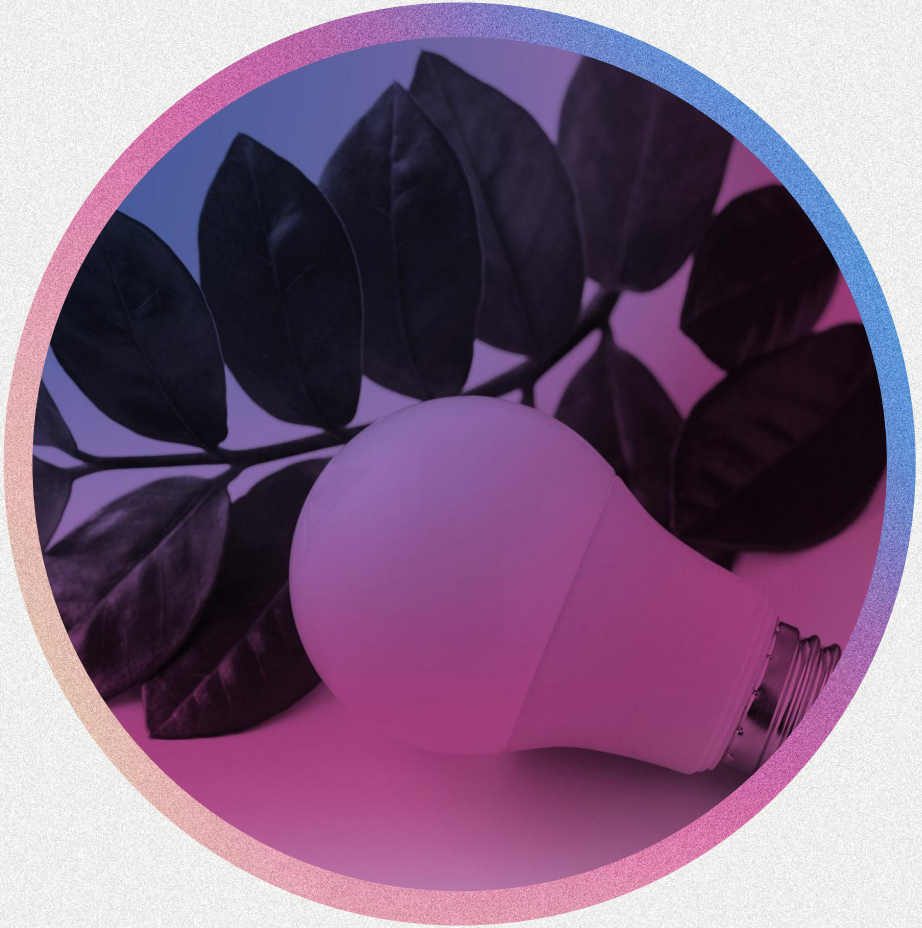
Synthetic Image Generation

Arpita Mangal, Suzana Amer, Siqi “Jessica” Chen, Daoxin “Alex” Wang

Table of Contents

Idea, Motivation & How it started



Idea & Motivation

1.Bridging Text and Image:

- Develop a generative model that translates textual descriptions/prompts into lifelike flower images.

2. User-Friendly Platform:

- Create a platform that enables users to express their thoughts in the form of flower images effortlessly.

3. Empower Verbally Impaired Individuals

- Enable and enhance the creativity, independence, and self-expression

Business Actions

01

Platform Development

- Prioritize ease of use, intuitive interfaces, and accessibility features to ensure that visually impaired individuals can easily navigate and utilize the system.

02

User Engagement

- Reach out to differently abled individuals and organizations
- Conduct awareness campaigns and organize workshops
- Collaborate with disability advocates and influencers

03

User Support

- Comprehensive user support and training resources
- Prompt and responsive customer support channels

How it started

Dataset Selection

Selecting a flower dataset that contains 6552 images and captions of different varieties of flowers.

Training GANs Model

Building a stacked GANs Model from scratched proved to be very computationally heavy

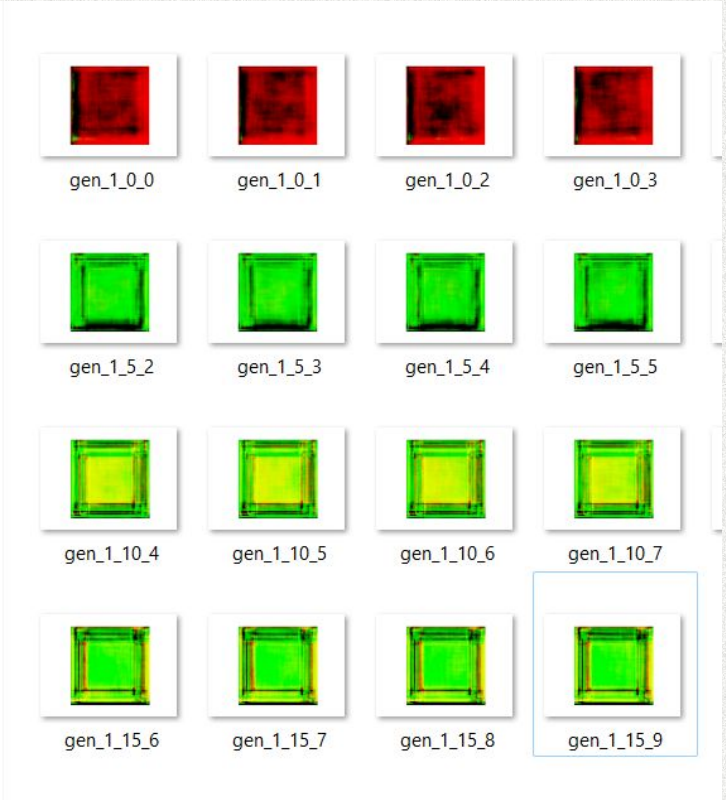
Fine-tuning Stable Diffusion Model

Applying a pre-trained text and image encoders to our diffusion model which provided us the tuned parameters.

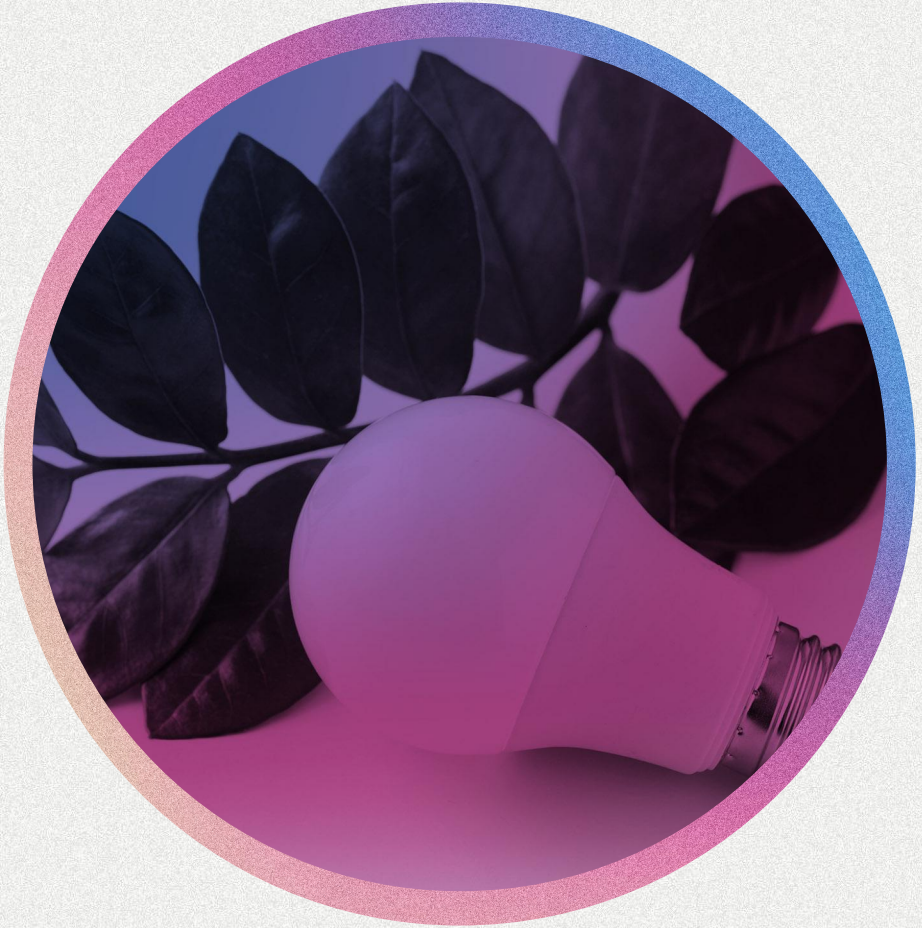
Loading Weights

Loading and saving the weights generated from our model to be used later for new flower image generation

Results of the Trained GANs Model

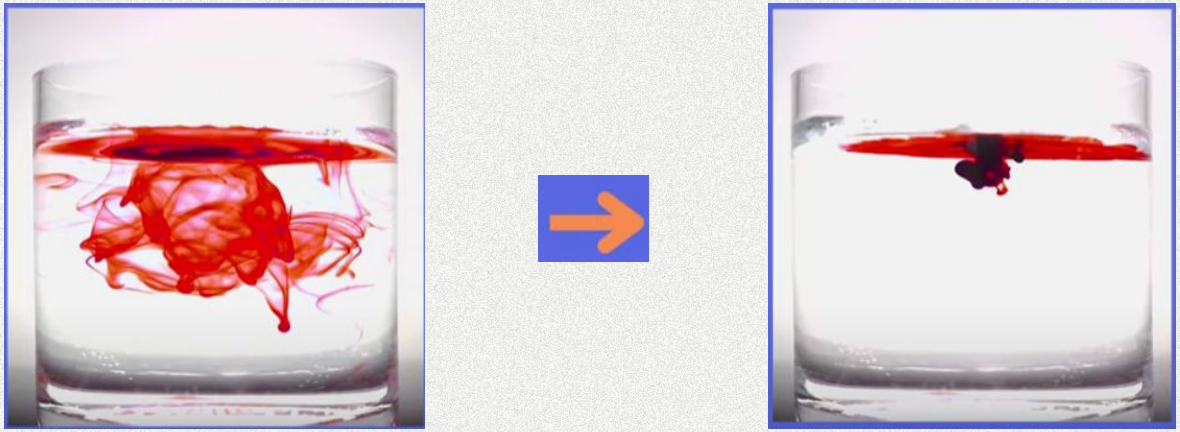


Methodology & Output



Literature Behind Diffusion Models

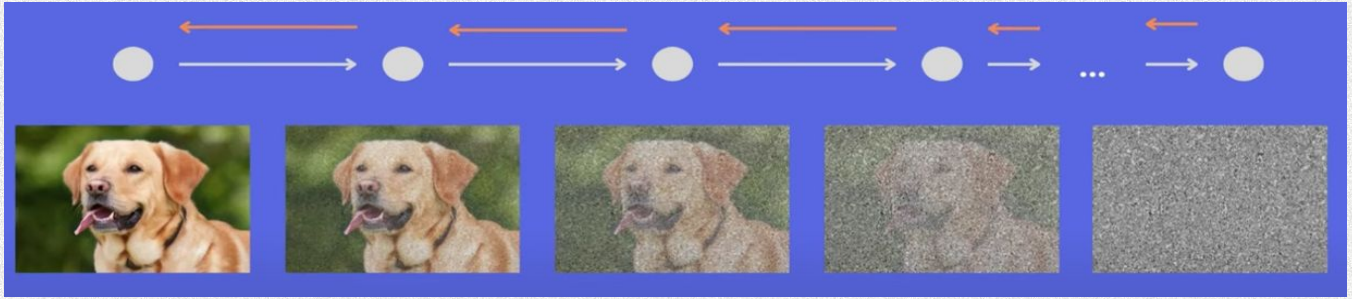
Diffusion models were inspired from non-equilibrium thermodynamics from physics.



- With diffusion models the goal is to learn a model that can reverse the diffusion process and bring the drop of paint to its original state.
- The drop being in one spot includes some information and as the diffusion process progresses, we lose this information.
- The one drop being the clear images we're interested in.

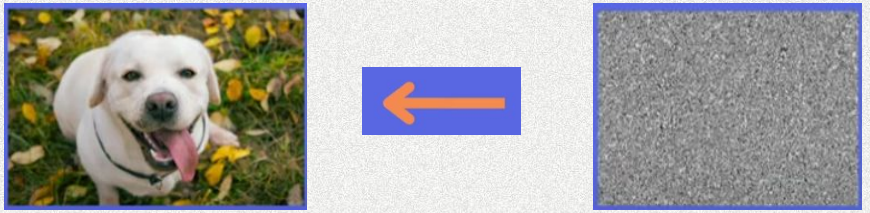
How do Diffusion models work?

Diffusion models work by adding noise to original images and later learning how to reverse this noise process through a markov chain



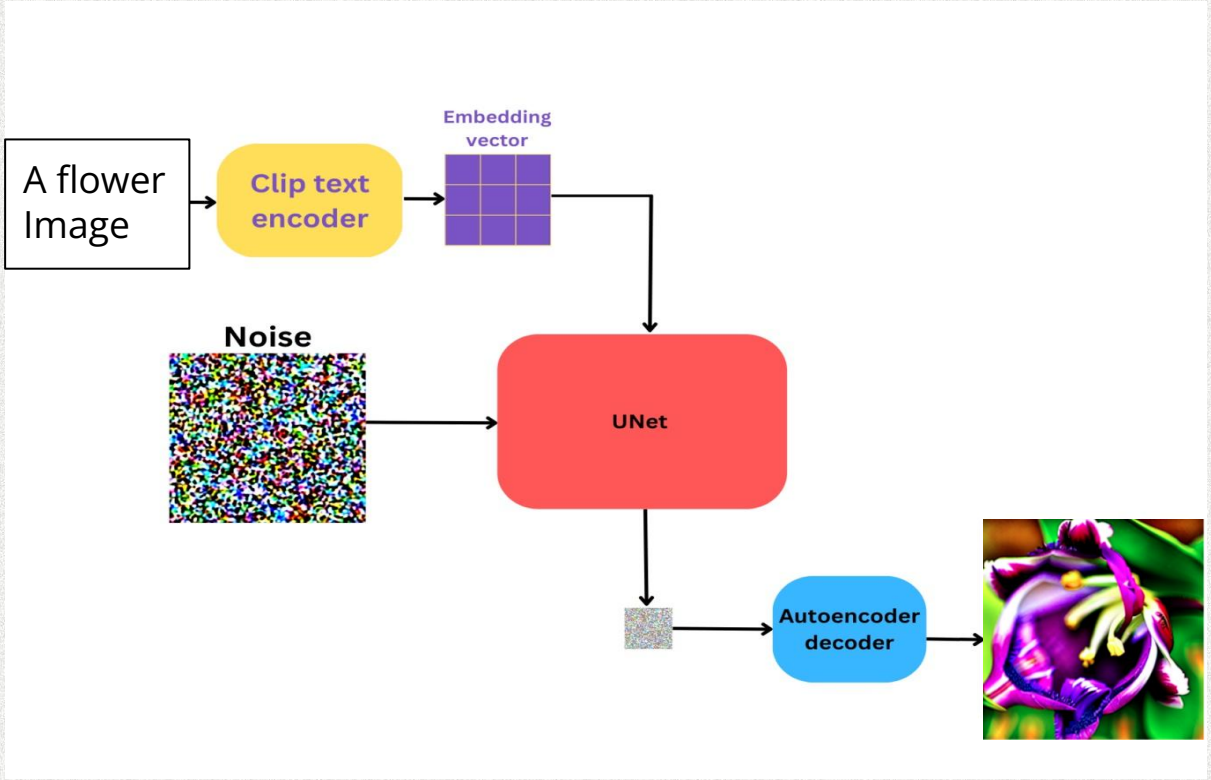
In each time step we are adding a little bit of noise to our image until the image only consist of noise and later learning how to reverse this noise adding process.

After the model is trained given only noise, the model is able to generate high resolution images.



What we did

1. Model


A flower Image
Clip text encoder
Embedding vector
Noise
UNet
Autoencoder decoder

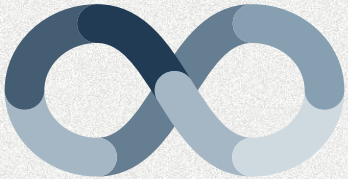
2. Data Preparation

- Selected a Pre-trained Diffusion Model:
- Pre-processed the data to match the input format expected by the diffusion model
- Total Images: 6550
- Textual Descriptions: ~9/image
- Tokenized the textual descriptions.
- Processed the images.
- Encoded the tokenized texts.
- Prepared the training dataset.

image (image)	text (string)
	"a close up of a passion flower"
	"a blue passion flower with a yellow flower in the middle of the flower"
	"a purple passion flower in a pot"
	"a purple passion flower with white flowers and green leaves"

3. Model Parameters

- Image Encoder, Text Encoder Form the original Pre-trained model
- Maximum length for a text prompt: 77
- Image Resolution: 256
- Generated Normal Random Noise
- Hyperparameters: Tutorial from “Hugging Face”
- Loss: 0.2285



4.Image Generation and Evaluation

- Loaded the weights of the fine-tuned diffusion model from the saved file.
- Generated flower images based on different prompts using the text_to_image function of the Stable Diffusion Model.
- Plotted and visualized the generated images along with their corresponding prompts.

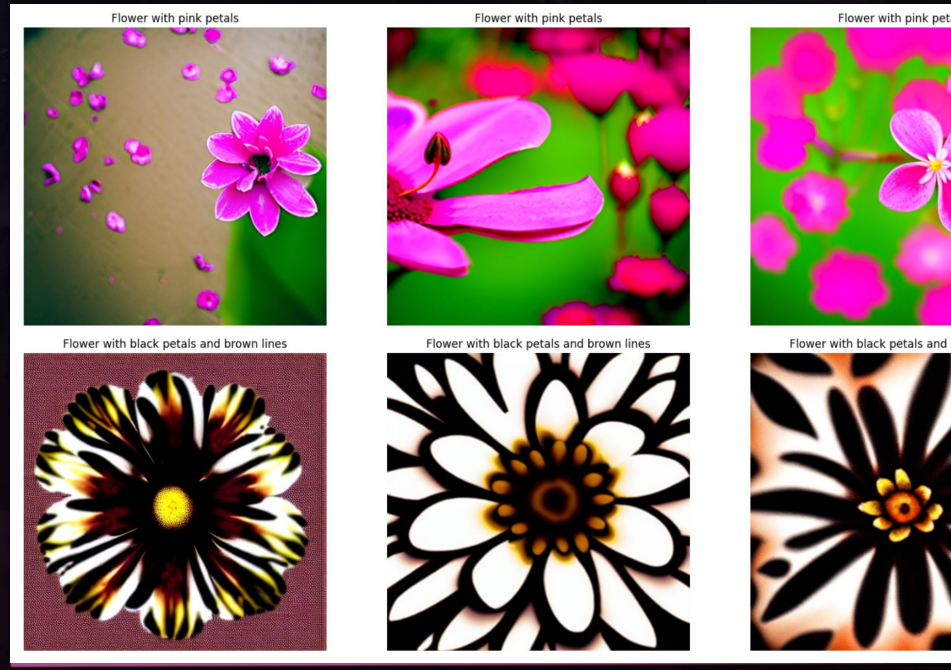
Model Output

Textual Prompts:

- ➔ *“Flowers with pink petals”*
- ➔ *“Flowers with black petals and brown lines”*

Image Output:

- ❑ Visually coherent and representative
- ❑ Strong alignment between the textual prompts and the generated flower images
- ❑ Produces a diverse range of flower images



Thank You!