REAL-TIME OBJECT DETECTION AND TRACKING FOR TRAFFIC FLOW ANALYSIS

Mohana Priya N (22CSEG19)

M. Sc Data Analytics

Agenda

- Abstract
- Introduction
- Problem Statement
- Implementation
- Methodology
- Applications

Abstract

This project employs a real-time object detection and tracking system using Deep Learning and Computer Vision techniques. The implementation integrates OpenCV for video processing, YOLO (You Look Only Once), a state-of-the-art model for object detection. In addition, the project utilizes the Deep SORT algorithm for tracking detected objects across video frames. The primary focus is on detecting and tracking vehicles, with supplementary features such as displaying the current time and providing a count of detected vehicles. The system serves as a foundation for traffic flow analysis, offering insights into optimizing resource allocation and enhancing road infrastructure. The report details the methodologies employed, presents experimental results, and discusses the practical applications of the developed system.

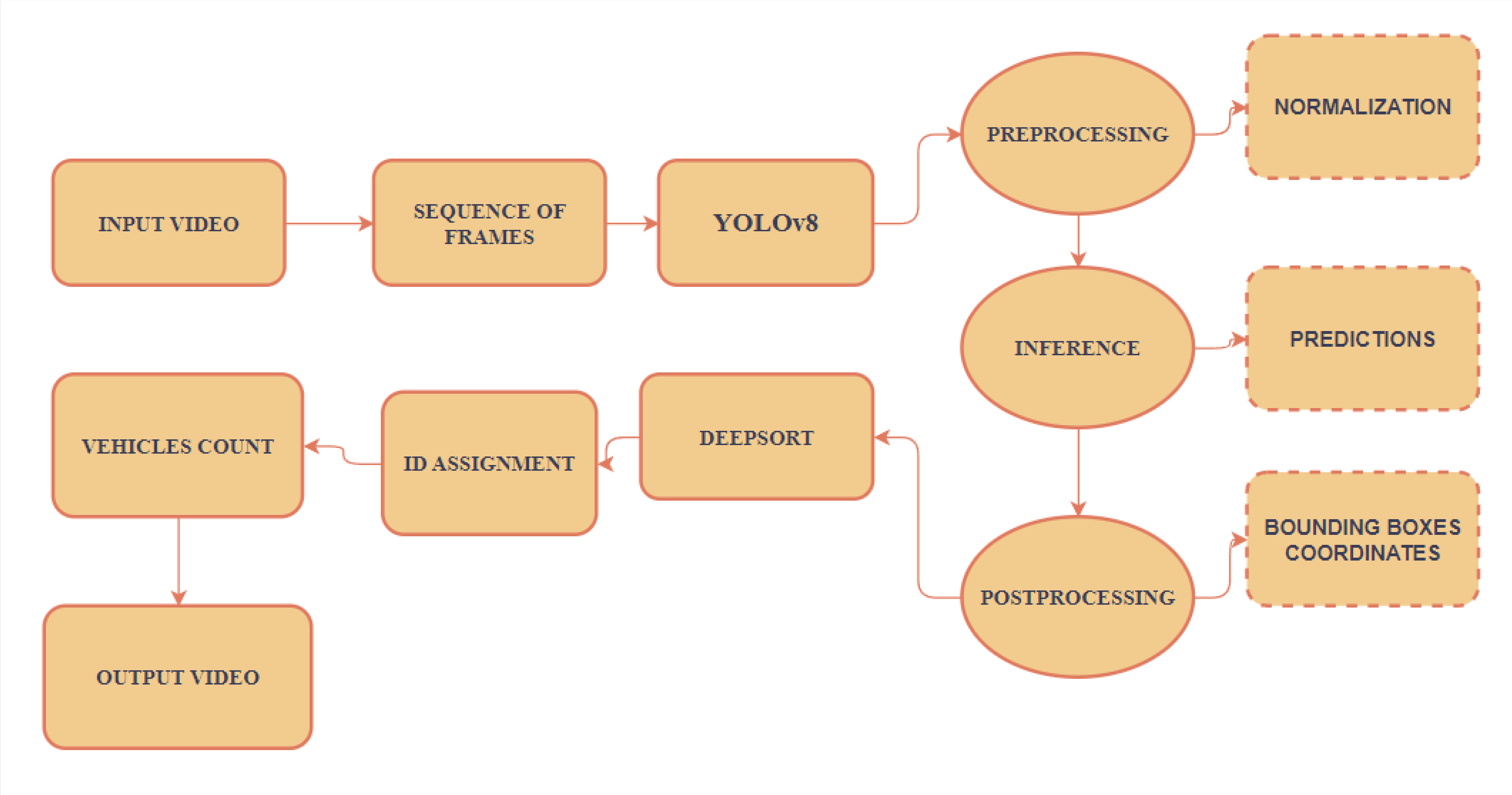
Introduction

- In the dynamic landscape of today's technology, the fusion of advanced computer vision and deep learning has become pivotal for addressing real-world challenges.
- Our primary aim is to harness cutting-edge technologies to enhance our understanding of dynamic environments, placing particular emphasis on detecting and tracking vehicles.
- The increasing prevalence of surveillance cameras and the need for efficient traffic control have generated a demand for intelligent systems capable of instant visual data analysis.
- To meet this need, our research merges OpenCV, a versatile computer vision toolkit, with the speed and accuracy of YOLO (You Look Only Once), a state-of-the-art object detection model along with the utilization of the Deep SORT method ensures stable tracking of identified objects across video frames.
- This project's core objective is to address challenges associated with real-time vehicle monitoring and detection, a critical element in traffic control systems.

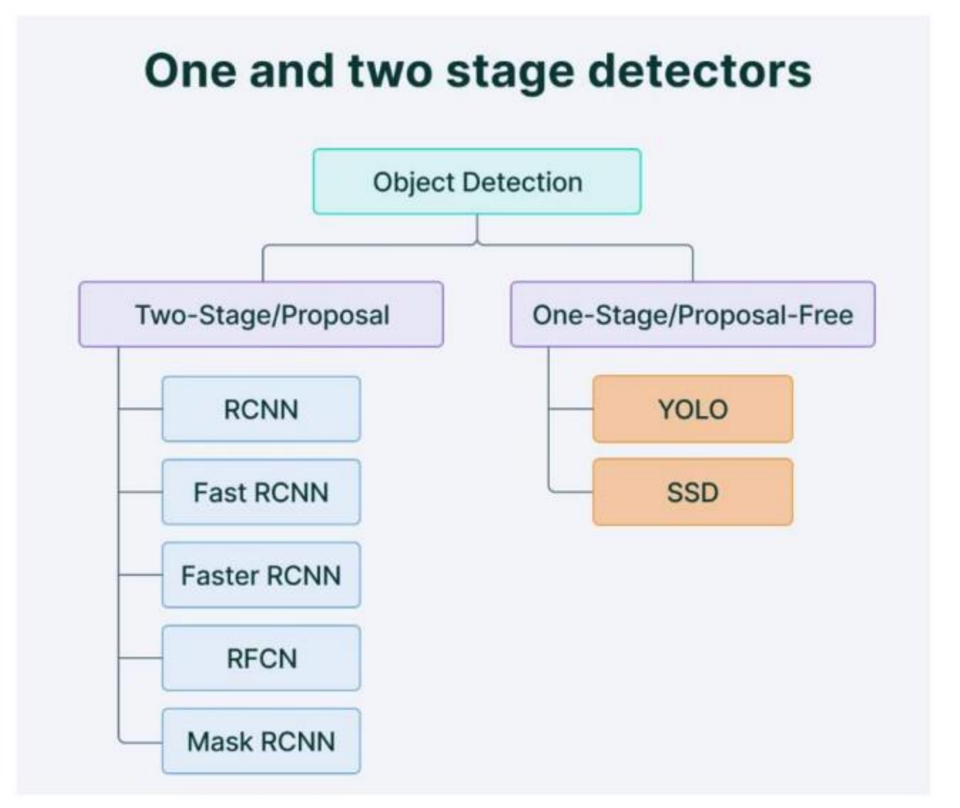
PROBLEM STATEMENT

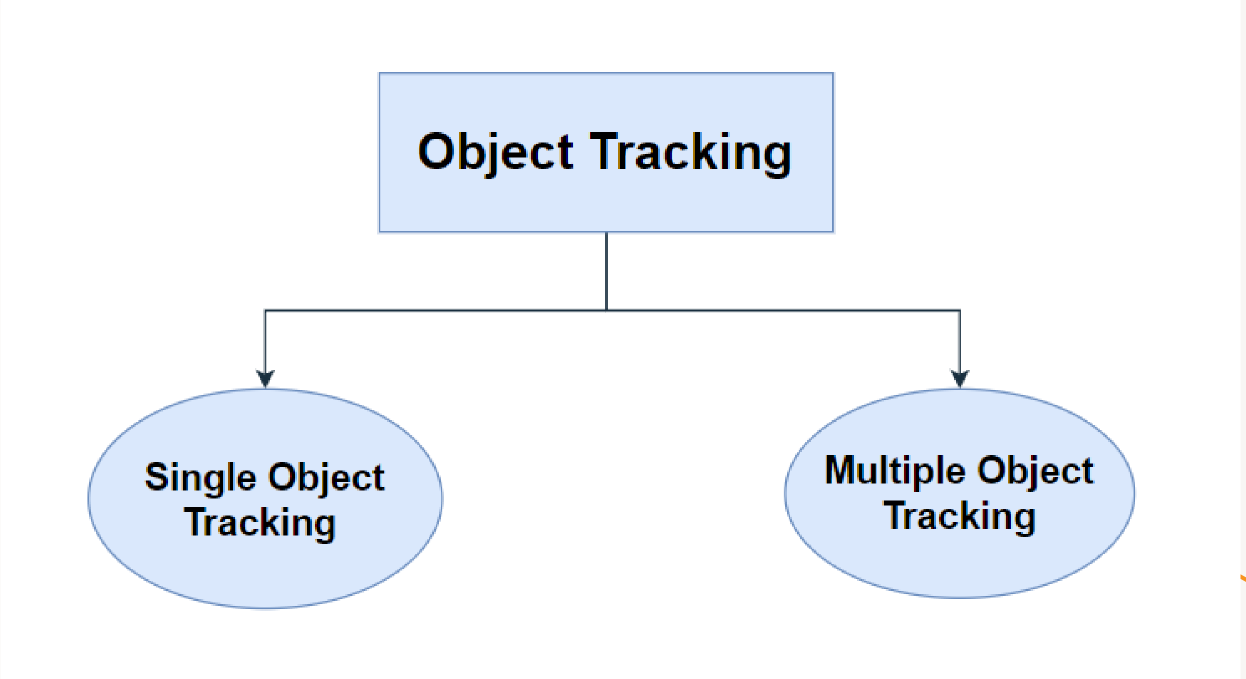
- Real-time object detection and tracking in dynamic traffic environments present formidable challenges to surveillance systems.
- The inherent complexities include delays in processing and analyzing video data, inaccuracies in detecting objects like vehicles and pedestrians, unreliable tracking mechanisms, and the crucial need for a more comprehensive understanding of traffic patterns.
- These challenges collectively hinder the efficiency of traffic management systems, compromising both the accuracy of surveillance data and the overall safety of road users.
- The primary goal of this project is to design and implement an advanced real-time object detection and tracking system tailored for dynamic traffic environments.

Implementation



Methodology





Object Localization:

Finding the location of the object in image.

But if we have multiple objects, then the vector size will increase.

For example, if we have 5 objects in single image then then vector size will be 10 x 7 = 70 vectors.

So, this method is not optimal.

$$\begin{bmatrix} P_c \\ B_x \\ B_y \\ B_w \\ B_h \\ C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 50 \\ 70 \\ 60 \\ 70 \\ 1 \\ 0 \end{bmatrix}$$

P_c = Probability of Class

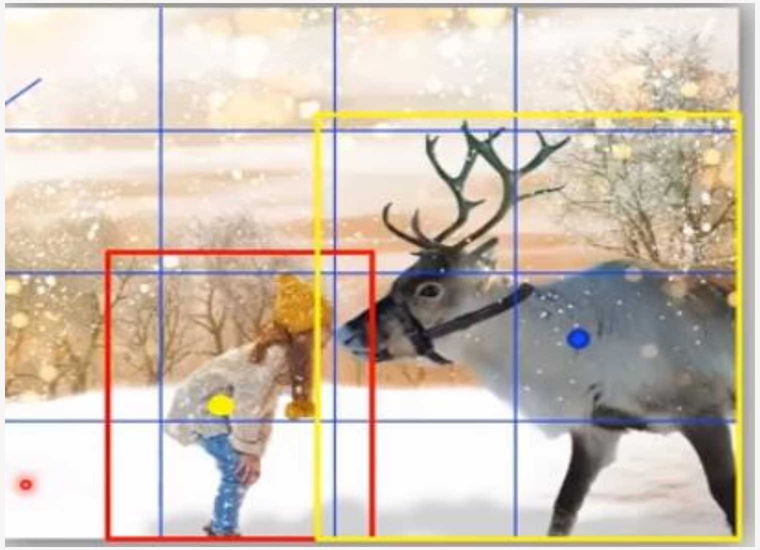
B_x, B_y = Represents coordinate for center which is indicated in yellow color

C_1 = Represents class 1 which is for Deer

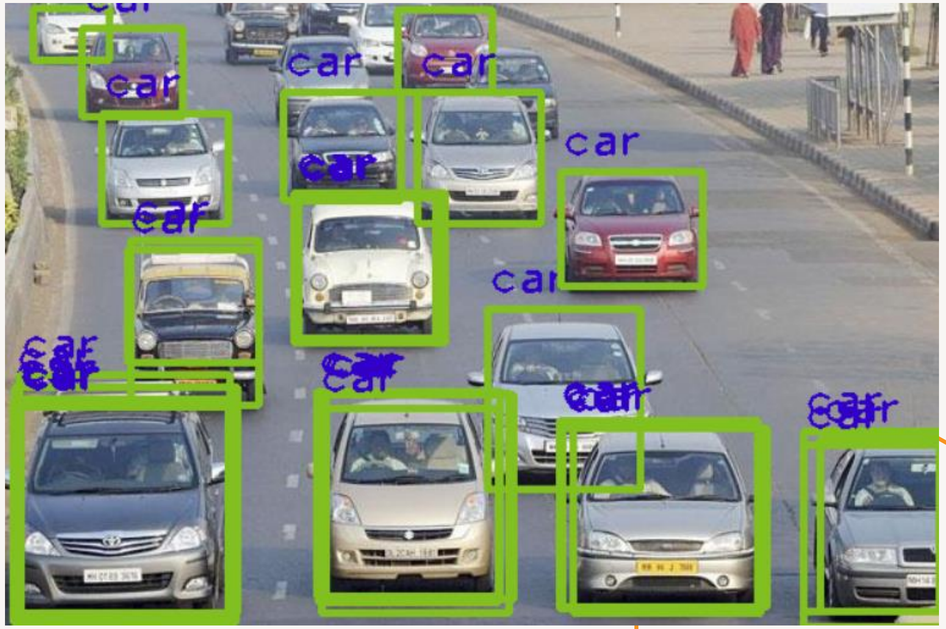
C_2 = Represents class 2 which is for Person

B_w, B_h = Represent the width and height of the Red Box

YOLO (You Look Only Once)

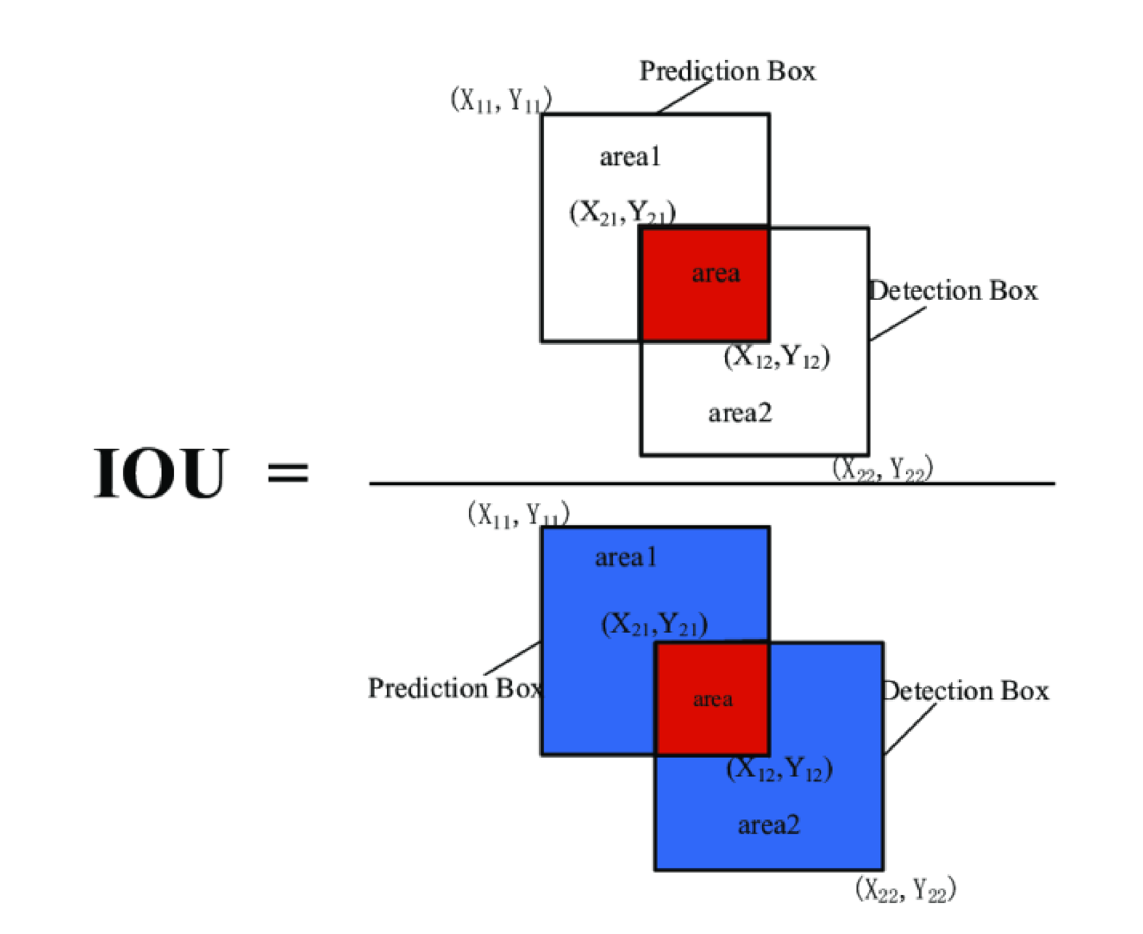


- Detects in single pass.
- Dividing the image into grids and then detecting objects in each grids.
- Each boundary box has centroid, width, height and confidence score.
- Confidence score = objectness score (the probability that there is an object within the box) and the class probability.
- Multiple bounding boxes issue
- We can’t take max of the confidence scores.
- Non-Maximum Suppression – solution



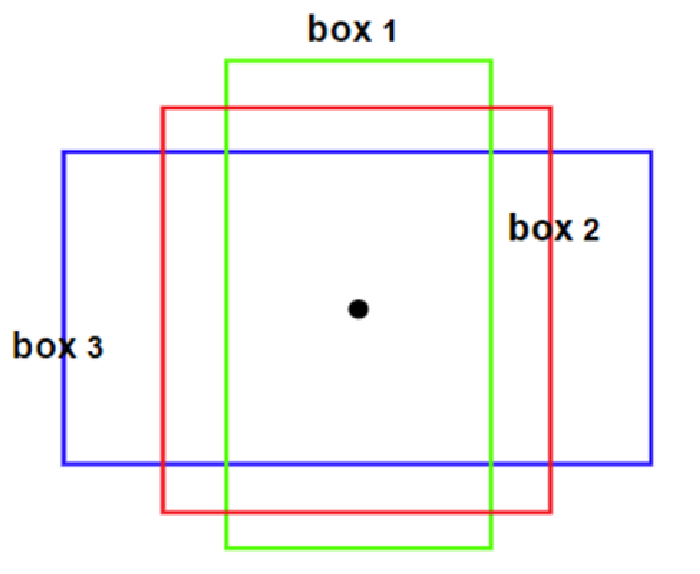
Non Maximum Suppression

- post-processing step to filter out redundant and low-confidence bounding box predictions.
- goal of NMS is to keep only the most confident and non-overlapping bounding boxes for a given object.
- Low confidence score is removed by threshold value.
- IoU is a measure of how much overlap there is between two bounding boxes.
- Bounding boxes with high threshold IoU values with a previously selected box a re considered redundant or overlapping.



Features of YOLOv8

- Able to perform classification, detection and segmentation.
- Multiple backbones supported. New backbone network introduced CSPDarknet53
- Anchor free detections – does not depends on bounding box for training
- Feature pyramid network for better detection for varying object sizes.
- Faster and accurate than previous model which can process 43 frame per second



DeepSORT (Deep Simple Online and Realtime Tracking)

- state-of-the-art object tracking algorithm designed for real-time tracking of objects in videos.

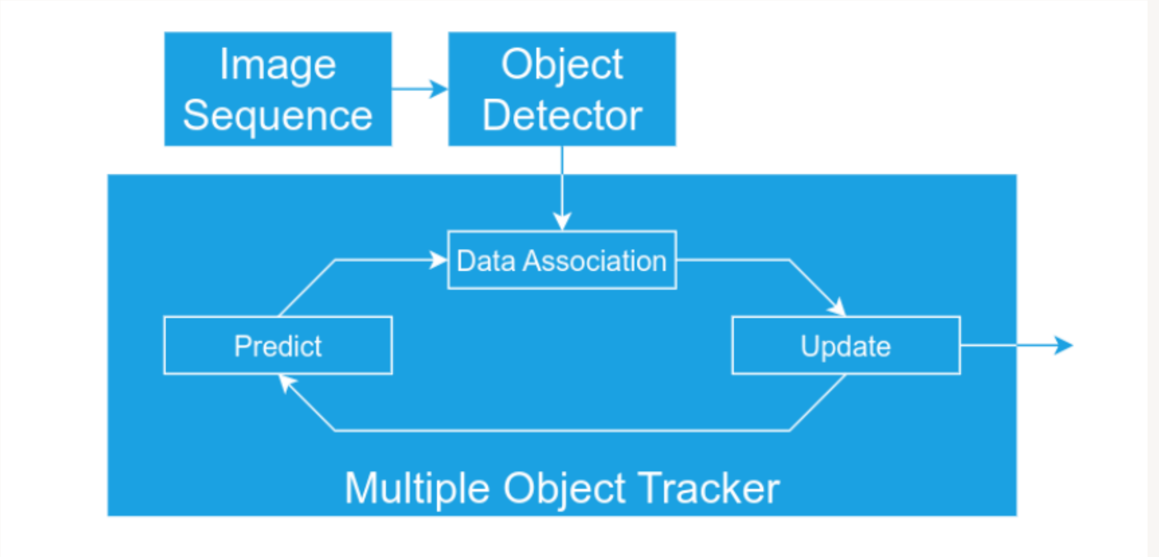
- relying solely on geometric information (position, velocity)

- Kalman Filter to predict the future position of each tracked object based on its historical motion. The Kalman Filter estimates the state (position, velocity, etc.) of each object.

- Hungarian algorithm to associate current object detections with existing tracks. cost matrix that combines Mahalanobis distance (based on Kalman filter predictions) and appearance feature distances (e.g., cosine distance).

- Minimize the total cost to find the optimal assignment of detections to tracks.

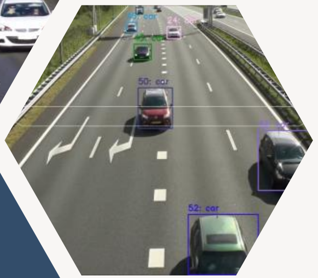
- Non-Maximum Suppression – for redundance



Applications

- **Healthcare:**
  - Monitoring patient movement within hospitals or eldercare facilities.
  - Analyzing social distancing and mask compliance in public spaces during health crises.
- **Environmental Monitoring:**
  - Wildlife tracking and conservation efforts.
  - Monitoring environmental changes and responding to natural disasters.
- **Entertainment:**
  - Augmented reality (AR) applications that interact with real-world objects.
  - Virtual reality (VR) environments with realistic and dynamic elements.
- **Human-Computer Interaction:**
  - Gesture recognition and tracking for controlling devices.
  - Enhancing virtual collaboration tools with real-time video analysis.
- **Sports Analytics:**
  - Player tracking in sports for performance analysis and fan engagement.
  - Automated referee assistance systems.
- **Education:**
  - Monitoring student engagement and attendance in classrooms.
  - Enhancing virtual learning environments with interactive elements.
- **Social Distancing and Health Safety:**
  - Monitoring social distancing compliance in public places.
  - Identifying potential health risks in crowded areas.






Thank you!