

Computer Vision Assignment – Dwell Time Estimation in a Retail Environment

Objective:

The goal of this assignment is to design a computer vision system that can detect, track, and estimate dwell time of individuals in a retail setting. The task simulates real-world analytics often used to understand customer behavior in physical stores.

Problem Statement:

Given a retail video (or any crowd/pedestrian-based input), your task is to:

1. Detect and track individuals across frames using any model and tracker of your choice. Can consider either the specific section of frame or the entire frame.
2. Calculate dwell time for each individual based on how long they remain in the frame.
3. Compute and report the average dwell time across all individuals.

Requirements & Expectations:

- You may use any object detection model (e.g., YOLO, EfficientDet, etc.) and tracking algorithm (e.g., SORT, DeepSORT, ByteTrack, etc.) of your choice.
- Your approach should handle ID consistency for each individual throughout the video.
- Accuracy expectation: A reasonably good result (~90%+ accuracy) is acceptable. Perfect results are not expected.

Final deliverables should include:

- Processed output video with tracking visualizations and overlay of dwell time info.
- Dwell time calculation output (e.g., CSV or structured logs showing time spent per ID and average dwell time).
- A brief report (PDF) outlining:
 - Tools and models used.
 - Step-by-step approach.
 - Any challenges faced and how they were solved.
 - Final observations and accuracy metrics.

Evaluation Criteria:

- Approach & reasoning
- Model and tracker selection
- Code quality and structure
- Result interpretation and reporting
- Documentation and clarity

Time Allotted:

3 days after receiving the assignment.

We have attached a sample retail surveillance [video](#) or you may use any publicly available dataset/video to demonstrate the concept. Please submit the solution along with required documents.