CRCV REU Week #6

Project: Geo-HiPO

Advisors:

- Rohit Gupta
- Parth Kulkarni
- Dr. Shah

David Rivard

07/03/2025

Project Overview



Motivation:

- Current models use flat classification
  - Don’t generalize well
  - Use cross-entropy loss
  - No distance-based error

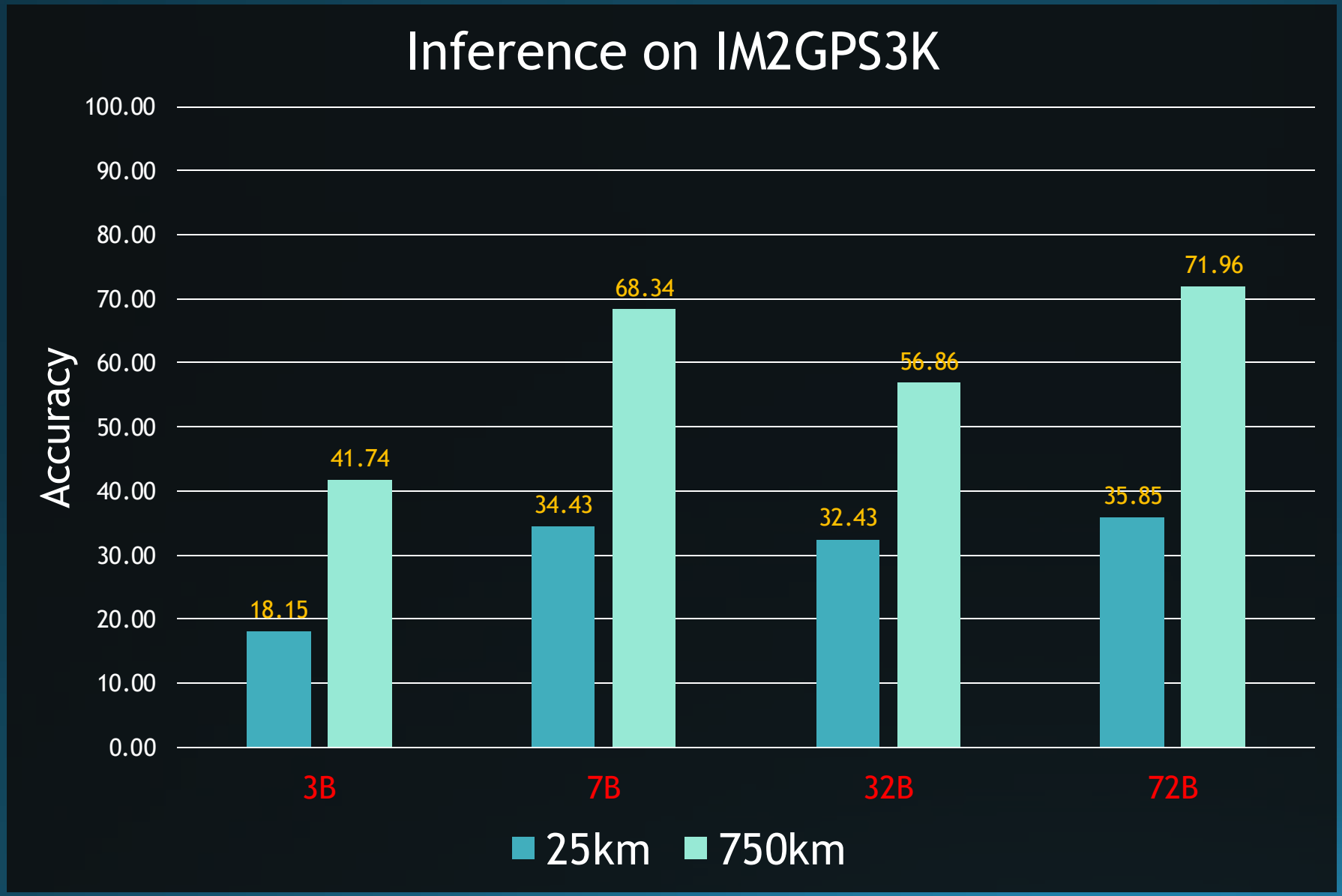
Research Question:

- Can we improve geolocation accuracy by **teaching vision-language models to reason location like humans?**

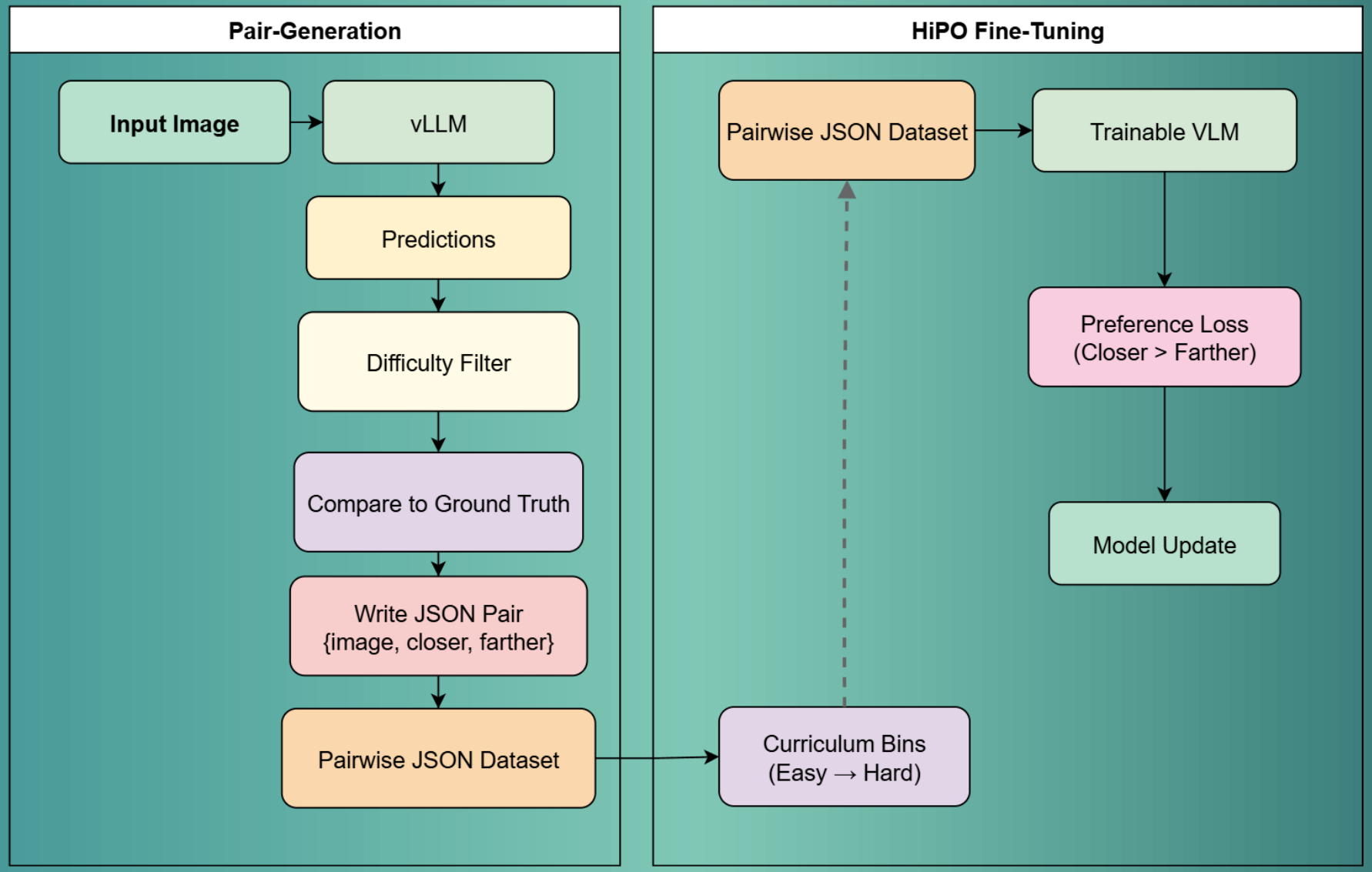
Key Progress

- Baseline SFT
  - Ran minimal supervised fine-tuning
- Filtering Training Samples
  - Filtering out easy and hard samples
- vLLM Integrated
  - Needed to finish HiPO in time
- Fixed Evaluation Pipeline
- Standardized Prompt

Corrected Accuracy



Training Plan



Preference Pairs


```
[
  {
    "image": "123456789.jpg",
    "closer": "Paris, France",
    "farther": "Rome, Italy"
  },
  {
    "image": "987654321.jpg",
    "closer": "Beijing, China",
    "farther": "Tokyo, Japan"
  }
]
```

Too Hard





Too Easy







Plan


~~Research Papers~~
~~Environment Setup~~
~~Benchmark Datasets~~
~~Baseline Inference~~
~~Deeper Model Inference~~
~~Setup vLLM~~
Filter Training Samples
Implement HiPO
Evaluation & Refinement
Finalize Report & Paper

Thank You

Questions ?