PAINTDIFFUSION

A Hybrid Generative AI Framework for Sketch-to-Image Synthesis

Objectives

- Convert rough sketches into realistic images
- Combine VAE, Transformer, Diffusion, and GAN techniques
- Preserve structural details from sketches
- Improve image realism and texture sharpness
- Support user control via guidance parameters
- Enable efficient processing on available hardware

Key Features

- Latent space generation reduces VRAM usage
- Mixed precision execution speeds up inference
- Classifier-free guidance controls image style/strength
- Stable, predictable results with diffusion models
- Edge detection ensures sketch fidelity

Applications

- Digital art
- Concept design
- Animation pre-visualization
- Film/Gaming asset creation
- AR/VR environments
- AI-assisted drawing tools

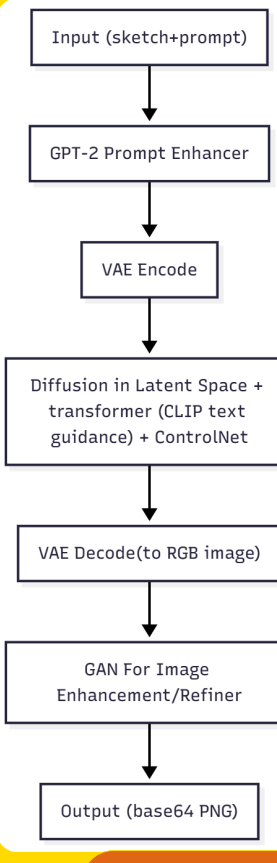
Future Enhancements

- Text + Sketch multimodal fusion
- Full-scene image generation
- Optimized inference via model quantization
- Mobile/real-time deployment

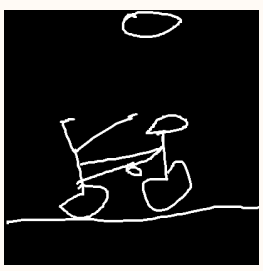
Ethical Considerations

- Avoid misuse of realistic AI-generated images
- Address dataset bias
- Transparency in generated content
- Ensure responsible deployment

FLOW



Results


Input Sketch

VAE Output

Transformer Output

Diffusion Output

GAN Output

COMPARATIVE ANALYSIS

Aspect	Variational Autoencoder (VAE)	SRGAN (Super-Resolution GAN)	Diffusion Model	GPT-2 (Transformer-based)	ControlNet (Conditional Diffusion)
FID ↓	45.70	18.40	**3.60**	7.90	4.20
IS ↑ (mean ± std)	5.20 ± 0.30	6.10 ± 0.40	**8.50 ± 0.20**	7.40 ± 0.30	8.10 ± 0.30
Precision ↑	0.62	0.74	**0.88**	0.84	0.86
Recall ↑	0.58	0.69	**0.83**	0.76	0.81
LPIPS ↓	0.28	0.21	**0.09**	0.11	0.10
Loss	0.0773 / 0.0223	0.6626 / 0.0607	**0.0088**	0.045	**0.0088** (similar to Diffusion)
Visual Quality	Low realism, blurry edges	Improved sharpness, minor artifacts	Photorealistic, excellent texture	High-quality, smooth tone balance	Realistic, well-aligned with input sketch
Best Use-case	Basic reconstruction, grayscale to color	Enhancing details in low-res images	High-fidelity image synthesis	Semantic-guided or text-aware generation	Conditional sketch-to-image generation

Student Details

- 202302040004 Nabil Ansari
- 202302040019 Gourav Sable
- 202201040015 Vaibhav Shinde
- 202201040135 Priyanshu Wagh

Course: Generative Artificial Intelligence (PEC)

Guide : Prof. Savita Mane