Anmol Agarwal

Roll No.: 23051
Bachelors of Science, Engineering Science
EECS (Electrincal Engineering and Computer Science)
Indian Institute of Science, Education and Research, Bhopal

+91-6396855860
anmolagrawal5704@gmail.com
anmol23@iiserb.ac.in
GitHub Profile
LinkedIn Profile

Education

•Indian Institute of Science, Education and Research, Bhopal *2023-Ongoing*
Bachelors of Science, Electrical Engineering and Computer Science CGPA: 8.5

•Jim Corbett Public School *2022*
CBSE, Uttar Pradesh Percentage: 91.8%

•St. Peter's College, Agra *2020*
ICSE, Uttar Pradesh Percentage: 96%

Experience

•Vision and AI Lab, Indian Institute of Science *December 2024 - Ongoing*
Research Intern Bangalore

- Creating a dataset for Object Relighting using BlenderProc API using procedural camera poses and Environment lighting maps.
- Literature Survey of cutting edge CVPR2025 papers on Video Diffusion models.
- Efficient LoRA finetuning of WanVideo 1.3B model on custom dataset using Diffusion-Pipe and inference using Huggingface Diffusers
- Exhaustive Literature Survey on SOTA Video Diffusion models and their applications.
- Analyzed model architectures, performance trade-offs, and innovative features, such as spatiotemporal attention, causal video generation and their application in sparse-view 3D reconstruction.
- Ran and compared three state-of-the-art models (Hunyuan by Tencent, Mochi1, and CogVideoX) on server-grade hardware, utilizing pipelines such as HuggingFace Diffusers, ComfyUI, and xDiT for multi-GPU inference..

Personal Projects

•DreamLab - Latent Diffusion Transformer based model for Text2Video Generation *GitHub Link*
Designed and inspired as a mini version of Sora text2video from OpenAI

- Tools & technologies used: PyTorch, Autoencoders, Transformer, Google Colab
- Experimented with multiple autoencoder architectures, eventually trained a model which used a combination of perceptual and reconstruction loss to get minimal quality loss.
- Studied reasearch papers such as DDPM, LDM, VAE, VQ-VAE, Vision Transformer, Diffusion Transformer
- Implemented denoising using DiT architecture for images without text conditioning
- Creating a DiT denoising model with spatio-temporal attention and text conditioning using CLIP embeddings, trained on a custom Minecraft gameplay dataset as a proof of context.

•Sim3D - Realtime Cloth Simulation and Physics Engine from Scratch *GitHub Link*
Implemented the entire rendering engine and physics from scratch with a teammate.

- Developed a 3D physics simulation engine in C++ using OpenGL, featuring textured cloth simulation with mass-spring systems.
- Implemented advanced lighting, including Phong shading, texture mapping for realistic rendering, and batch rendering for efficient GPU utilization.
- Particle system with efficient collision detection using Uniform grid partitions.
- Interactive scene editor using ImGui for easy configuration and gizmos for intuitive scene manipulation.
- Applied OOP concepts to abstract and integrate different engine components, improving engine design.

•Image Captioning using CNN + RNN on Flickr8k Dataset *GitHub Link*
Flickr8k contains 8000 images with 5 captions/image

- Tools & technologies used: PyTorch, CNN, RNN, Google Colab
- Generated natural-sounding captions, though 50% of the time the captions were incorrect
- Integrated both the CNN and RNN components into a pipeline without using any pre-trained models
- CNN architecture very close to VGG-16 model optimised for feature extraction
- Simple 1 layer Reccurrent Neural Network for text generation

•Deep Learning Model from Scratch in Java
Trained on MNIST handwritten digit recognition

- Implemented forward propagation, backpropagation, and gradient descent using standard Java Libraries.
- Trained the model on the MNIST digit dataset, achieving up to 80% test accuracy.

Technical Skills and Interests

Languages: C++, C, Java, GLSL(shading Language), Python, MATLAB
Developer Tools: Visual Studio, Git, Github, Linux
Libraries and APIs: HuggingFace Diffusers, PyTorch, OpenGL, DearImGUI, NumPy, Matplotlib
Soft Skills: Presentation skills, Teamwork, Quick Learner
Coursework: Linear Algebra, single Variable Calculus, Introduction to Programming in C
Areas of Interest: Computer Vision, GenAI, Diffusion Models, Transformers, Graphics Programming, Simulation

Positions of Responsibility

•**Coordinator,** EECS Club, IISER Bhopal *2024-2025*