SOMAIYA VIDYAVIHAR UNIVERSITY

K. J. Somaiya School of Engineering

Department of Computer Engineering

Batch: A4	**Roll No.: 16010122147**
Experiment No: 06	
Group No: 03	

Title: Chapter 06 Result and discussion

Expected Outcome of Experiment:

CO4: Design of test cases to meet the desired specifications.

Books/ Journals/ Websites referred:

[Students can mention websites/ books used in their project implementation]

This write-up will expect students to prepare chapter no 6 in the format given below

Chapter 6

Result and discussion

This chapter presents the results obtained from the implementation of the prototype/application and provides an in-depth analysis of its performance. The findings are evaluated based on predefined metrics, including accuracy, efficiency, usability, and reliability. A comparative analysis with existing systems or methodologies is conducted to assess improvements and innovations introduced by the proposed approach. Additionally, key observations, challenges encountered during testing, and their implications are discussed. The insights gained from the results are critically examined to validate the effectiveness of the system and highlight areas for further improvement.

Implementation Summary:

- The code (`2-1.py`) uses the `"stabilityai/stable-diffusion-2-1"` model.
- The pipeline is optimized with attention slicing and `DPMSolverMultistepScheduler`.
- Four different prompts were tested:
  1. A programmer touching grass
  2. A dreamlike landscape with floating islands
  3. A Roman soldier in front of the Colosseum
  4. A cyberpunk character in a neon-lit alley
- Each output image is generated with:

1. `512x512` resolution
2. `50 inference steps`
3. `guidance scale = 3.5`

Output Images and Observations:

Prompt	Observation
A programmer touching grass	Realistic outdoor environment with human-like character. Slight ambiguity in facial details.
Dreamlike floating islands	Visually rich with abstract patterns, vibrant colors, and atmospheric lighting.
Roman soldier at sunset	Strong historical resemblance and correct architectural context (Colosseum).
Cyberpunk character	High detail in background textures and neon reflections. Stylized aesthetics achieved.

Performance Metrics:

Metric	Result
Average Inference Time	~6–7 seconds per image (with CUDA)
Image Resolution	512x512
Prompt Adherence	High (especially for descriptive prompts)
Usability	CLI-based, easily configurable
Resource Usage	Moderate GPU (8–12 GB VRAM recommended)

Comparative Analysis:

Compared to traditional GAN-based models or older versions of diffusion models, Stable Diffusion 2.1 offers:

- **Higher prompt accuracy** due to better text encoders
- **More detailed and photorealistic images**
- **Faster inference with reduced memory usage** using latent space diffusion

While **Stable Diffusion 3.5** could offer even better results, it was not used in this experiment due to its intensive hardware requirements, as noted in the README.

Challenges and Resolutions:

Challenge	Resolution
High VRAM usage	Used attention slicing and optimized scheduler
Access issues for model	Hugging Face CLI and access token setup used
Some prompts had vague results	Refined prompts for better semantic input

Conclusion:

The **NeuraPix** system built with Stable Diffusion 2.1 successfully generates diverse and high-quality images from text. It demonstrates the potential of diffusion-based models in creative applications like digital art, marketing content, and educational visualization. Future enhancements include integrating UI and experimenting with ControlNets or Stable Diffusion 3.5 for advanced control.