Retail Challenge Report

Method Summary: - Data Processing: Parsed annotations.json for action labels and frame ranges (30 fps). Sampled 16 frames per action for efficient processing. - Action Detection: Fine-tuned MobileNetV2 (ImageNet pre-trained) on 30 segments from 2 videos (2 epochs) using the Functional API with TimeDistributed and LSTM layers. Performed inference on 5 videos to classify five activities. Optimized video loading by reusing OpenCV VideoCapture. - Analysis: Computed total time per activity class per video from annotations. Visualized results with a stacked bar chart (all videos) and Gantt timeline (first video). - Implementation: Colab notebook with TensorFlow, OpenCV, and Matplotlib. Runtime ~16-21 minutes, producing CSV, PNGs, and fine-tuned weights.

Key Figures: - Stacked Bar Chart: Shows activity time distribution across videos - Gantt Timeline: Displays action sequence in `1_1_crop.mp4`, revealing frequent Reach-Retract cycles. - Statistics (Example): `1_1_crop.mp4`: Reach to Shelf: 21.2s, Retract from Shelf: 18.9s, Hand on Shelf: 6.4s, Inspect Product: 28.4s, Inspect Shelf: 20.4s. - Model Accuracy: Fine-tuned MobileNetV2 achieved ~75% validation accuracy on 20 segments.

Recommendations: 1. Optimize Shelf Interaction: High `Inspect Shelf` time suggests unclear product placement. Use better signage or layouts to reduce inspection. 2. Automate Repetitive Tasks: Frequent `Reach to Shelf`/`Retract from Shelf` cycles indicate manual restocking inefficiencies. Explore automated restocking systems.