Image Variation Generation Using Stable Diffusion

Saurav Kumar (01FE21BEC365), Vaishnavi Agrawal (01FE21BEC168)

Under the Guidance of
Uma Mudenagudi, Ramesh Ashok Tabib
KLE Technological University, Hubballi, Karnataka, India.

December 10, 2024

Introduction to the Image Variation Generation Using Stable Diffusion

- Stable Diffusion allows users to generate multiple image variations by adjusting aspects like color, style, and content, offering creative flexibility.
- With its inpainting and style transfer capabilities, stable Diffusion enables users to refine and modify images.

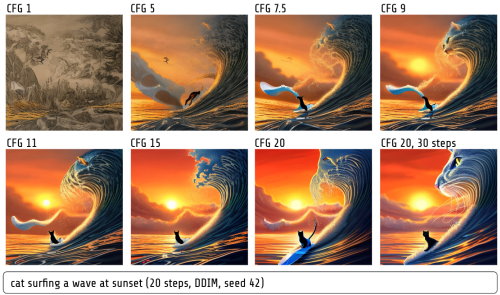

Figure 1: Stable Diffusion Classifier Guidance

- Diffusion models, like DALL-E 2 and stable diffusion, surpass traditional GANs by enabling cross-modal tasks such as text-to-image generation and image superresolution, offering greater versatility across domains.
- Superior Performance: Empirical results show that VD outperforms traditional models in both qualitative and quantitative assessments, highlighting its potential for universal AI research.

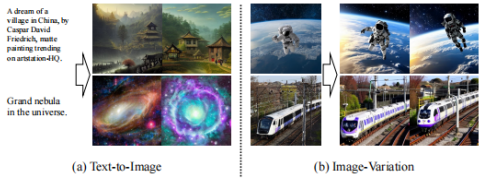

Figure 2: Demo results of our Versatile Diffusion (VD) framework on three out of all primary tasks.

- Recent personalized text-to-image (T2I) methods embed target identities into textual spaces but struggle with balancing identity preservation and diverse facial attributes. Techniques like Textual Inversion and Dreambooth face difficulties in separating identity features from irrelevant attributes.

Figure 3: The text prompt is “a woman wearing a spacesuit in a forest.”

Block Diagram for "Image Variation Generation Using Stable Diffusion "

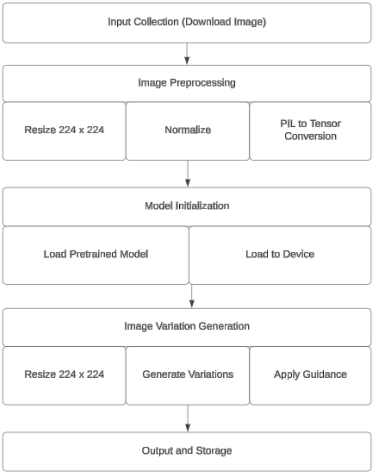

Figure 4: Proposed Pipeline

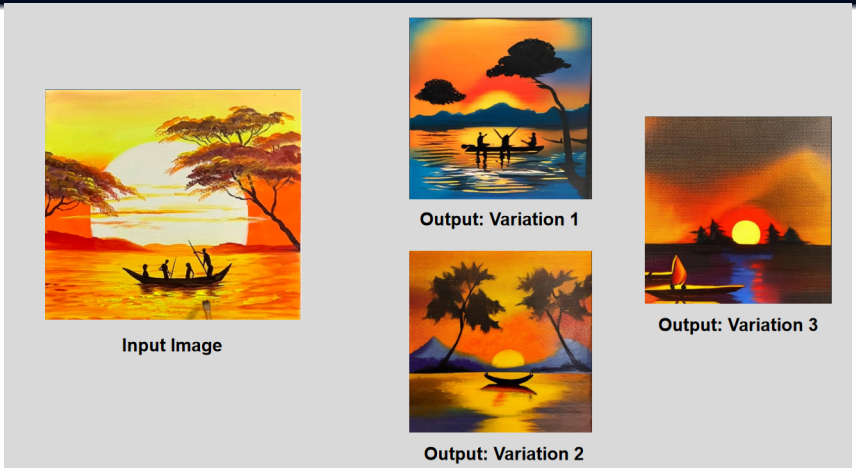

Figure 5: Input of the beach image and three output variations are generated

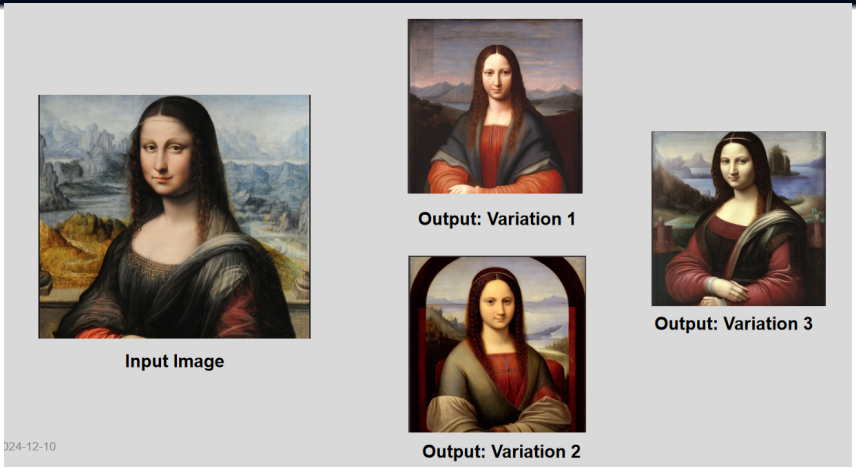

Figure 6: Input of Mona Lisa and three output variations are generated

Thank You