Project Name:

Person Re-identification

Course Name: CSE641 Computer Vision
Professor Name: Prof Mehul Raval
University: Ahmedabad University

Team Members:

1. **Shlok Shelat** - AU2240025
2. **Shrey Salvi** - AU2240033
3. **Rushi Moliya** - AU2240020

Summary of Work Done This Week:

Weekly Report 5:

Training & Evaluation on Flickr30kAccomplishments:

Trained multiple CLIP model variations on Flickr30k data with:Self-attention, cross-attention and sparse attention. ResNet-50 and FastViT-T8 encoders.

Evaluated model(s) based on:

Recall@1, @5, @10. Validation loss. Inference and memory usage. Best performance achieved was using FastViT-T8 with cross attention (Recall@10: 1.0000).

Barriers:

Sparse attention had high memory usage with slow inference time (1.7s). ResNet-50 performed very well, but required considerably more memory compared to FastViT-T8.

Future work:

Start training and testing on RSTP-ReID. Attempt to improve efficiency while maintaining high retrieval accuracy.