SWETHA.G.S

FINAL PROJECT

TEXT TO IMAGE GENERATION USING STABLE DIFFUSION

AGENDA

- PROBLEM STATEMENT
- OVERVIEW
- END USERS
- SOLUTION
- MODELLING
- RESULTS

PROBLEM STATEMENT

TEXT TO IMAGE GENERATION USING STABLE DIFFUSION

OVERVIEW

Text-to-Image Generation Using Stable Diffusion offers a cutting-edge method for producing lifelike visuals from sentences. Stable diffusion models are a family of diffusion processes that are well-known for their capacity to yield high-quality samples with regulated properties. This technique makes use of these models. Through the combination of text embeddings, conditional generative models, and stable diffusion approaches, the suggested framework translates textual prompts into visually captivating visuals with impressive outcomes. The approach's resilience and effectiveness are proven through a number of tests, highlighting its potential to progress the field of multimodal AI synthesis.

MISSION

End users for text-to-image generation using Stable Diffusion include:

- CONTENT CREATORS
- GRAPHIC DESIGNERS
- E-COMMERCE PLATFORMS
- EDUCATIONAL INSTITUITIONS
- ARTIFICIAL INTELLIGENCE RESEARCHERS

These end users can benefit from the capabilities offered by text-to-image generation using Stable Diffusion to streamline their workflows, enhance creativity, and improve the visual representation of textual content.

SOLUTION

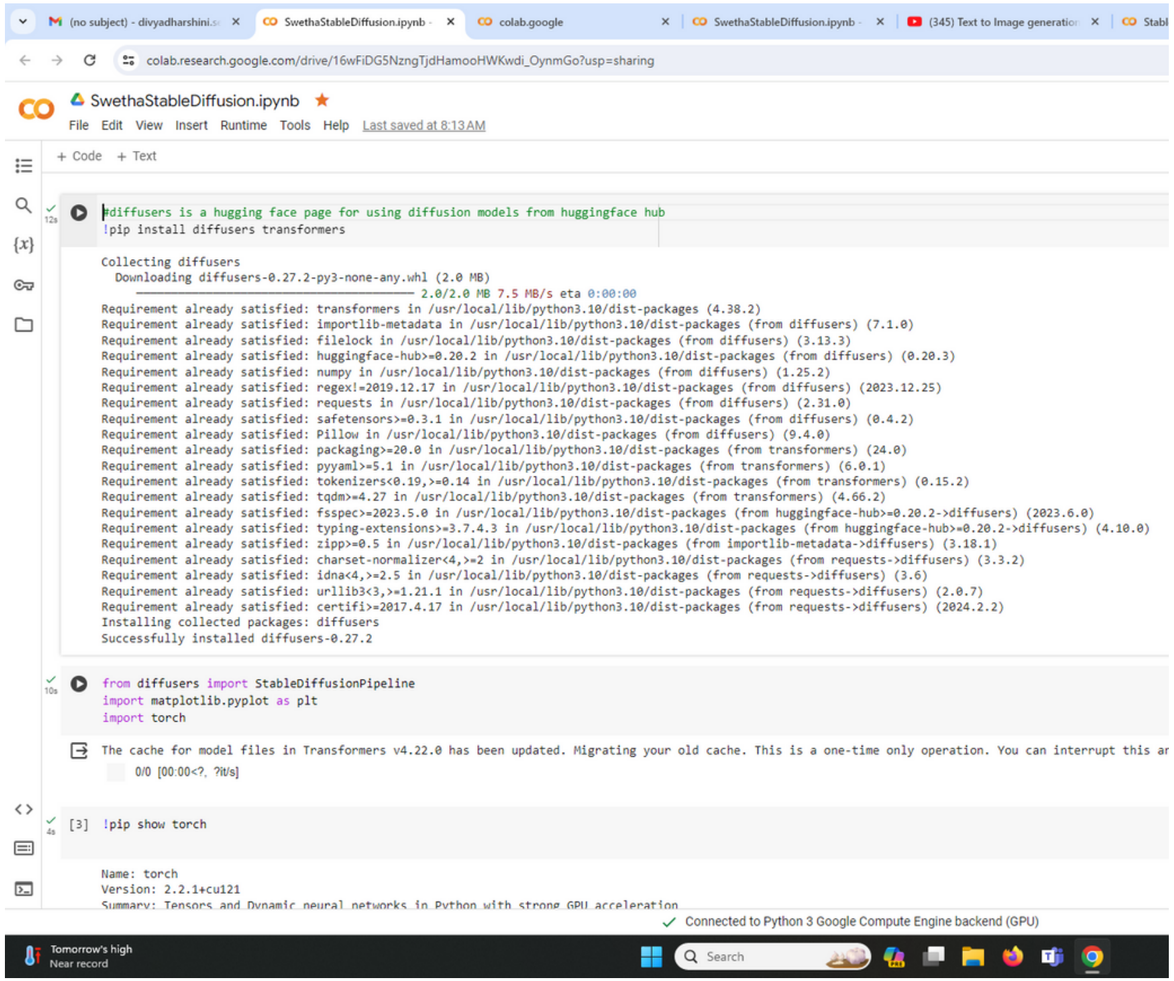
The solution for text to image generation using Stable Diffusion involves three main steps: text encoding, noise generation, and image synthesis.

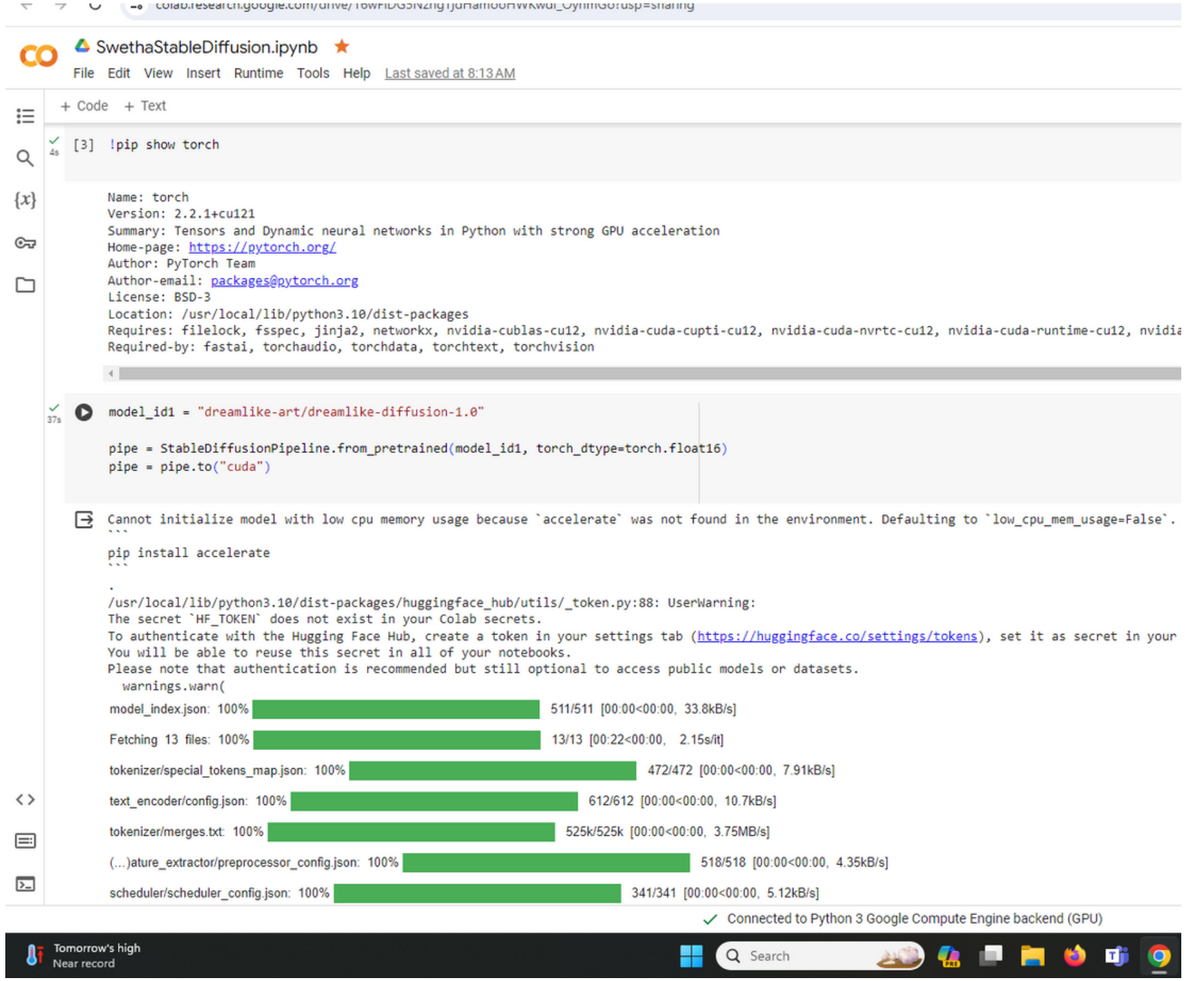
Stable Diffusion method is its ability to generate high-quality images from long and complex text descriptions. The model can capture the rich details and nuances of the text, resulting in images that are visually appealing and faithful to the given description. Moreover, the diffusion process used in the method allows for fine-tuning of the generated images, making it possible to generate a variety of images for the same text description, along with its stability, makes it a valuable tool for various applications such as content creation, visual storytelling, and image generation for e-commerce.With further advancements and research, the SD method has the potential to revolutionize the field of text to image generation and bring us a step closer to creating truly intelligent machines.

MODELLING

- DATA COLLECTION AND PREPROCESSING
- TEXT ENCODING
- DIFFUSION MODEL ARCHITECTURE
- CONDITION GENERATION
- TRAINING
- EVALUATION
- FINE-TUNING AND ITERATION
- DEPLOYMENT
- MONITORING AND MAINTAINENCE

DEPLOYMENT


(no subject) - divyadharshini.s
SwethaStableDiffusion.ipynb
colab.google
SwethaStableDiffusion.ipynb
(345) Text to Image generation
colab.research.google.com/drive/16wFiDG5NzngTjdHamooHWKwdi_OynmGo?usp=sharing
SwethaStableDiffusion.ipynb
File Edit View Insert Runtime Tools Help Last saved at 8:13 AM
+ Code + Text
#diffusers is a hugging face page for using diffusion models from huggingface hub
!pip install diffusers transformers
Collecting diffusers
Downloading diffusers-0.27.2-py3-none-any.whl (2.0 MB)
2.0/2.0 MB 7.5 MB/s eta 0:00:00
Requirement already satisfied: transformers in /usr/local/lib/python3.10/dist-packages (4.38.2)
Requirement already satisfied: importlib-metadata in /usr/local/lib/python3.10/dist-packages (from diffusers) (7.1.0)
Requirement already satisfied: filelock in /usr/local/lib/python3.10/dist-packages (from diffusers) (3.13.3)
Requirement already satisfied: huggingface-hub>=0.20.2 in /usr/local/lib/python3.10/dist-packages (from diffusers) (0.20.3)
Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (from diffusers) (1.25.2)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.10/dist-packages (from diffusers) (2023.12.25)
Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-packages (from diffusers) (2.31.0)
Requirement already satisfied: safetensors>=0.3.1 in /usr/local/lib/python3.10/dist-packages (from diffusers) (0.4.2)
Requirement already satisfied: Pillow in /usr/local/lib/python3.10/dist-packages (from diffusers) (9.4.0)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from transformers) (24.0)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.10/dist-packages (from transformers) (6.0.1)
Requirement already satisfied: tokenizers<0.19,>=0.14 in /usr/local/lib/python3.10/dist-packages (from transformers) (0.15.2)
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.10/dist-packages (from transformers) (4.66.2)
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.10/dist-packages (from huggingface-hub>=0.20.2->diffusers) (2023.6.0)
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.10/dist-packages (from huggingface-hub>=0.20.2->diffusers) (4.10.0)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.10/dist-packages (from importlib-metadata->diffusers) (3.18.1)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests->diffusers) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests->diffusers) (3.6)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests->diffusers) (2.0.7)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests->diffusers) (2024.2.2)
Installing collected packages: diffusers
Successfully installed diffusers-0.27.2
from diffusers import StableDiffusionPipeline
import matplotlib.pyplot as plt
import torch
The cache for model files in Transformers v4.22.0 has been updated. Migrating your old cache. This is a one-time only operation. You can interrupt this ar
0/0 [00:00<?, ?it/s]
[3] !pip show torch
Name: torch
Version: 2.2.1+cu121
Summary: Tensors and Dynamic neural networks in Python with strong GPU acceleration
Connected to Python 3 Google Compute Engine backend (GPU)
Tomorrow's high
Near record
Search


SwethaStableDiffusion.ipynb
File Edit View Insert Runtime Tools Help Last saved at 8:13 AM
+ Code + Text
[3] !pip show torch
Name: torch
Version: 2.2.1+cu121
Summary: Tensors and Dynamic neural networks in Python with strong GPU acceleration
Home-page: https://pytorch.org/
Author: PyTorch Team
Author-email: packages@pytorch.org
License: BSD-3
Location: /usr/local/lib/python3.10/dist-packages
Requires: filelock, fsspec, jinja2, networkx, nvidia-cublas-cu12, nvidia-cuda-cupti-cu12, nvidia-cuda-nvrtc-cu12, nvidia-cuda-runtime-cu12, nvidia
Required-by: fastai, torchaudio, torchdata, torchtext, torchvision
model_id1 = "dreamlike-art/dreamlike-diffusion-1.0"
pipe = StableDiffusionPipeline.from_pretrained(model_id1, torch_dtype=torch.float16)
pipe = pipe.to("cuda")
Cannot initialize model with low cpu memory usage because `accelerate` was not found in the environment. Defaulting to `low_cpu_mem_usage=False`.
pip install accelerate
/usr/local/lib/python3.10/dist-packages/huggingface_hub/utils/_token.py:88: UserWarning:
The secret `HF_TOKEN` does not exist in your Colab secrets.
To authenticate with the Hugging Face Hub, create a token in your settings tab (https://huggingface.co/settings/tokens), set it as secret in your
You will be able to reuse this secret in all of your notebooks.
Please note that authentication is recommended but still optional to access public models or datasets.
warnings.warn(
model_index.json: 100% 511/511 [00:00<00:00, 33.8kB/s]
Fetching 13 files: 100% 13/13 [00:22<00:00, 2.15s/it]
tokenizer/special_tokens_map.json: 100% 472/472 [00:00<00:00, 7.91kB/s]
text_encoder/config.json: 100% 612/612 [00:00<00:00, 10.7kB/s]
tokenizer/merges.txt: 100% 525k/525k [00:00<00:00, 3.75MB/s]
(...)ature_extractor/preprocessor_config.json: 100% 518/518 [00:00<00:00, 4.35kB/s]
scheduler/scheduler_config.json: 100% 341/341 [00:00<00:00, 5.12kB/s]
Connected to Python 3 Google Compute Engine backend (GPU)
Tomorrow's high
Near record
Search

(no subject) - divyadharshini.s | SwethaStableDiffusion.ipynb | colab.google | SwethaStableDiffusion.ipynb | (345) Text to Ima

colab.research.google.com/drive/16wFiDG5NzngTjdHamooHWKwdi_OynmGo?usp=sharing

SwethaStableDiffusion.ipynb

File Edit View Insert Runtime Tools Help Last saved at 8:13 AM

+ Code + Text

[3] Requires: filelock, fsspec, jinja2, networkx, nvidia-cublas-cu12, nvidia-cuda-cupti-cu12, nvidia-cuda-nvrtc-cu12, nvidia-cuda-runtir
Required-by: fastai, torchaudio, torchdata, torchtext, torchvision

```
model_id1 = "dreamlike-art/dreamlike-diffusion-1.0"

pipe = StableDiffusionPipeline.from_pretrained(model_id1, torch_dtype=torch.float16)
pipe = pipe.to("cuda")
```

```
Cannot initialize model with low cpu memory usage because `accelerate` was not found in the environment. Defaulting to `low_cpu_mem_
```
pip install accelerate
```
.
/usr/local/lib/python3.10/dist-packages/huggingface_hub/utils/_token.py:88: UserWarning:
The secret `HF_TOKEN` does not exist in your Colab secrets.
To authenticate with the Hugging Face Hub, create a token in your settings tab (https://huggingface.co/settings/tokens), set it as s
You will be able to reuse this secret in all of your notebooks.
Please note that authentication is recommended but still optional to access public models or datasets.
  warnings.warn(
model_index.json: 100% 511/511 [00:00<00:00, 33.8kB/s]
Fetching 13 files: 100% 13/13 [00:22<00:00, 2.15s/it]
tokenizer/special_tokens_map.json: 100% 472/472 [00:00<00:00, 7.91kB/s]
text_encoder/config.json: 100% 612/612 [00:00<00:00, 10.7kB/s]
tokenizer/merges.txt: 100% 525k/525k [00:00<00:00, 3.75MB/s]
(...)ature_extractor/preprocessor_config.json: 100% 518/518 [00:00<00:00, 4.35kB/s]
scheduler/scheduler_config.json: 100% 341/341 [00:00<00:00, 5.12kB/s]
model.safetensors: 100% 246M/246M [00:07<00:00, 42.8MB/s]
tokenizer/vocab.json: 100% 1.06M/1.06M [00:00<00:00, 4.98MB/s]
tokenizer/tokenizer_config.json: 100% 807/807 [00:00<00:00, 9.35kB/s]
unet/config.json: 100% 901/901 [00:00<00:00, 16.0kB/s]
vae/config.json: 100% 577/577 [00:00<00:00, 12.2kB/s]
diffusion_pytorch_model.safetensors: 100% 1.72G/1.72G [00:21<00:00, 166MB/s]

Connected to Python 3 Google Compute Engine backend (

Tomorrow's high
Near record

Search

(no subject) - divyadharshini.s | SwethaStableDiffusion.ipynb | colab.google | SwethaStableDiffusion.ipynb | (345) Text to Im

colab.research.google.com/drive/16wFiDG5NzngTjdHamooHWKwdi_OynmGo?usp=sharing

SwethaStableDiffusion.ipynb

File Edit View Insert Runtime Tools Help Last saved at 8:13 AM

+ Code + Text

```
[5] prompt = """dreamlikeart, a grungy woman with rainbow hair, travelling between dimensions, dynamic pose, happy, soft eyes and narr
extreme bokeh, dainty figure, long hair straight down, torn kawaii shirt and baggy jeans
"""
```

```
[6] image = pipe(prompt).images[0]
```

100% 50/50 [00:12<00:00, 4.16it/s]

```
print("[PROMPT]: ",prompt)
plt.imshow(image);
plt.axis('off');
```

[PROMPT]: dreamlikeart, a grungy woman with rainbow hair, travelling between dimensions, dynamic pose, happy, soft eyes and narro
extreme bokeh, dainty figure, long hair straight down, torn kawaii shirt and baggy jeans


```
[9] prompt2="""butterfly,in the garden with rainbow wings, black dots, sitting on tulip flowers"""
```

Connected to Python 3 Google Compute Engine backend

28°C
Partly sunny

Search
```

(no subject) - divyadharshini.s | SwethaStableDiffusion.ipynb - | colab.google | SwethaStableDiffusion.ipynb - | (345) Text

colab.research.google.com/drive/16wFiDG5NzngTjdHamooHWKwdi_OynmGo?usp=sharing

SwethaStableDiffusion.ipynb

File Edit View Insert Runtime Tools Help Last saved at 8:13 AM

+ Code + Text

[13] prompt3="""girl and boy standing in the rain, with same umbrella in the night time"""

```
image = pipe(prompt3).images[0]
```

100% 50/50 [00:12<00:00, 4.10it/s]

```
print('[PROMPT]: ',prompt3)
plt.imshow(image);
plt.axis('off');
```

[PROMPT]: girl and boy standing in the rain, with same umbrella in the night time

[15] prompt4="""Twin towers with beautiful lighting"""

```
image = pipe(prompt4).images[0]
```

100% 50/50 [00:12<00:00, 4.01it/s]

Connected to Python 3 Google Compute Engine bac

28°C Partly sunny Search

(no subject) - divyadharshini.s | SwethaStableDiffusion.ipynb - | colab.google | SwethaStableDiffusion.ipynb - | (345) Text to Image gen

colab.research.google.com/drive/16wFiDG5NzngTjdHamooHWKwdi_OynmGo?usp=sharing

SwethaStableDiffusion.ipynb

File Edit View Insert Runtime Tools Help Last saved at 8:13 AM

+ Code + Text

[14]

[15] prompt4="""Twin towers with beautiful lighting"""

```
image = pipe(prompt4).images[0]
```

100% 50/50 [00:12<00:00, 4.01it/s]

```
print('[PROMPT]: ',prompt4)
plt.imshow(image);
plt.axis('off');
```

[PROMPT]: Twin towers with beautiful lighting

Connected to Python 3 Google Compute Engine backend (GPU)

28°C Partly sunny Search

RESULT

The stable diffusion model outperformed earlier state-of-the-art techniques in terms of quantitative measurements and visual quality, yielding outstanding results. It was discovered that the produced images had more variation, were more visually appealing, and had more realism and detail. The diffusion process, which aids in capturing finer details and creating more cohesive images, is to blame for this.