

WEEK 5,6

Tiếp tục Object Tracking : Multi Object Tracking



Đặt vấn đề :

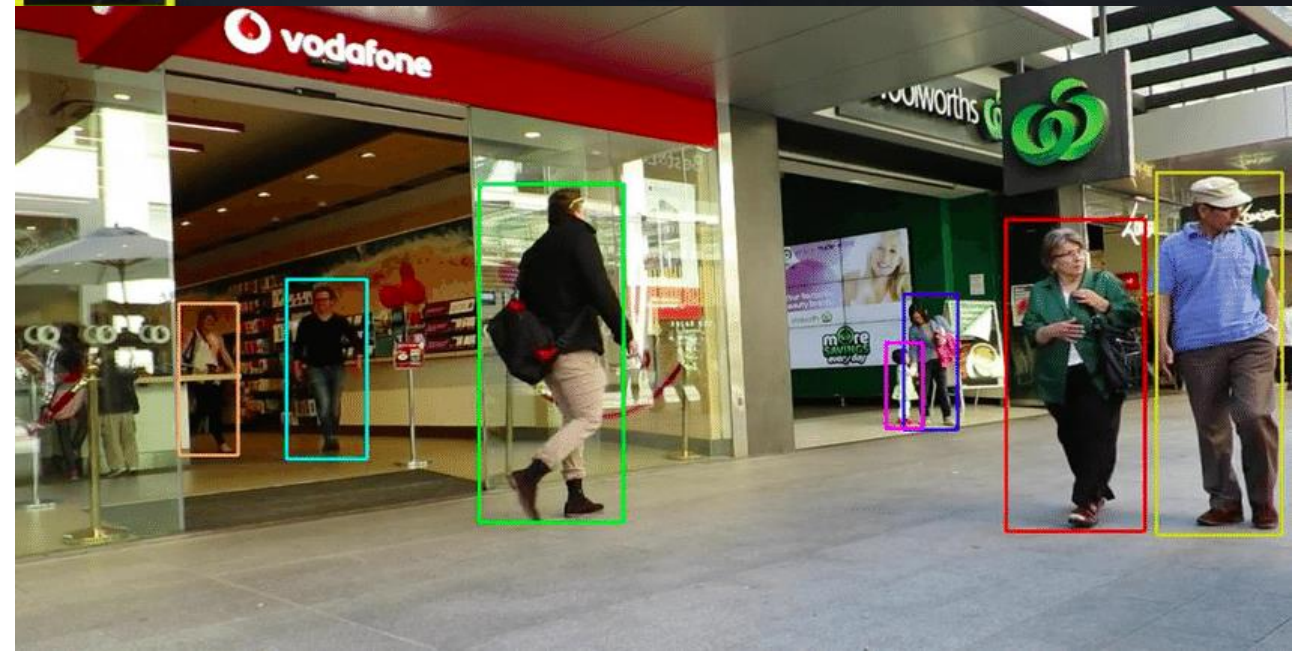
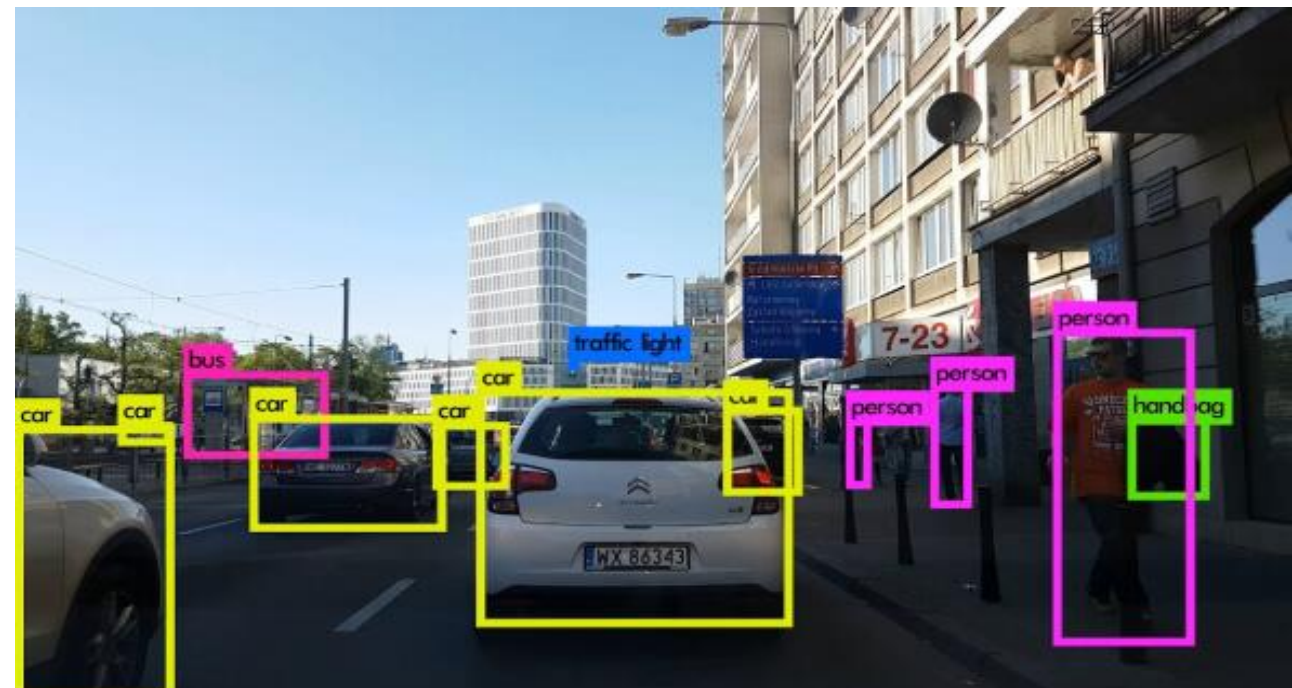
Lấy một ví dụ đơn giản: Việc lắp đặt camera giám sát những nơi đường phố công cộng là hoàn toàn cần thiết để quản lý trật tự xã hội.

Object Tracking là một trong những giải pháp hiệu quả giúp ta theo dõi từng đối tượng trong frame hình từ đó giúp ta dễ dàng phát hiện ngăn chặn , quản lý tội phạm, những hành vi gây mất trật tự xã hội.

Giới thiệu: *Object tracking là gì?*

Theo vết đối tượng (Object Tracking) là bài toán thuộc lĩnh vực thị giác máy tính.

Object tracking là bài toán theo dõi một hoặc nhiều đối tượng chuyển động theo thời gian thực trong một video. Hiểu một cách đơn giản, nó là bài toán cao hơn nhận dạng vật thể, khi đối tượng cần được xử lý không đơn giản là một hình ảnh mà là một chuỗi các hình ảnh, video.



Giới thiệu:

Nhận diện đối tượng thường được sử dụng để làm gì và ứng dụng như thế nào?

Đây là một trong những bài toán quan trọng và phổ biến trong lĩnh vực Machine Vision. Object tracking được ứng dụng rất nhiều và đem lại hiệu quả cao trong các lĩnh vực trong cuộc sống của chúng ta.

Trong kiểm soát an ninh: VD: ỨNG DỤNG KỸ THUẬT THEO DÕI ĐỐI TƯỢNG CHO BÀI TOÁN NHẬN DẠNG HÀNH VI CỦA KHÁCH HÀNG TRONG SIÊU THỊ .

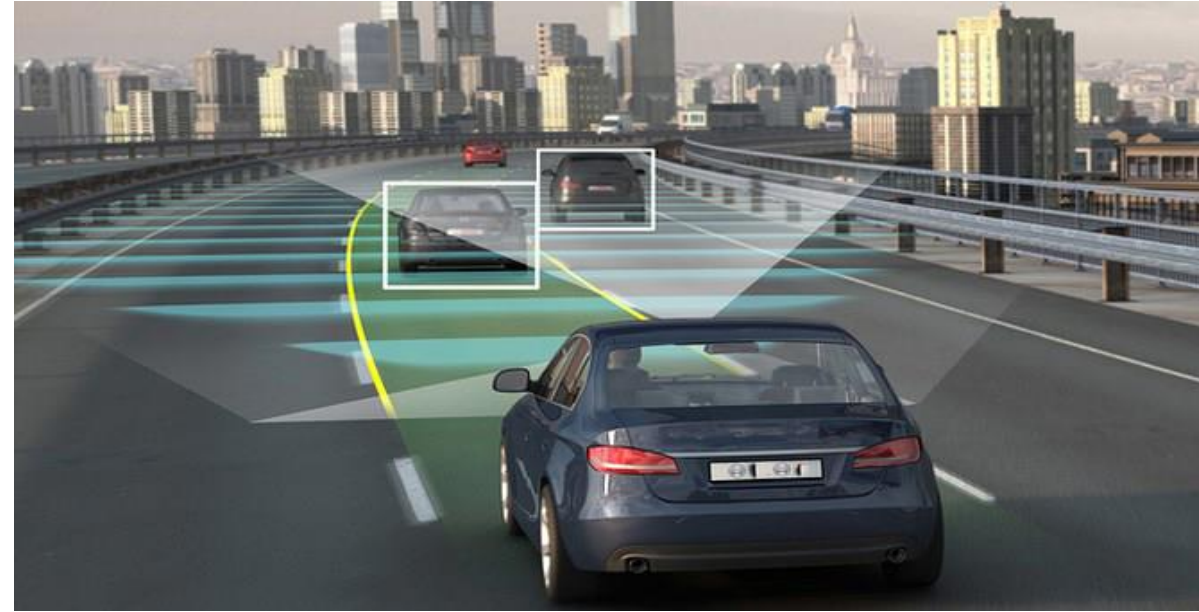
Hệ thống theo dõi đối tượng nhận vào các khung hình video thu nhận từ các camera, qua một số bước xử lý, phân tích và cuối cùng là đưa ra quỹ đạo đường đi của đối tượng theo thời gian làm cơ sở cho việc tiếp theo là nhận biết hành vi. Điều này giúp ta tiết kiệm rất nhiều thời gian và dễ dàng quản lý an ninh những khu đông người



Giới thiệu:

Nhận diện đối tượng thường được sử dụng để làm gì và ứng dụng như thế nào?

Trong xe tự hành: với xe tự hành, hệ thống điều hướng (navigation) cần phải nhận biết được chướng ngại vật (obstacle) trên đường đi. Và nếu đó là những đối tượng di chuyển, robot cần kích hoạt hệ thống theo vết thời gian thực để tránh va chạm.



Face Tracking – Theo dõi khuôn mặt

- Chức năng theo dõi khuôn mặt tức là camera sẽ tự động chụp khung hình để thấy được khuôn mặt của mọi người đang vào khu vực mà camera quan sát được.
- Chỉ tự động chụp hình và lưu lại các khuôn mặt vào thẻ nhớ SD. Ví dụ: các sân bay, khu vực nhạy cảm trong trường hợp muốn ghi nhận kẻ gian để sau này phát hiện người này lại đi qua camera vào lần sau
- Chính vì thế, với chức năng này, hệ thống camera IVS rất hay được sử dụng trong các sân bay hoặc trong các khu vực kiểm soát người ra vào để phòng trường hợp các kẻ xấu đột nhập

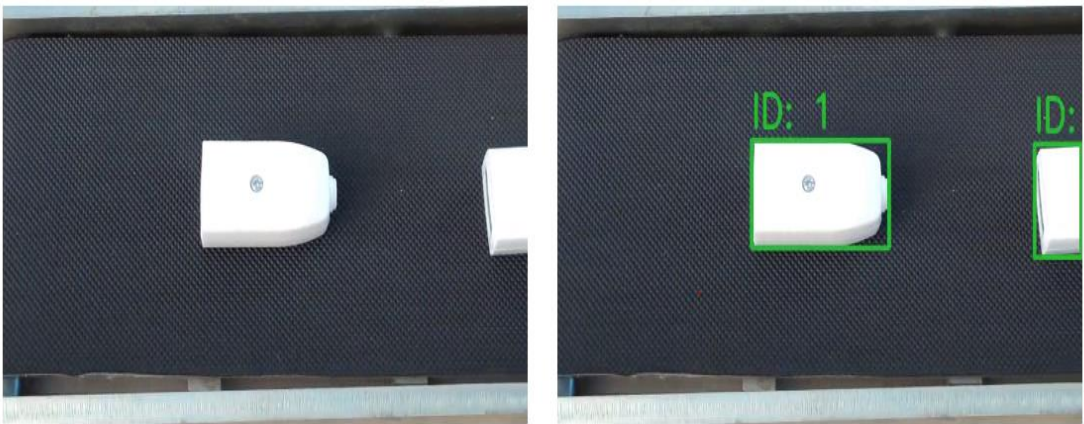


Giới thiệu:

Nhận diện đối tượng thường được sử dụng để làm gì và ứng dụng như thế nào?

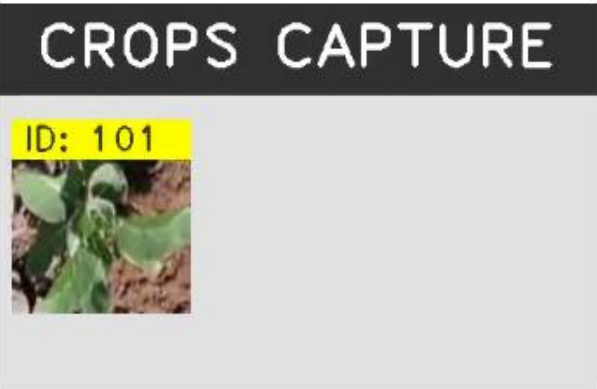
Trong công nghiệp: Theo dõi sản phẩm trên các dây chuyền Sản xuất

Project 2: Conveyor Belt

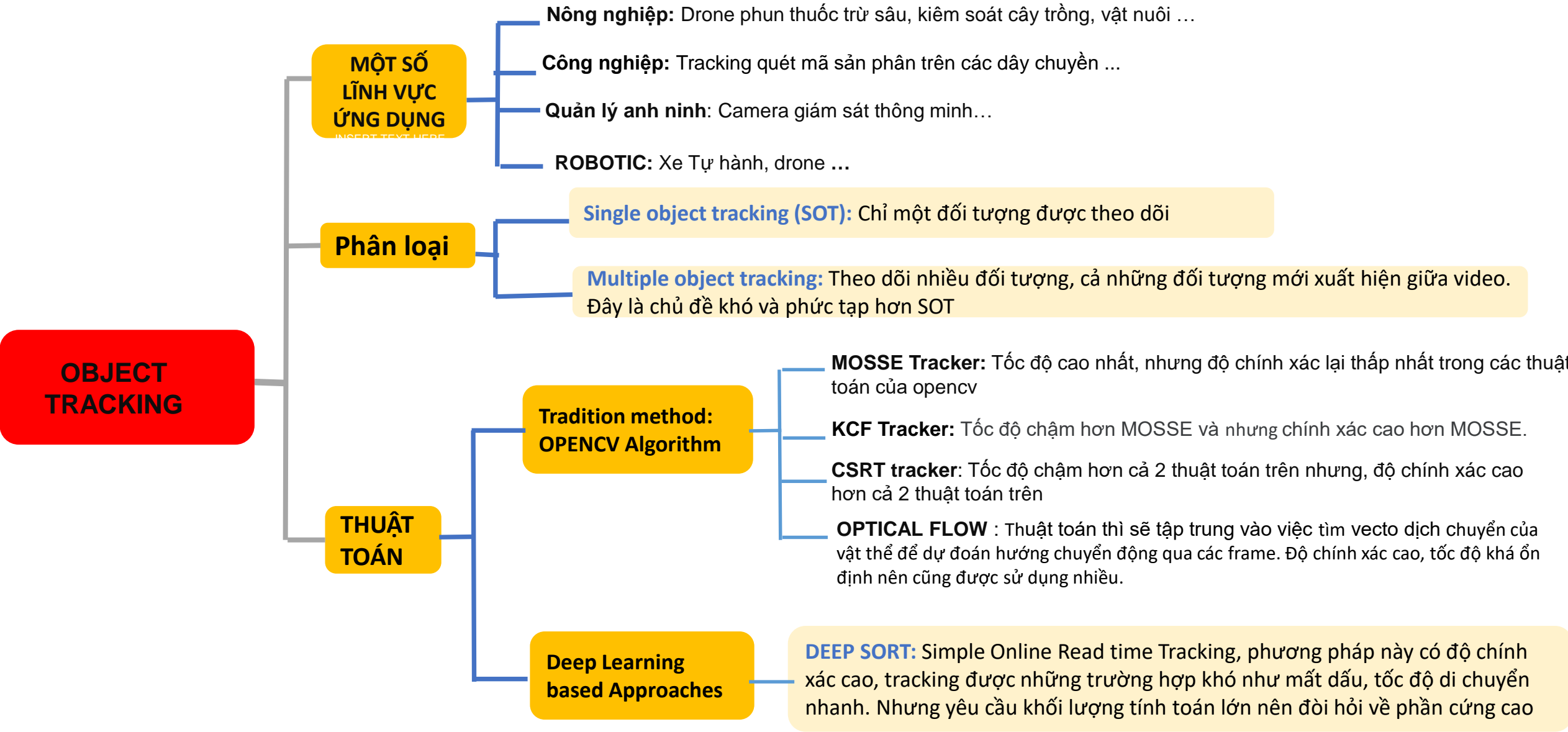


Trong nông nghiệp: Sử dụng drone để quét và đếm nhanh Số lượng cây trồng, và gán ID cho từng cây và lưu trữ lại giúp ta dễ dàng quản lý.

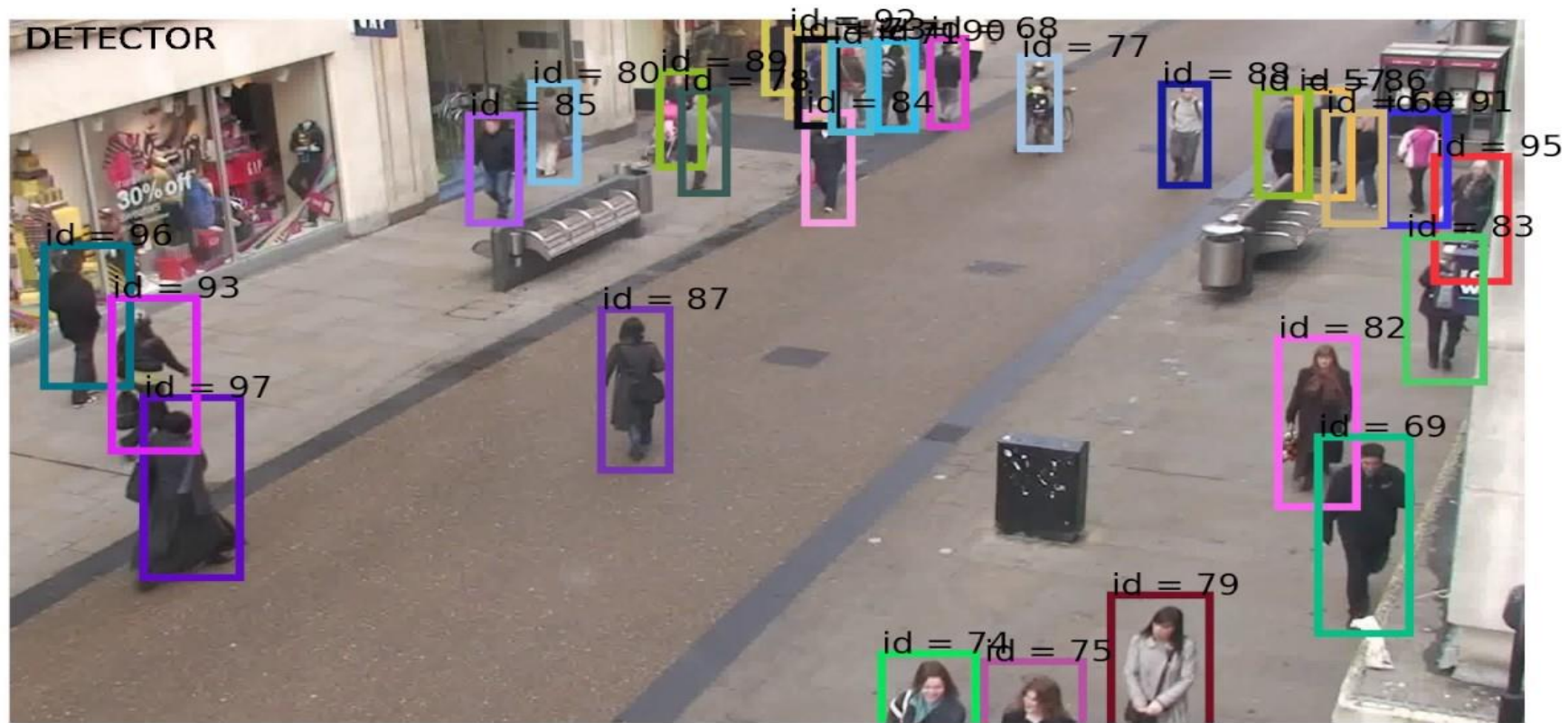
Tracking from Drone Footage



TỔNG QUAN



How to implement tracking multi object?



1.Tracking without detection: Traditional way to track object

- Tracking without Detection:

The type of tracking algorithm where the coordinates of the object are manually initialized and then the object is tracked in further frames. This type is mostly used in traditional computer vision algorithms as discussed earlier.



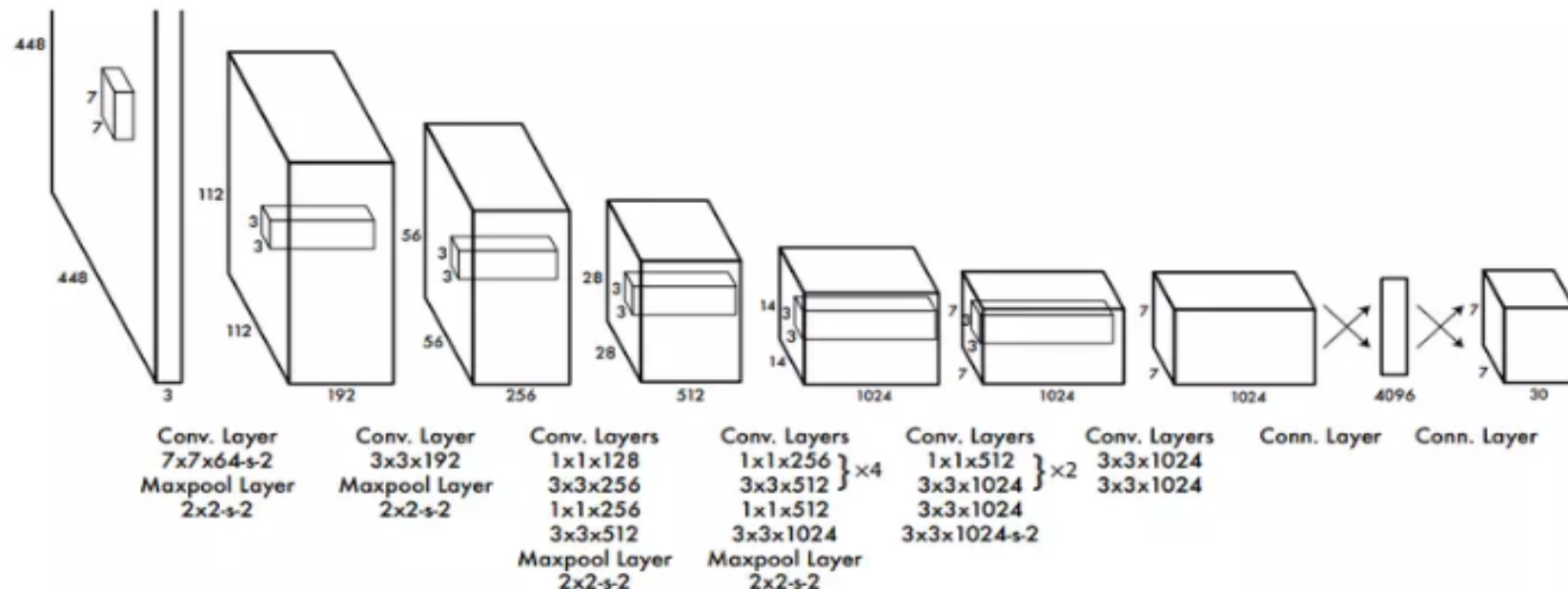
2.Tracking by detection: METHOD: Using YOLOv4, Center point

Yolo Introduction:

Yolo là một mô hình mạng CNN cho việc phát hiện, nhận dạng, phân loại đối tượng.

Yolo được tạo ra từ việc kết hợp giữa các convolutional layers và connected

layers. Trong đó các convolutional layers sẽ trích xuất ra các feature của ảnh, còn full-connected layers sẽ dự đoán ra xác suất đó và tọa độ của đối tượng.

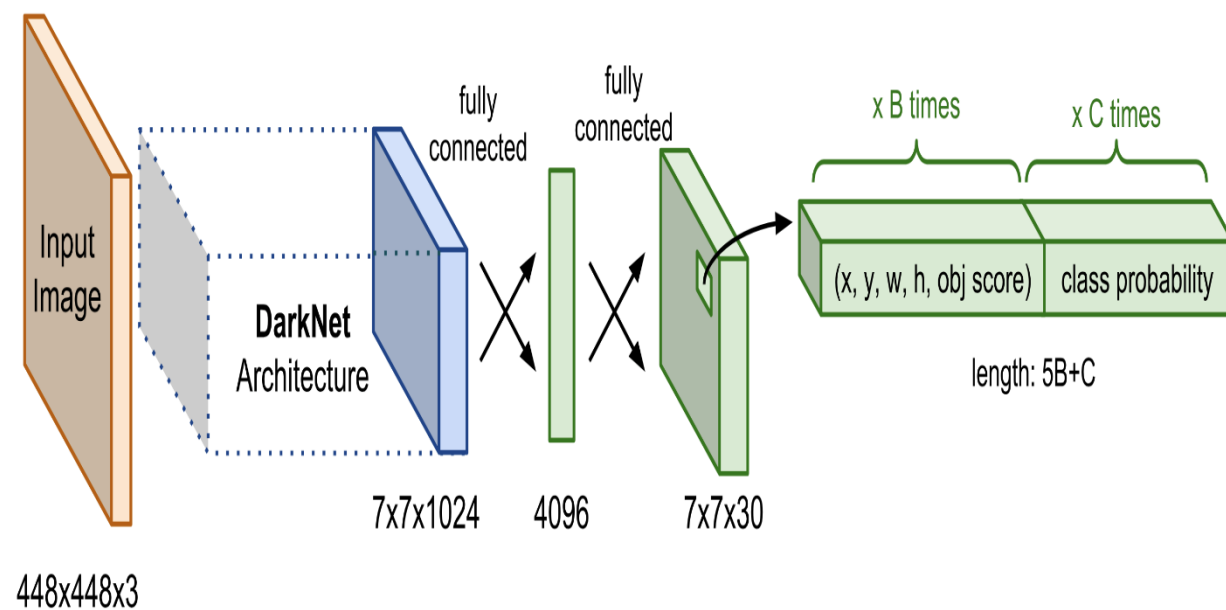


2.Tracking by detection : Using Yolov4,Center point:

Yolo Introduction:

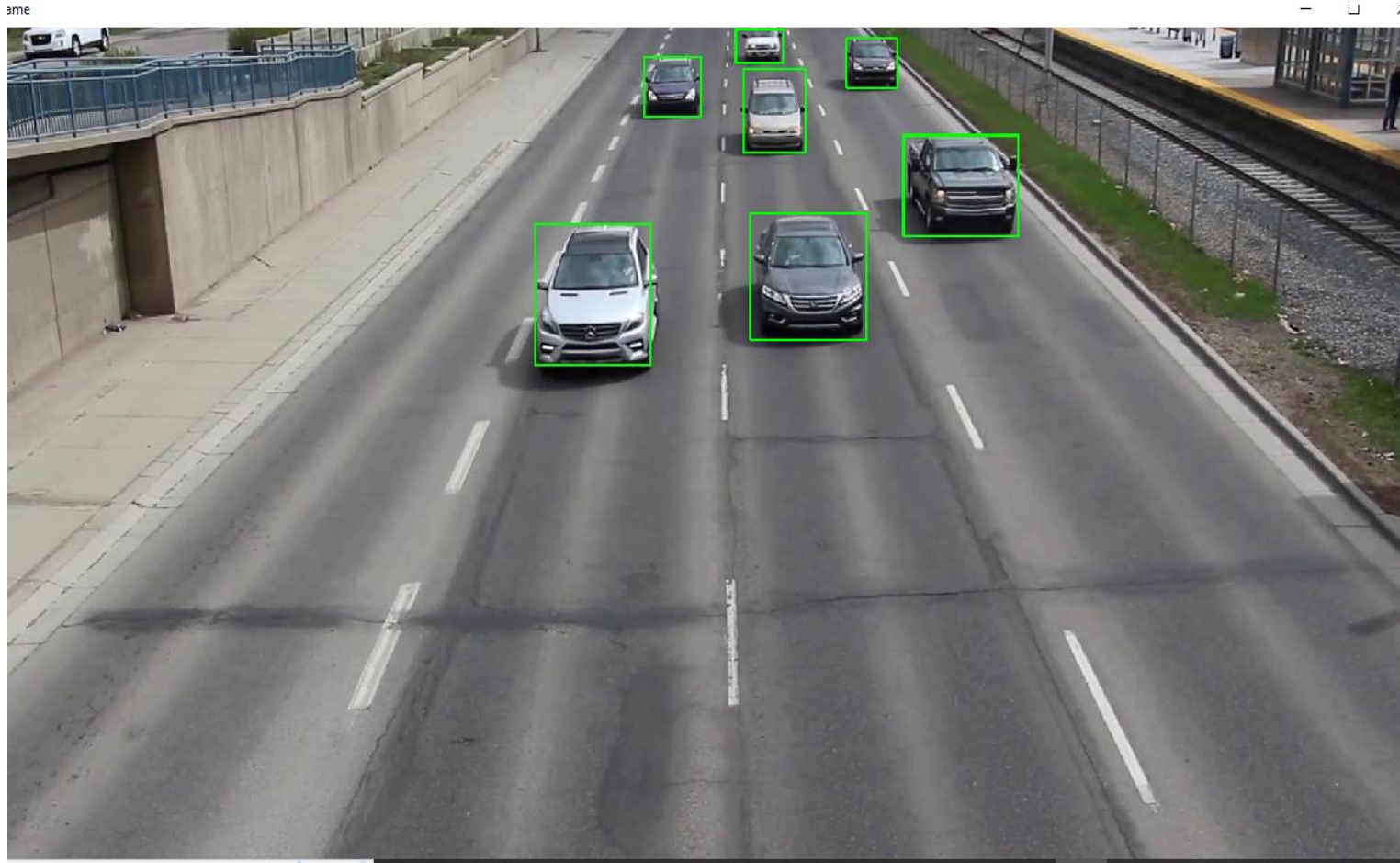
YOLO trong object detection có nghĩa là “You only look once”. Tức là chúng ta chỉ cần nhìn 1 lần là có thể phát hiện ra vật thể.

Về độ chính xác thì YOLO có thể không phải là thuật toán tốt nhất nhưng nó là thuật toán nhanh nhất trong các lớp mô hình object detection. Nó có thể đạt được tốc độ gần như real time mà độ chính xác không quá giảm so với các model thuộc top đầu.



2.Tracking by detection : Using Yolov4,Center point:

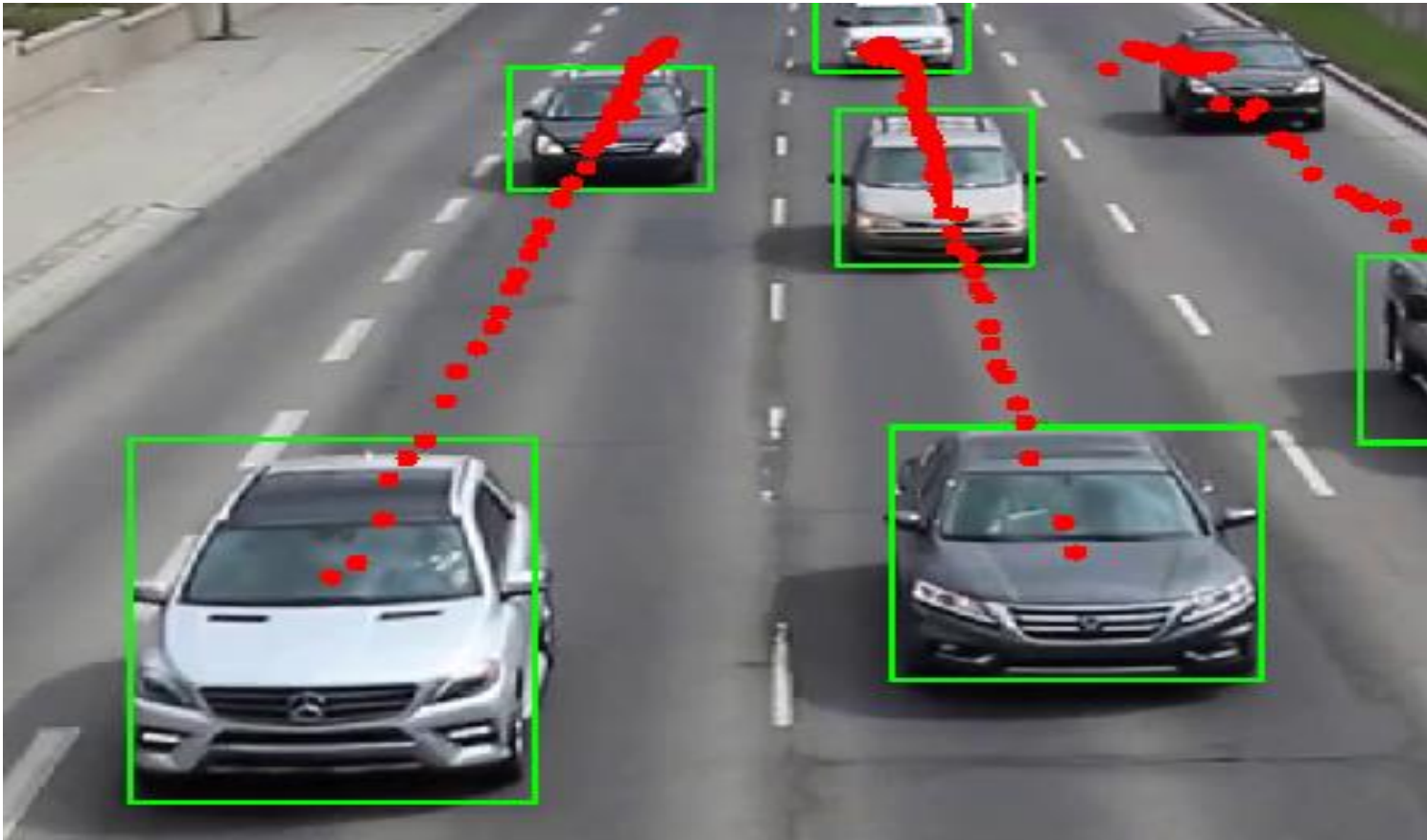
Bước 1: Sử dụng yolov4 model để detect object trong từng frame hình một



2.Tracking by detection : Using Yolov4,Center point:

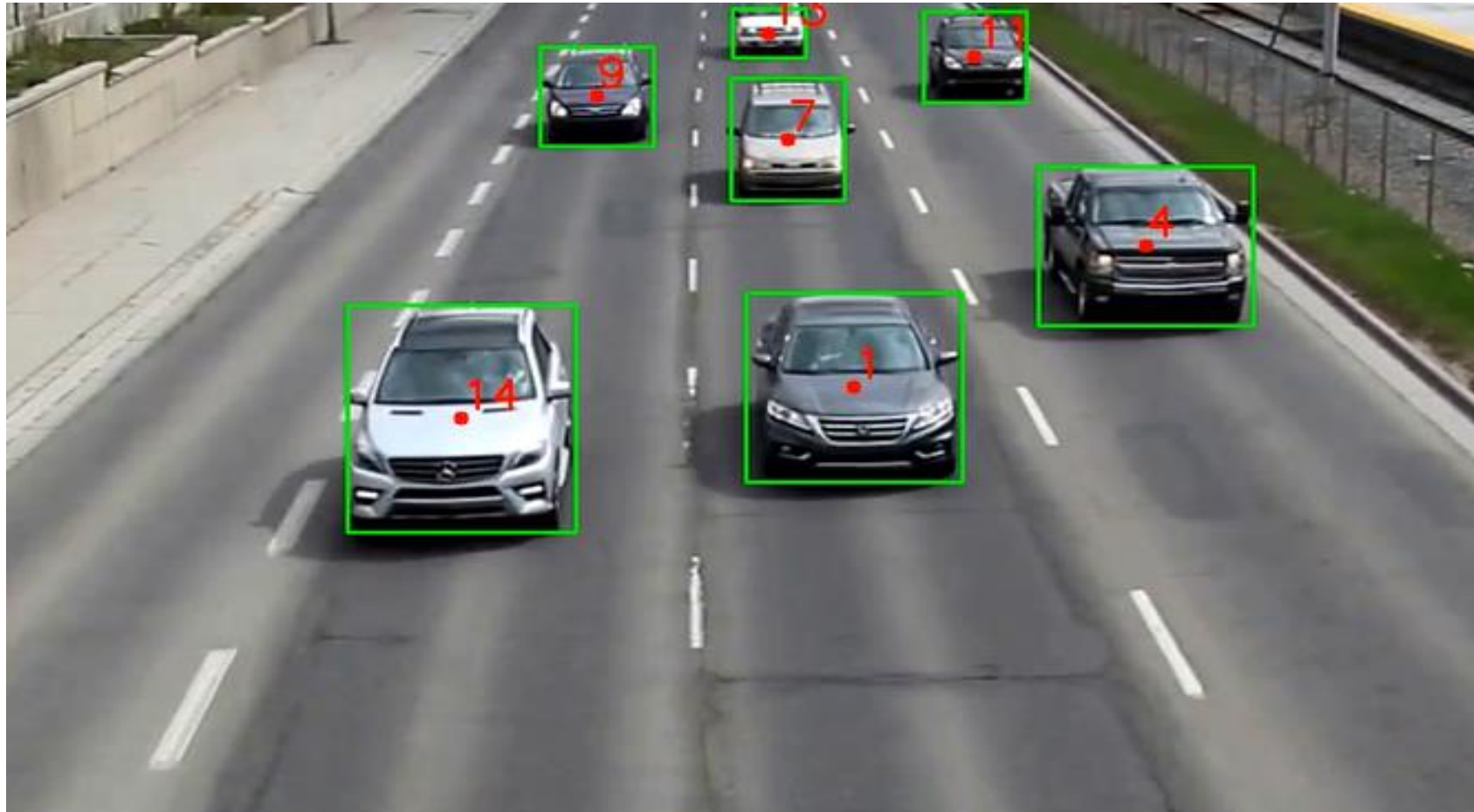
Bước 2: Dựa vào khoảng cách giữa các center point của từng frame để xác định các bounding box là cùng một object.

Các center point của các bounding box gần nhau thì các bounding box cùng 1 object



2.Tracking by detection : Using Yolov4,Center point:

Bước 3: Gán ID cho từng object qua từng frame 1



2.Tracking by detection : Using Yolov4, Deep sort

DeepSORT

Để hiểu thế nào là DeepSort thì chúng ta sẽ tìm hiểu SORT - Simple Online Real-time Tracking là gì?

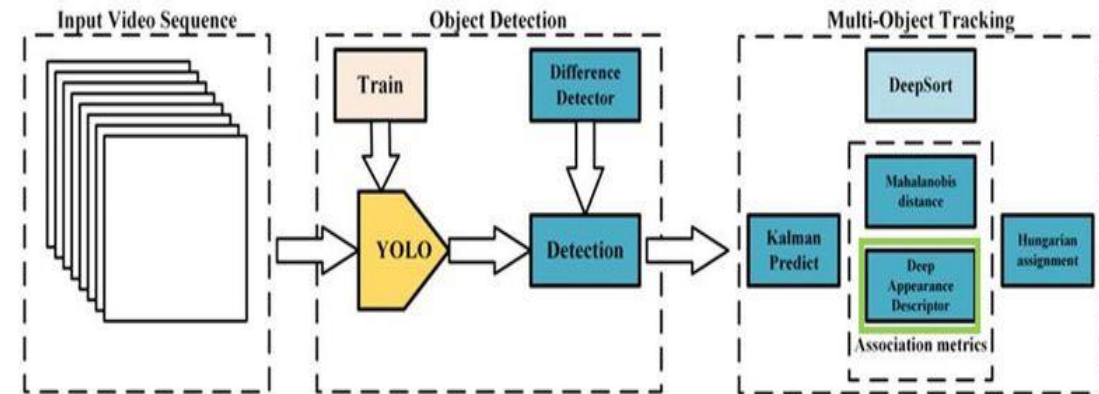
Và như cái tên của nó, SORT được sinh ra để có thể tracking đối tượng trên thời gian thực dựa trên bốn yếu tố:

- Phát hiện đối tượng (Detection)
- Đánh giá và dự đoán (Estimation)
- Liên kết các đối tượng với nhau (Association)
- Theo dõi và hủy bỏ danh tính đối tượng (Track Identity creation and destruction)

Về cơ bản SORT rất đơn giản và dễ implement. Nhưng nó tồn tại điểm yếu chết người về vấn đề ID Switches và dự đoán đối tượng dựa trên Kalman Filter . Để giải quyết vấn đề đó thì DeepSORT được tạo ra.

DeepSORT

Where is the Deep Learning in all of this?



2.Tracking by detection : Using Yolov4, Deep sort

Luồng xử lý của Deep SORT được thực hiện tuần tự qua các bước dưới đây :

- {**Bước 1**}: Sử dụng Faster Region CNN (với backbone là VGG16) để phát hiện các đối tượng trong khung hình hiện tại.
- {**Bước 2**}: Deep SORT sử dụng Kalman Filter để dự đoán các trạng thái track mới dựa trên các track trong quá khứ. Các trạng thái này lúc mới khởi tạo sẽ được gán 1 giá trị mang tính thăm dò (tentative). Giá trị này nếu vẫn đảm bảo duy trì được trong 3 frame tiếp theo, trạng thái sẽ chuyển từ thăm dò sang xác nhận (confirmed), và sẽ cố gắng được duy trì theo dõi trong 30 frame tiếp theo. Ngược lại, nếu mất dấu khi chưa đủ 3 frame, trạng thái sẽ bị xóa khỏi trình theo dõi.
- {**Bước 3**}: Sử dụng những track đã được xác nhận, tiến hành đưa vào chiến lược đối sánh phân tầng (matching cascade) nhằm liên kết với các detection phát hiện được dựa trên độ đo về khoảng cách và đặc trưng.
- {**Bước 4**}: Các track và các detection chưa được liên kết sẽ được đưa đến 1 lớp lọc tiếp theo. Sử dụng giải thuật Hungary giải bài toán phân công với ma trận chi phí IOU để liên kết lần 2
- {**Bước 5**}: Xử lý, phân loại các detection và các track
- {**Bước 6**}: Sử dụng Kalman filter để hiệu chỉnh lại giá trị của track từ những detection đã được liên kết với track và khởi tạo các track mới.

DeepSORT

Where is the Deep Learning in all of this?

