RESEARCH ABSTRACT

Project no : P6
Project title : Computer Vision Applications based on deep learning techniques.

Topic: AI-Generated Content (AIGC) for Text-to-Image/Video Generation.

Artificial Intelligence-Generated Content (AIGC) has the potential to revolutionize automated image and video generation using deep learning models. This project aims to explore text-to-image and text-to-video generation through Stable Diffusion, Stable Video Diffusion, and Generative Adversarial Networks (GANs). The objective is to design a structured workflow that could enable the creation of high-quality visual content from textual descriptions.

Workflow and Implementation:

The proposed workflow would involve **environment setup, model deployment, fine-tuning, and inference optimization**. The implementation could leverage **Python 3.10**, **PyTorch**, and **Hugging Face Diffusers** for model training and inference. Pre-trained models such as **Stable Diffusion v1.5**, **DreamBooth** (for fine-tuning), and **ControlNet** (for structure-guided generation) might be utilized.

To ensure computational efficiency, **CUDA and PyTorch AMP (Automatic Mixed Precision)** could be employed for GPU acceleration. Image and video processing might integrate **OpenCV, PIL (Pillow),** and **MoviePy**. Training optimizations could involve **Accelerate** and **WandB (Weights & Biases)** for performance tracking. The system could potentially be deployed through **Gradio** for an interactive web-based UI or **Flask APIs** for integration into larger applications.

Future evaluations might focus on metrics such as **FID** and **CLIP Similarity Score** to assess image quality and text coherence. This research could provide a scalable and efficient approach for AIGC applications in digital art, content creation, and media production.