PRANAY MESHRAM

pranay23@iiserb.ac.in | Website | Github | LinkedIn

EDUCATION

- **Indian Institute of Science Education and Research Bhopal** — *2023 – Present*
  B.S. in Data Science Engineering — Bhopal, India

PUBLICATIONS

- **COMBAT: Conditional World Models for Behavioral Agent Training**
  Anmol Agarwal, **Pranay Meshram**, Sumer Singh, Saurav Suman, Andrew Lapp, Shahbuland Matiana, Louis Castricato, Spencer Frazier
  International Conference on Computer Vision (ICCV) *Workshop on Reliable and Interactable World Models (Oct 2025)*

EXPERIENCE

- **Research Intern** — *July 2025 – Present*
  Indian Institute of Technology Roorkee — Remote
  - Developing novel methods for **Visual Place Recognition (VPR)** and **Image Geolocation** under *Dr. Gaurav Kumar Nayak*.
  - Created a pipeline for generating synthetic 2D viewpoints from 3D indoor environments using *Infinigen*, producing a dataset of over 5,00 scenes for VPR training.
  - Conducted a systematic literature review of top-tier A* papers in VPR and visual geolocation.

- **Summer Research Intern** — *May 2024 – July 2025*
  Multi-Agent Simulation & Control Lab, IISER Bhopal — On-site
  - Implemented Vision-Language Models for segmentation tasks and analyzed SOTA methods including ZegCLIP, ZegFormer, and CATSeg.
  - Explored cross-modal feature alignment and contrastive loss optimization for improved image-text correspondence.
  - Conducted a literature review on Open-Vocabulary Segmentation and multi-modal learning.

PROJECTS

- **SigLIP-Seg** [⟳]
  Developed a cost aggregation framework for efficient segmentation using the SigLIP2 encoder
  - Trained CAT-Seg (Cost Aggregation for Open-Vocabulary Semantic Segmentation) using the COCO-Stuff dataset.
  - Implemented cost volume aggregation and cross-attention layers for fine-grained segmentation.
  - Designed custom preprocessing and evaluation pipelines for large-scale segmentation experiments.
  - **Tools:** PyTorch, Hugging Face, Detectron2, Matplotlib

- **Occupancy Estimation using Thermal Imagery** [⟳]
  Developed a real-time occupancy detection system using thermal imagery and edge deployment
  - Built a custom thermal dataset (1,000+ samples) using Raspberry Pi + MLX90640 thermal sensor.
  - Trained a YOLOv8-based detection model with data augmentation and quantization for edge deployment.
  - Deployed real-time inference on Raspberry Pi for lightweight occupancy monitoring.
  - **Tools:** PyTorch, Matplotlib, Raspberry Pi, NumPy

- **4-bit LLaMA2-7B** [⟳]
  Fine-tuned LLaMA2 on the Puffin Dataset for creative text generation
  - Applied LoRA and 4-bit quantization to improve model efficiency while reducing memory consumption.
  - Utilized PyTorch, Transformers, PEFT, and QLoRA; configured BitsAndBytes for quantization on Google Colab.
  - Modified a custom dataset from Hugging Face to match the LLaMA2 prompt template.

- **VLM-PaliGemma (from Scratch)** [⟳]
  Implemented the PaliGemma Vision-Language Model architecture from scratch
  - Reproduced the Transformer, CLIP, and SigLIP architectures in PyTorch.
  - Implemented Rotary Positional Embeddings, Key-Value Caching, etc.

SKILLS

- **Programming Languages:** Python, C, Wolfram, MATLAB, JavaScript, SQL, Bash
- **Libraries:** PyTorch, OpenCV, NumPy, Matplotlib, Pandas, Scikit-Learn
- **Tools and Platforms:** Git, GitHub, Docker, Linux, Jupyter, VS Code, Vim
- **Areas of Interest:** Computer Vision, GenAI, Vision-Language Models, Transformers, Diffusion Models, World Models

Positions of Responsibility

- **Coordinator** *2024 – Present*
  Electrical Engineering and Computer Science Club, IISER Bhopal [🌐]
- **Media Lead & Core Committee Member** *2024 – 2025*
  Computer and Networking Council, IISER Bhopal [🌐]