Object Detection and People Counting

Introduction

This document explores the comprehensive workflow for enhancing object detection and people counting. The workflow includes object detection, object tracking, unique ID assignment, centre coordinate identification, ROI assignment, movement direction maintenance, and entry/exit counting. We'll also discuss the models used for object detection and tracking, their advantages, disadvantages, and improvements made for accuracy and efficiency.

Object Detection and Tracking Models

Object Detection Model

YOLO (You Only Look Once) is a real-time object detection system renowned for its rapid and precise identification of multiple objects in a single pass within images and videos, finding extensive application in areas like autonomous vehicles, surveillance, and image analysis. It has gained widespread recognition for its performance on benchmark datasets like COCO (Common Objects in Context), which provides a comprehensive range of annotated images, enabling YOLO to excel in various computer vision tasks. This powerful combination of YOLO and the COCO dataset makes it a top choice for applications

Multiple versions of YOLO are available for object detection, each with its own set of features and improvements - YOLOv1,YOLOv2 (YOLO9000),YOLOv3,YOLOv4,YOLOv5,YOLOv6, YOLOv7, YOLOv8

Object Tracking Model

The primary purpose of object tracking is to follow the movement and location of objects in a video sequence over time. It enables the continuous monitoring, analysis, and understanding of object behaviour and interactions . Some Object tracking algorithms include -

- Mean Shift
- CamShift (Continuously Adaptive Mean Shift)
- Centroid based
- SORT (Simple Online and Realtime Tracking)
- DeepSORT (Deep Learning for Object Tracking with SORT)

Solution in Focus

Object Detection

The solution combines YOLOv8 trained on COCO dataset for object detection with DeepSORT for tracking and counting. DeepSORT assigns unique IDs to objects, improving people counting accuracy. It also updates entry/exit counts in real-time based on an object's position relative to the reference line.

Centroid-based Object Tracking

This approach involves calculating the centroid (center of mass) of an object in each frame. It tracks objects by monitoring how much the centroid has moved between frames. If the centroid moves within a certain threshold distance, the object is considered the same. This method is relatively simple and computationally efficient but may struggle with occlusions and changes in object appearance.

Advantages:

- Simple to implement
- Remarkable speed compared to other experimented solutions (9 FPS).

Disadvantages:

- Challenges arise in ID assignment, if the person takes a longer step, the ID is lost and assigned a new tracking ID. If the threshold for distance of centroids in consecutive frames is increased, IDs get exchanged for two people very close to each other.
- Performance struggles when objects are occluded.

SORT (Simple Online and Real-time Tracking)

SORT (Simple Online and Real-time Tracking) is an efficient object tracking algorithm that uses data association techniques to track objects in real-time video streams. It employs methods like the Kalman filter for state estimation and adapts tracking windows to follow objects as they move. SORT is known for its simplicity, making it suitable for various real-time tracking applications

Advantages:

- SORT could dynamically adjust the tracking window size, making it suitable for tracking people in motion.
- Removed disadvantage of centroid based method of losing ID and had consistent tracking

Disadvantages:

- With 6 FPS Less efficient than Centroid-Based Tracking
- Adversely affected by occlusion scenarios. Algorithm lost track of people whenever they entered through the gate because of occlusion.

DeepSORT (Deep Learning for Object Tracking with SORT)

DeepSORT (Deep Learning for Object Tracking with SORT) is an advanced object tracking algorithm that combines deep learning techniques with the Simple Online and Realtime Tracking (SORT) algorithm. It enhances object tracking accuracy and robustness by:

- Utilizing Deep Learning: DeepSORT incorporates deep neural networks to extract features from object detections, improving tracking performance, especially in complex scenarios.
- Data Association: It employs data association techniques to assign consistent tracking IDs to objects, even in situations with occlusions or multiple objects in close proximity.
- Handling Occlusions: DeepSORT is capable of maintaining object tracks when objects are temporarily hidden or occluded, ensuring continuous tracking.
- Predictive Tracking: The algorithm can predict object positions, making it resilient to abrupt changes in object movement.
- Multi-Object Tracking: DeepSORT excels in tracking multiple objects simultaneously, making it suitable for crowded environments.

Advantages:

- Notably improved tracking accuracy.

Disadvantages:

- Tends to slow down (3 FPS) computational systems.

FPS Optimization

Implemented:

- **Frame Skipping**: Skipping frames helped boost the FPS rate from 2 to 5.5.
- **Resolution Reduction**: Lowering the resolution of incoming frames contributed to the FPS improvement.

Proposal:

- **GPU Utilization**: Employing a GPU can further enhance FPS by leveraging its processing power.
- **Multicore CPU Usage**: Utilizing multiple CPU cores can parallelize tracking, improving frame rates.
- **DeepSORT Parameter Tuning**: Fine-tuning DeepSORT parameters can optimize processing efficiency, resulting in higher FPS.

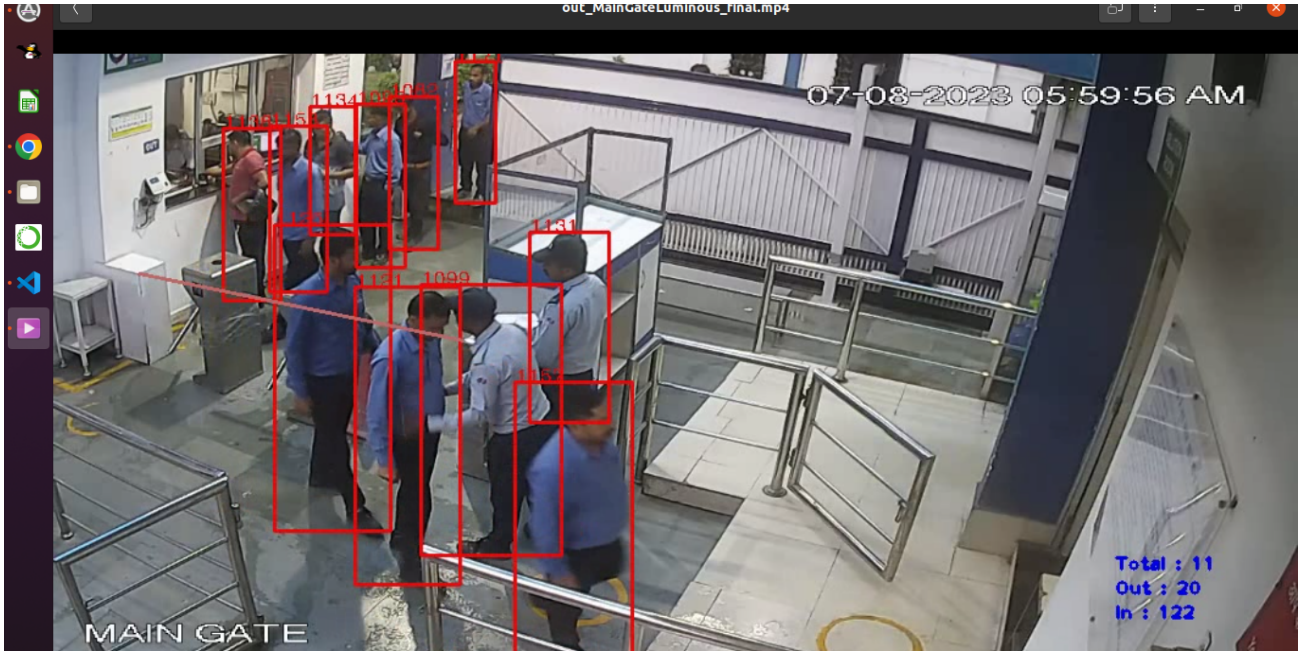
RESULTS

Actual in count : 127
Calculated by model : 122

Actual out count : 2
Calculated by model : 20

The accuracy drops when individuals are in close proximity. We need to enhance the detection and tracking capabilities, especially when two or more people are in close proximity to each other.

Additionally, the in/out counting logic should be refined to reduce the chances of miscounting individuals who have exited.



CONCLUSION

In conclusion, the project effectively fulfills the outlined requirements. The object detection model, powered by YOLOv8, accurately identifies individuals in video frames. For tracking, DeepSORT is employed to differentiate between new entries and exits. The counting logic keeps a meticulous record of individuals entering and exiting, and real-time counts are displayed on the video stream. Significant optimization efforts have improved the Frames Per Second (FPS) from 2 to 5.5, with further optimization possibilities proposed. The project is hosted on GitHub and provides an insightful exploration of different tracking models, their strengths, and limitations, leading to an informed model selection.