STABLE DIFFUSION

SOHILA AYMAN MOHAMED LASHIEN

41210053

—

DEEP GENERATIVE AI

—

ENG/OMAR EL-NAHAS

Prompt Exploration:-

-Prompt one First attempt :-

"a shallow muddy stream with red-brown textured ground, winding through a foggy forest of tall pine trees, soft diffused morning light, cinematic natural environment, ultra-realistic, photorealistic landscape, wet soil with visible flow marks"

-Prompt one Second attempt :-

"A narrow, reddish-brown stream winds through a misty, serene forest of tall, pale pine trees. Soft, diffused light from above casts an even glow across the mossy forest floor, enhancing the peaceful and ethereal atmosphere. The scene is viewed from a slightly elevated perspective, with the stream as the focal point framed by slender trees and muted vegetation."

-Prompt one third attempt :-

"a realistic muddy stream flowing through a dense foggy pine forest, soft morning light, tall trees with mist in the background, cinematic atmosphere, highly detailed, photorealistic, natural colors"
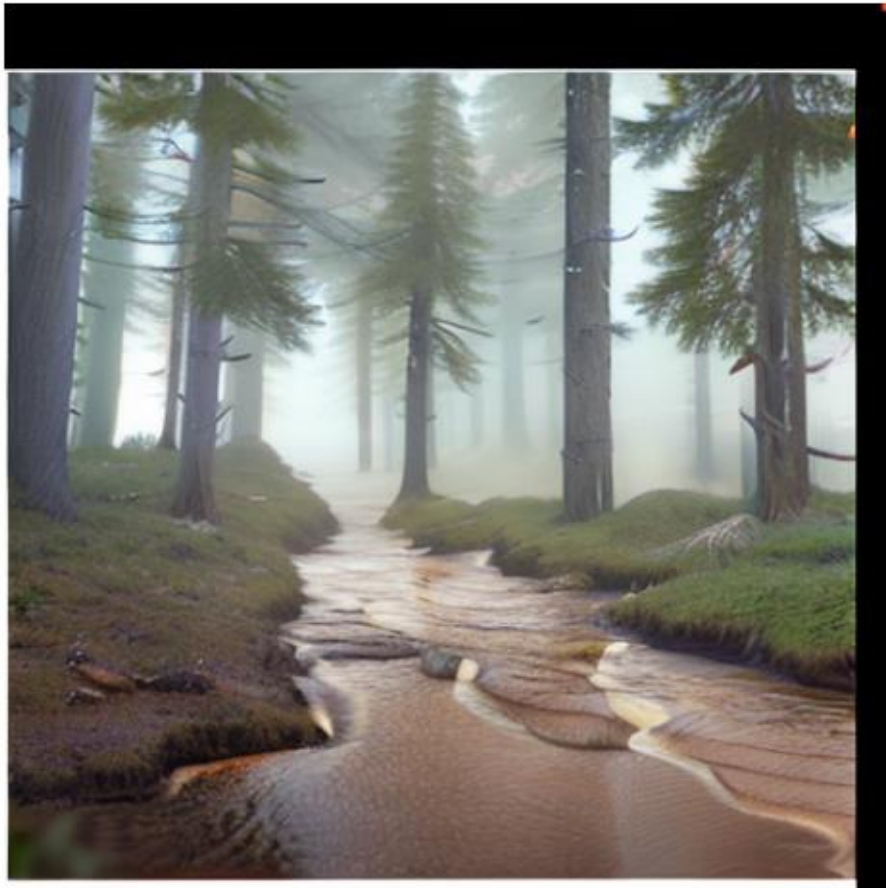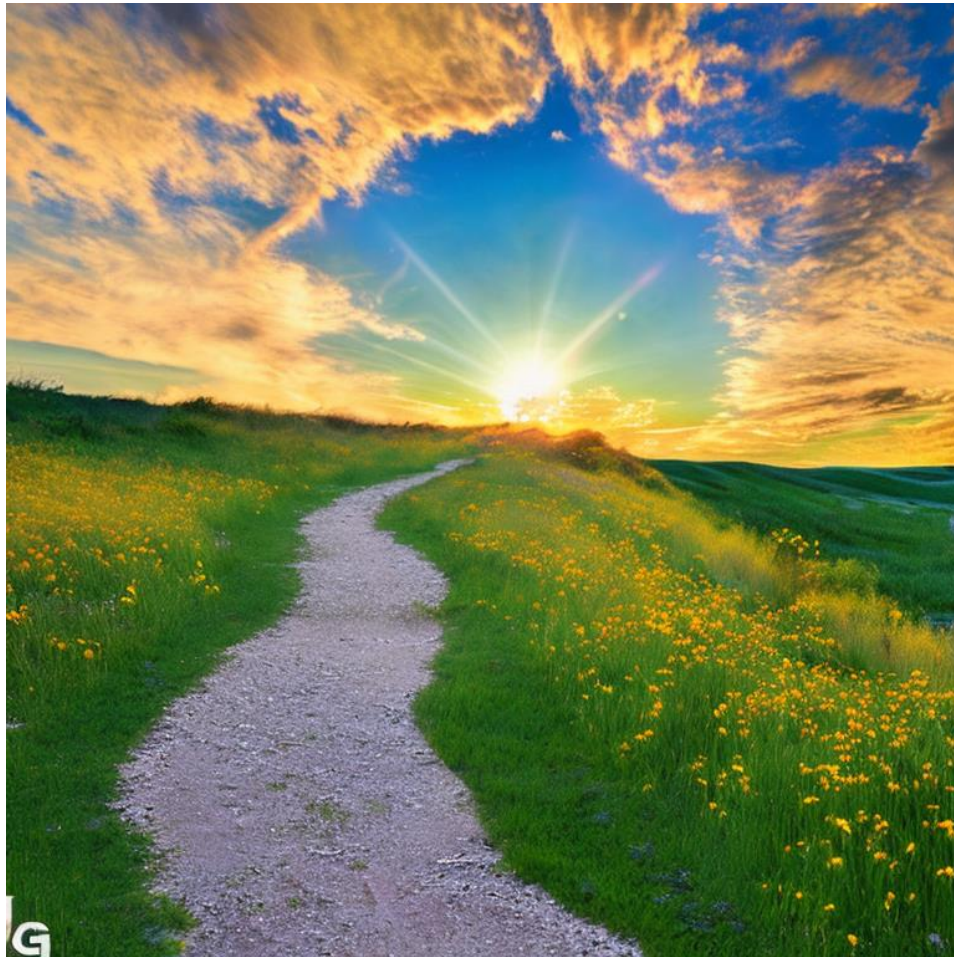
-Prompt Two First attempt :-

"Winding dirt path through lush green meadow with yellow wildflowers, rocky hill, dramatic sunset sky, sweeping clouds, soft golden light, wide-angle view, photorealistic, painterly style, magical realism"
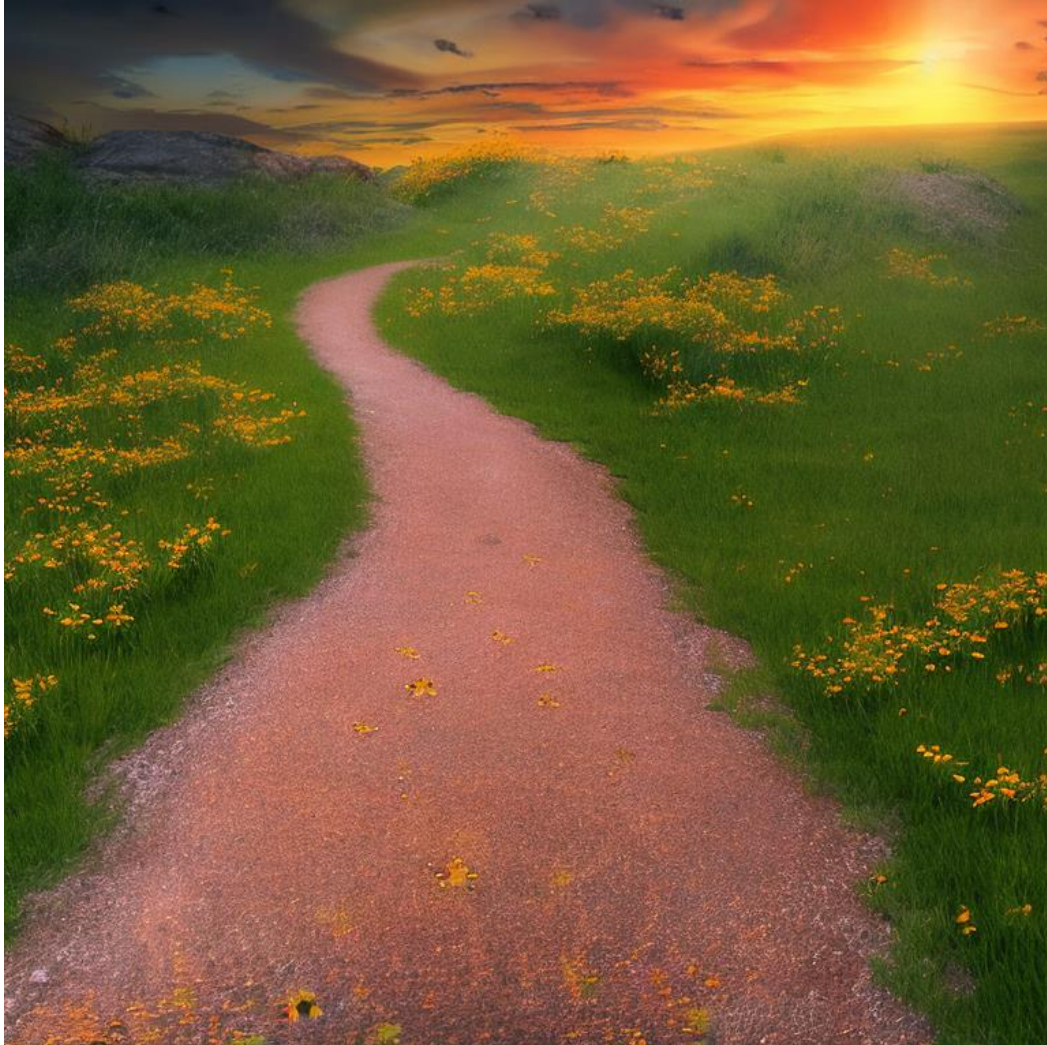
-Prompt two , second attempt :-

"Serene countryside path winding through a lush green meadow filled with blooming yellow wildflowers, dramatic sunset sky with swirling clouds, golden sunlight casting a warm glow, peaceful nature scene, springtime beauty, natural landscape photography style, vibrant colors, wide-angle view, gentle hills, soft shadows, tranquil atmosphere, realistic lighting, untouched nature, high-detail grass and flora, poetic scenery, painterly sky, horizon depth, idyllic rural trail, cinematic composition, immersive environment, photorealistic"

-Prompt Two third attempt :-

"Winding dirt path through lush green meadow with yellow wildflowers, rocky hill, dramatic sunset sky, sweeping clouds, soft golden light, wide-angle view, photorealistic, painterly style, magical realism"

-Prompt three First attempt :-

"A large dark brown sailing ship with white sails, centered on a calm sea at sunset, golden light reflecting on water, rich muted tones of gold, orange, deep blue, detailed rigging, serene atmosphere, 19th-century maritime painting style, sky fading from yellow-orange to bluish-gray, gentle waves, slight top-down perspective, balanced composition, distant smaller ship in background"

-Prompt three Second attempt :-

"A majestic old sailing ship with golden sails reflecting the warm light of a dramatic sunset, calm ocean waves with gentle reflections, cinematic lighting, highly detailed rigging, romantic maritime atmosphere, realistic water textures, 18th-century warship, atmospheric depth, oil painting style, ornate gold frame, classic marine art, seascape masterpiece, stormy sky clearing into sunlight, fine art, high realism, vintage nautical scene, museum quality, ultra-detailed"

-Prompt three third attempt :-

prompt = "A majestic old sailing ship with golden sails reflecting the warm light of a dramatic sunset, calm ocean waves with gentle reflections, cinematic lighting, highly detailed rigging, romantic maritime atmosphere, realistic water textures, 18th-century warship, atmospheric depth, oil painting style, ornate gold frame, classic marine art, seascape masterpiece, stormy sky clearing into sunlight, fine art, high realism, vintage nautical scene, museum quality, ultra-detailed"

-Parameter Sweep:

Fav_prompt = "A majestic old sailing ship with golden sails reflecting the warm light of a dramatic sunset, calm ocean waves with gentle reflections, cinematic lighting, highly detailed rigging, romantic maritime atmosphere, realistic water textures, 18th-century warship, atmospheric depth, oil painting style, ornate gold frame, classic marine art, seascape masterpiece, stormy sky clearing into sunlight, fine art, high realism, vintage nautical scene, museum quality, ultra-detailed"
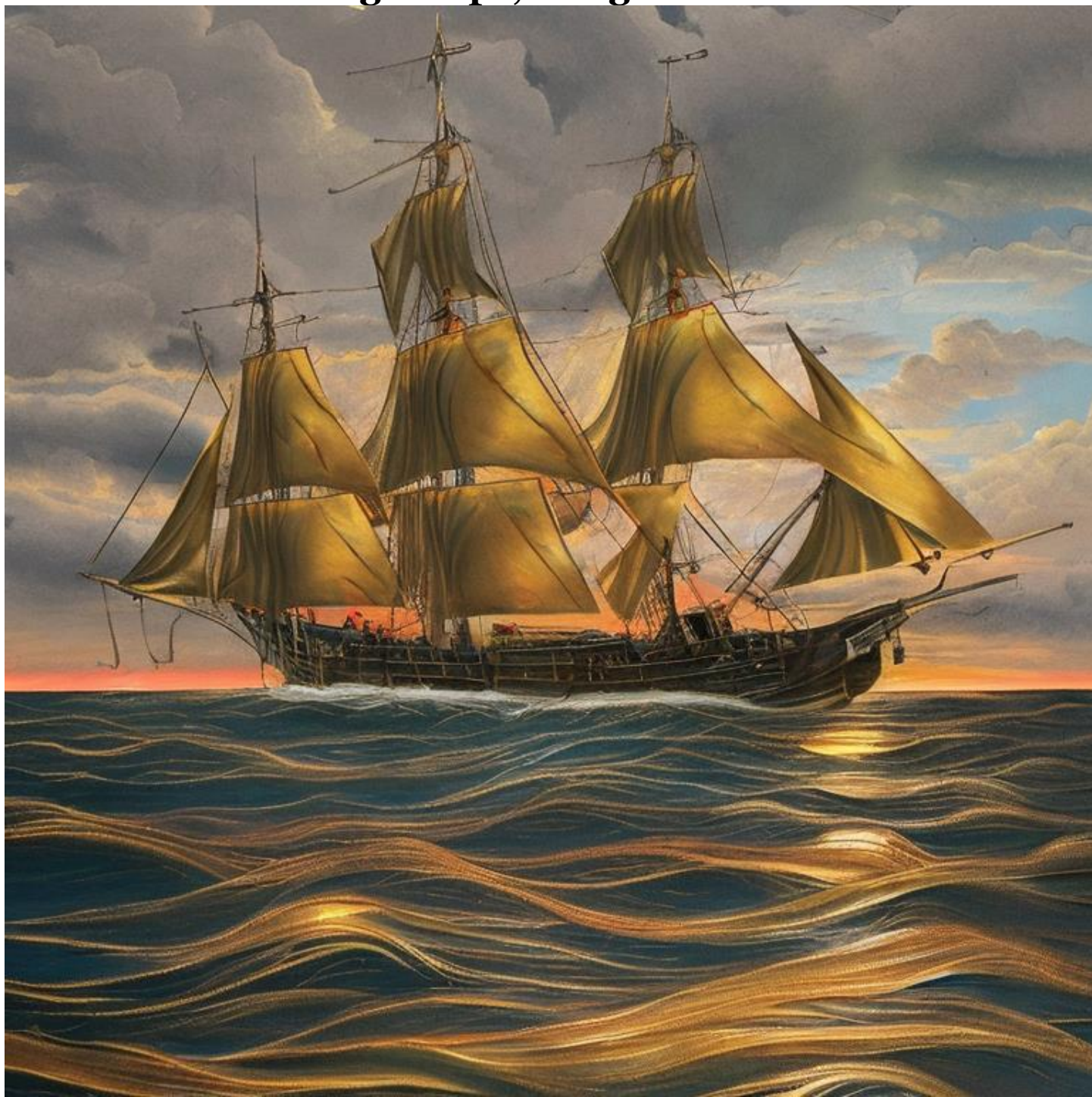
Low steps, low guidance

Low steps, high guidance

high steps, low guidance

high steps, high guidance

Reflection

1. What visual differences did you notice as you increased num_inference_steps?

With a low number of inference steps (e.g., 20), images often appear blurry or lack fine details. Sometimes, the overall structure is incomplete or "muddy," and more visual noise or artifacts may be visible. When increasing the steps to a higher value (e.g., 80), images become sharper and more detailed. The composition is more coherent, with clearer edges and textures, and usually less noise, resulting in a more polished output. However, increasing the steps also increases the generation time, so there is a trade-off between quality and efficiency.

2. How did raising guidance_scale change the match between prompt and image?

At a low guidance scale (e.g., 5), the AI behaves more creatively and may drift away from the original prompt. Images might not strictly match the description and can be more abstract or surprising. Increasing the guidance scale to a higher value (e.g., 12.5) forces the model to adhere very closely to the prompt. The resulting images are more literal and aligned with the text. However, very high guidance values can sometimes cause images to look unnatural or introduce odd artifacts because the model is "forcing" specific features to appear.

3. Explain why using manual_seed makes your run reproducible.

Using manual_seed ensures reproducibility by initializing the random number generator with a fixed value. Identical seeds, prompts, and settings (e.g., model version, parameters) yield the same image output, enabling precise replication for testing or iterative refinement. For example, torch.Generator("cuda").manual_seed(0) guarantees the same noise pattern starts the denoising process, making results deterministic.

4. Name one potential ethical concern when deploying text-to-image models.

A major ethical issue is copyright ambiguity: models trained on unlicensed data (e.g., copyrighted artwork) risk generating infringing content, raising legal and ownership disputes. For example, Stable Diffusion's training dataset includes scraped images without explicit consent, potentially replicating artists' styles without attribution. Additionally, biased outputs (e.g., racial/gender stereotypes) and misuse for deepfakes or harmful content are significant risks.

5. Suggest a real-world application where Stable Diffusion could be helpful—and one where it might cause harm.

Helpful: Medical visualization – Generating anatomical diagrams or disease progression simulations for patient education.
Harmful: Deepfake propaganda – Creating hyper-realistic fake images/videos to spread misinformation or manipulate public opinion

Helpful: Educational content creation – Generating visual aids, illustrations, or infographics to support learning materials in schools or online courses.
Harmful: Fake product advertising – Creating misleading images of products that do not exist or exaggerate features, potentially deceiving consumers.

Conclusion:-

This assignment gave me practical experience with prompt engineering and showed how different parameters like num_inference_steps, guidance_scale, and seed affect image generation quality and reproducibility.
Using <u>Stable Diffusion 2.1</u> improved the realism and detail of the outputs.
The most effective part was using <u>the custom stable diffusion img2img pipeline</u>
 to guide generation with a reference image — this gave results closer to the target image than text prompts alone.
I now understand the importance of prompt wording, parameter tuning, and pipeline choice in achieving high-quality, realistic images. I also gained insight into ethical considerations and potential real-world applications of generative models.