GOVERNMENT COLLEGE OF ENGINEERING SALEM-11

NAAN MUDHALVAN ASSESSMENT

CREATING IMAGES FROM TEXT DESCRIPTION

NAME : SUBIKSHA M

REGISTER NO: 61772241049

BRANCH: ELECTRICAL AND ELECTRONICS ENGINEERING

TEAM MEMBERS:

1) SRI HARINI S - 61772241047
2) MAITHILI M P - 61772241019
3) MEGAVARSHINI K - 61772241024
4) MEGAVARTHINI G - 61772241025

CREATING IMAGES FROM TEXT DESCRIPTION

1.ABSTRACT:

This project explores the development of an AI-based system capable of generating high-quality images from textual descriptions. Utilizing advancements in natural language processing (NLP) and computer vision, the system will interpret descriptive prompts and create corresponding visual outputs. This type of system has applications in various fields, including digital art, marketing, gaming, and content creation. By combining neural networks, deep learning models, and image generation techniques, the project aims to provide a platform that can automate creative processes through AI-generated imagery.

Generative AI has made significant strides in recent years, especially in text-to-image generation, where AI models can create visuals from user input descriptions. This project aims to build a system that integrates NLP and image generation algorithms to produce high-quality images based on natural language prompts. This application of AI can simplify tasks such as concept art creation, visual content production, and design ideation. Using neural networks trained on massive datasets of text and images, the system will generate accurate and detailed images that reflect the descriptions provided.

The core theme of the project is to leverage generative AI technologies, specifically focusing on translating text into visuals. The goal is to develop a user-friendly interface that allows users to input text prompts, which the AI will process to generate images that match the textual description as closely as possible. The system will incorporate techniques from both NLP and generative adversarial networks (GANs) or diffusion models to ensure image quality and semantic alignment with the prompt.

In conclusion, Building an AI system capable of generating images from text descriptions holds great potential in revolutionizing creative industries. The system could be used in various sectors, including advertising, entertainment, and even education. By combining state-of-the-art models and tools, this project will demonstrate how artificial intelligence can contribute to creative processes, making it easier for individuals and businesses to produce visual content.

2.SYSTEM REQUIREMENTS:

2.1 HARDWARE REQUIREMENTS:

1) High performance computing hardware (e.g., Multi-core CPU , GPU)
2) RAM: At least 4 GB of RAM

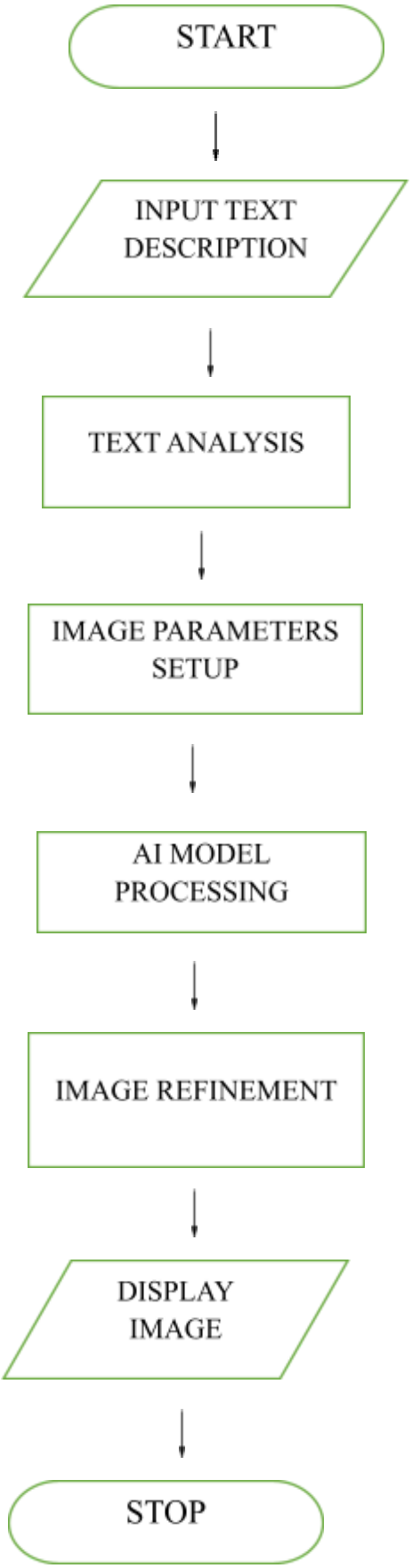
2.2 SOFTWARE REQUIREMENTS:

1) Operating System: Linux (Ubuntu 20.04 or later) or Windows 10
2) Programming Language: Python 3.8 or higher
3) Deep Learning Framework: TensorFlow 2.x or PyTorch 1.10+
4) Text Preprocessing Libraries: NLTK, SpaCy, or Hugging Face's Transformers
5) Image Generation Model: DALL-E, Stable Diffusion, or a custom GAN model
6) Database: MongoDB, SQLite, or PostgreSQL for managing datasets
7) IDE: VS Code, PyCharm, or Jupyter Notebook for code development

3.TOOLS AND VERSIONS

1) CUDA: Version 11.2
2) cuDNN: Version 8.1+
3) Python: Version 3.8+
4) TensorFlow: Version 2.6+
5) Hugging Face Transformers: Version 4.11+

FLOW CHART:


START
INPUT TEXT DESCRIPTION
TEXT ANALYSIS
IMAGE PARAMETERS SETUP
AI MODEL PROCESSING
IMAGE REFINEMENT
DISPLAY IMAGE
STOP

CODE IMPLEMENTATION (SAMPLE CODE):

```
import torch
from transformers import pipeline
from diffusers import StableDiffusionPipeline
from PIL import Image

def input_text():
    text_description = input("Enter the text description: ")
    return text_description

def analyze_text(text):
    print(f"Analyzing text: {text}")
    return text
    image_size = (512, 512)  # 512x512 pixels
    return image_size

                                                    model                =
StableDiffusionPipeline.from_pretrained("runwayml/stable-diffusion-v1-5")
    model.to("cuda" if torch.cuda.is_available() else "cpu")

    print("Generating image...")
    image = model(text, guidance_scale=7.5).images[0]
     image = image.resize(image_size)
    return image

def refine_image(image):
    print("Refining image...")
```

```
    return image
def display_image(image):
    image.show()


def main():
    text_description = input_text()
    analyzed_text = analyze_text(text_description)
    image_size = setup_image_params()
    generated_image = generate_image_from_text(analyzed_text, image_size)
    refined_image = refine_image(generated_image)
    display_image(refined_image)


    if __name__ == "__main__":
    main()
```

PROJECT HURDLES:

1) **Ambiguity in Language**: Natural language can be vague or ambiguous, making it difficult to interpret what the user intends. For example, descriptors like "tall" or "colorful" can vary widely in meaning.

2) **Complexity of Visual Elements**: Describing complex scenes with multiple elements (people, objects, backgrounds) can lead to challenges in accurately capturing relationships, proportions, and perspectives.

3) **Style and Artistic Interpretation**: Users may have specific artistic styles in mind, which can be challenging to replicate without a clear description. Different interpretations of style can lead to varying results.

4) **Technical Limitations**: Image generation models may have limitations in resolution, color accuracy, and the ability to represent certain concepts or styles.

5) **Computational Resources**: Generating high-quality images can be resource-intensive, requiring significant computational power and memory.

OUTPUT:

Image Description:

"A sunset over the mountains with a river flowing through a forest"