Week 2 Update

Shriyadita Medavaram

Background

- Mechanical Engineering at Virginia Tech
  - Rising Junior
  - Coursework: Calculus 1 and 2, Multivariable Calculus, Linear Algebra, Differential Equations
- Experience
  - Virginia Tech Astrobotics – Robotic Manipulator Mechanical Lead
  - ISAME Project (In-Space 3D printing) – Mechanical Team – Enclosure
  - Analyzing State-Level Environmental Legislature with LLMs



Project 12: Efficient and Robust VLA Models For Real-Time Robotic Manipulation

- Problem
  - Open-source models are strong, but slow and computationally expensive for real-time manipulation
- Goal: Create Efficient and Robust VLA model for robotic manipulation
  - Apply to laparoscopic surgery



Summer Schedule

- Week 1: Load Pi-0.5 and Libero on clusters and evaluate
- Week 2-3: Look into Pi-0.5 model and find bottleneck (Run profiles)
- Week 4-5: Run compression experiments (INT8, INT4 Quantization, LLaVA-PruMerge)
- Week 6: Action-Head Distillation, Full Libero evaluation
- Week 7: Robustness under Async inference, fine-tune model
- Week 8: Make poster and release code

This Week

- Goals
  - Load Pi-0.5 (Model)
  - Load Libero
  - Re-implement Inference
- Progress
  - Loaded Pi-0.5 and Libero on clusters
  - Need to re-implement inference

Short Term Plan

- End of this week
  - Re-implement inference
  - Finish adapting arm for surgery use
- Next week
  - Run latency and memory profiles
  - Identify bottleneck