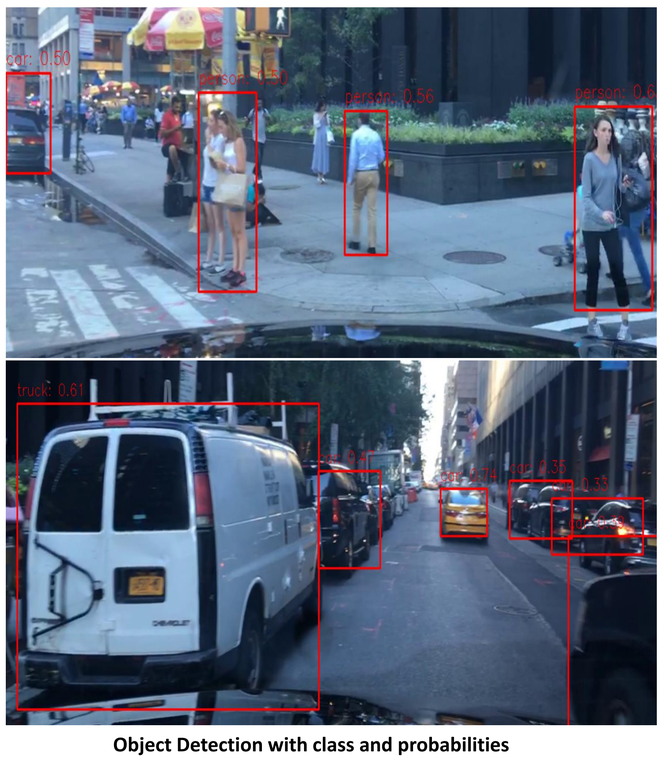
Moving Object detection and tracking with Bird's Eye View Simulation

Dhiraj Lala (22200259) and Dhaarna Waghaye (22201256)

School of Mathematics and Statistics, University College Dublin, Ireland

Summary

- Python script: This article offers a flexible Python script for in-the-moment object identification, dynamic tracking, and logical bird's-eye perspective simulation.
- Integrated Techniques: The YOLOv5m model, dynamic object tracking, and translation to a bird's-eye view are seamlessly integrated.
- Real-time processing: Enables tracking and object detection in real-time while processing video.



Object Detection with class and probabilities

Introduction

- Fusion of Cutting-Edge Techniques: The script effortlessly combines simulation of a bird's eye perspective, object tracking, and object identification.
- Holistic video analysis: gives in-depth explanations of the dynamics of objects in video streams.
- YOLOv5 for Object Identification: Utilizes YOLOv5 for real-time object detection and annotation.
- Immersive Bird's Eye View: Presents a simple top-down simulation for the display of spatial interaction.
- Multipurpose Applications: Used across a variety of industries, including traffic control, surveillance, and more.

Bird's Eye View

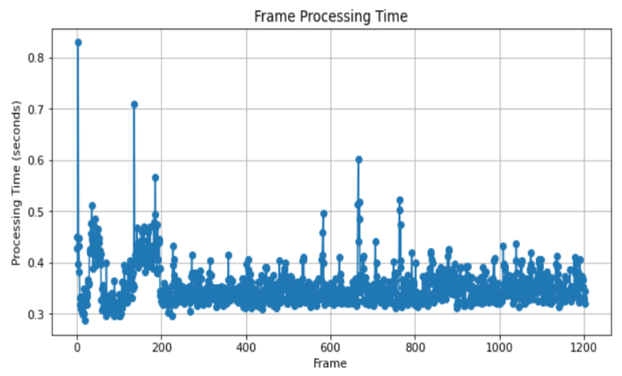
Motivation

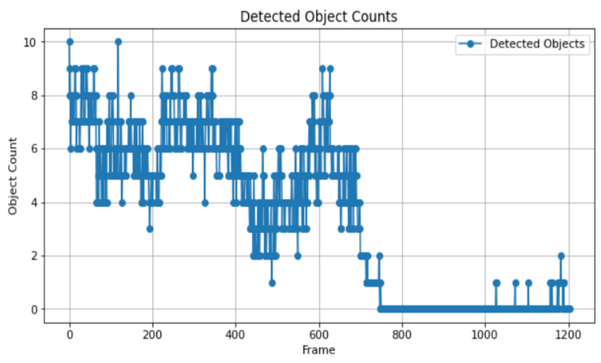
- Enhancing Situational Understanding: Traditional object detection struggles to capture changing object behaviour in dynamic situations. By including dynamic tracking, which provides a more full picture of item interactions and trajectories, this script seeks to close that gap.
- Spatial Insights Through Simulation: A comprehensive spatial viewpoint is frequently missing from traditional video analysis. This script offers a novel method to perceive item movements and spatial interactions by introducing a bird's eye view simulation, leading to deeper understandings.
- Real-world Application Needs: From traffic control to surveillance, there is a wide range of applications where accurate, real-time analysis is required. These requirements are met by this script's smooth integration of YOLOv5 detection, dynamic tracking, and a fresh simulation strategy.

Features

- Real-time Object Detection: YOLOv5 is used to quickly and precisely identify objects in video frames.
- Dynamic Object Tracking: Establishes associations between objects across frames, enabling precise trajectory tracing and interaction analysis.
- Customizable Parameters: Offers adjustable input, output, and tracking settings for versatile analysis.
- Rich annotations: These include annotated bounding boxes, labels, and trajectories, improving visual comprehension.
- User-friendly Interface: This interface was made to be accessible for practitioners, analysts, and researchers.
- Real-time Video Processing: Enables on-the-fly analysis for dynamic scenarios requiring timely insights.
- Interdisciplinary Insights: Combines computer vision and practical knowledge to help make wise decisions.

Performance Measures





Processing Time per Frame:

- This plot shows how long the object tracking and simulation system took to process each frame of the movie.
- The y-axis shows the processing time in seconds, and the x-axis corresponds to the number of frames.
- According to the plot, processing takes longer at start, peaking at about 0.8 seconds for the first frame, then stabilizing at 0.4 seconds for the majority of frames.
- These variances imply possible variations in the computing load or resource use during various stages of video processing.

Detected Object Counts:

- This graph shows how the number of detected objects has changed over the course of the video's frames.
- The y-axis shows the number of items, while the x-axis depicts the frame number.
- Up to 10 objects can be initially detected by the system, but this number gradually drops to 9.
- The number of detected objects then drops even more, to 6, before stabilizing at 2 in the final frame.
- The plot emphasizes how dynamic object detection is and makes suggestions about potential difficulties in maintaining consistent and precise tracking.

References

- R. B. Girshick, J. Donahue, T. Darrell, and J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp. 580–587, 2014.
- J. Redmon, S. K. Divvala, R. B. Girshick, and A. Farhadi, You only look once: Unified, real-time object detection, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788, 2016.
- P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun, Overfeat: Integrated recognition, localization and detection using convolutional networks, CoRR, vol. abs/1312.6229, 2014.