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:

Tasks Completed:

Review of the literature related to the CLIP model, DistilBERT, ResNet-50, FastViT-T8, and attention mechanisms (self, cross, sparse).Basic understanding of text-based person reidentification. Learned about the structure and limitations of the Flickr30k and RSTP-ReID datasets.

Foundations have been laid down to analyze implementations on optimizing attention in multimodal settings.

Tasks Completed:

Read and comprehend the complexities of cross-attention and the interaction of sparse attention with CLIP.Selected an image encoder (in a dual-encoder setup), by initial standards, that achieves a compromise between performance and speed.

Right now, We are planning to:

Start setting up the environment and preprocessing the datasets.Finalize architecture for a CLIP capable dual-encoder.