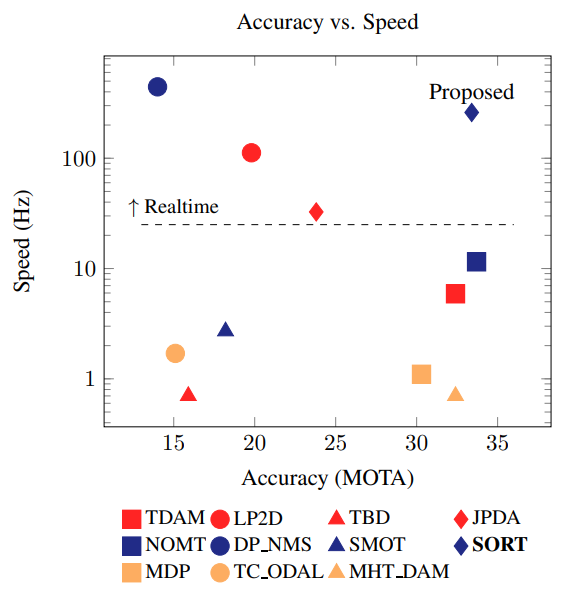
Nghiên cứu tuần 1 SCS

Đặng Anh Quân – 20203909

# SORT (Simple online realtime traking):

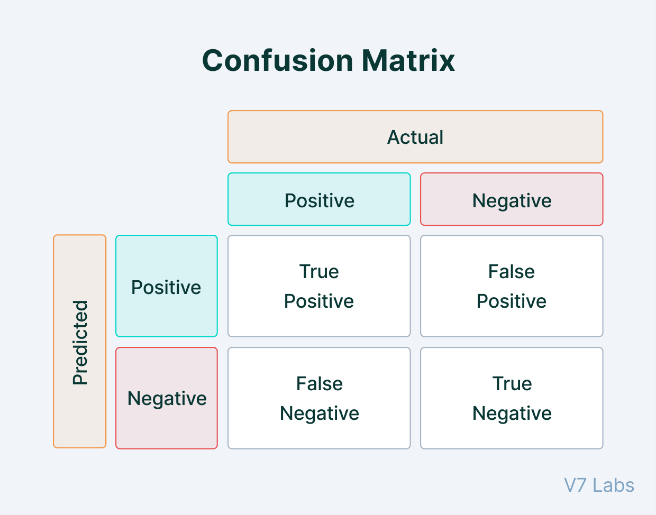
* Link paper: https://arxiv.org/pdf/1602.00763.pdf
* Online tracking where only detections from the previous and the current frame are presented to the tracker.

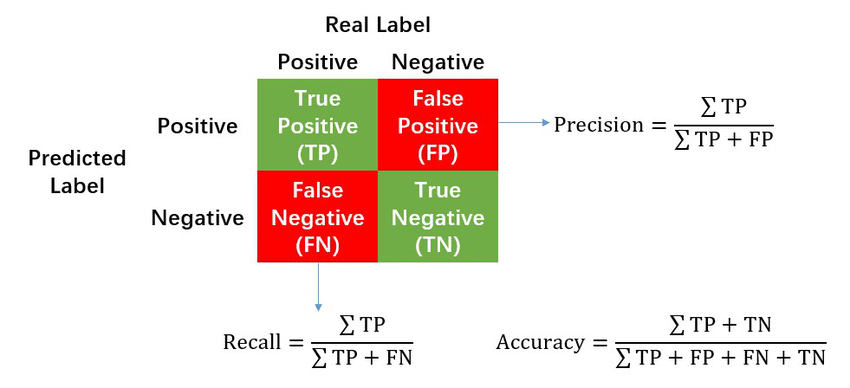


* Phương pháp mà SORT sử dụng gồm các yếu tố chính:
  + Detection (xác định)
  + Propagating object states into future frames/ Estimation (truyền trạng thái vật thể vào trong frames tiếp theo)
  + Associating current detections with existing objects/ Data association(liên kết giữa vật xác định và vật hiện tại)
  + Managing the lifespan of tracked objects/ Creation and deletion of track identities (quản lý dòng đời của vật thể được theo dõi)

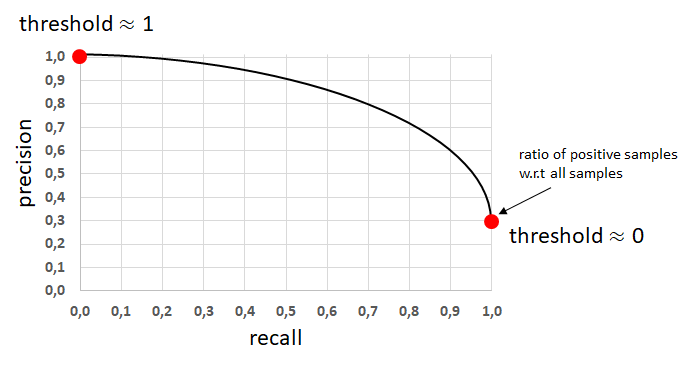
## Detection (xác đinh):

* Xác định vật thể được theo dõi. Trong project này YOLOv8 được sử dụng là bộ xác định (detector).
  + YOLOv8 được ra mắt vào 10/01/2023 là một YOLO model mới nhất có thể sử dụng cho xác định vật thể (object detection), phân loại ảnh (image classification), và một số nhiệm vụ phân đoạn(segmentation task). Được phát triển bởi Ultralytics, có nhiều phát triển về kiến trúc cũng như trải nghiệm của lập trình viên so với YOLOv5
  + Một số lý do sử dụng YOLOv8:
    - Có chỉ số chính xác cao đo bởi COCO và Roboflow100. Model trung bình đạt 50.2% mAP khi đo trên COCO.
      * Cách tính mAP:
        + Confusion Matrix,
        + Intersection over Union(IoU),
        + Recall,
        + Precision





A graph of a graph

Description automatically generated- A graph of average precision

Description automatically generated

* + - * + mAP: là giá trị trung bình của AP trong các class khác nhau
    - Có hỗ trợ nhiều tính năng cho lập trình, từ dễ sử dụng như CLI đến, các package phức tạp Python.
    - Có cộng đồng lớn nên được hỗ trợ nhiều.

A screenshot of a computer

Description automatically generated

## Estimation (estimation model):

* Truyền danh tính của mục tiêu vào frame tiếp theo. Ước lượng vị trí trong nhiều frame của từng vật thể bằng model linear constant velocity model độc lập với các vật thể khác và với cả chuyển động của camera.
* Mỗi trạng thái của mục tiêu được mô hình hóa bởi:

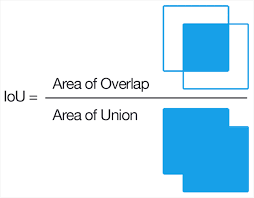
A group of black letters

Description automatically generated

* Trong đó, u và v biểu diễn ngang và dọc của pixel chính giữa mục tiêu, s và r biểu diễn diện tích(scale), tỉ lệ bounding box. Khi mục tiêu được xác định, bounding box được sử dụng để cập nhật trạng thái của mục tiêu (trong đó thành phần velocity được ước lượng qua Kalman filter). Nếu trong trường hợp mục tiêu không được xác định, trạng thái của nó được xác định đơn giản mà không có chính xác thông qua linear velocity model.

## Data Association:

* Mục tiêu chính: Gán detections vào mục tiêu có sẵn, mỗi bounding box của mục tiêu được dự đoán bởi vị trí mới của mục tiêu trong frame hiện tại.



* Ma trận chi phí (cost matrix) được tính toán thông qua IOU (intersection-over-union) – một dạng khoảng cách, IOU được tính giữa mỗi dự đoán và tất cả các bounding box của mục tiêu.
* Sau đó, việc gán được giải quyết thông qua Hungarian algorithm, Thêm vào đó, IOU tối thiệu được áp dụng để từ chối việc gán khi mà vùng dự đoán và vùng của mục tiêu bé hơn IOU min.

## Creation and Deletion of track Identities:

* Giải thích cơ bản: Khi mà vật thể vào và ra khỏi khung ảnh thì ID của chúng cần phải được tạo và xóa bỏ dựa trên IOUmin. Track được xóa bỏ nếu chúng ko được xác định trong 1 số frame.
* Quy trình tạo 1 IDs:
  + Để tạo 1 theo dõi, xét đến tất cả các xác định (detection) có IOU nhỏ hơn IOUmin.
  + 1 theo dõi được khảo tạo sử dụng hình dạng của bounding box với velocity đặt về 0. Do velocity chưa có nên học đặt covariance của velocity giá trị lớn, phản ảnh sự không chắc chắn.
  + Sau đó, theo dõi mới trải qua giai đoạn thử nghiệm, khi mà mục tiêu cần được liên kết với xác định để đưa ra đủ bằng chứng (nhằm ngăn chặn FP).
* Quy trình xóa 1 IDs:
  + Theo dõi bị hủy bỏ nếu không được xác định trong Tlost Frame. (thường được set là 1)

# DeepSORT:

* Link paper: <https://arxiv.org/pdf/1703.07402.pdf>
* SORT có khả năng rất tốt trong độ chính xác và độ tụ (precision). Tuy nhiên SORT trả lại 1 số nhiều các ID và thất bại trong việc bị che lấp. Điều này là do SORT sử dụng ma trận liên kết (association matrix).
* DeepSORT sử dụng ma trận liên kết tốt hơn kết hợp cả chuyển động và mô tả xuất hiện. Có thể được mô tả là thuật toán theo dõi vật thể dựa trên cả chuyển động và mô tả xuất hiên.
* Do đó có thể chia làm 2 branch framework:
  + Appearance branch:
    - Bộ mô tả xuất hiện sâu (deep appearance descriptor) được áp dụng để trích xuất đặc trưng xuất hiện (appearance features) từ detections.
    - Có cơ chế ngân hàng đặc trưng (feature bank) dùng để lưu trữ đặc trưng của 100 frames.
    - Khi mà detections mới, khoảng cách cosine giữa ngâng hàng đặc trưng và đặc trưng của detections được tính toán. Khoảng cách được sử dụng để làm matching cost trong giai đoạn liên kết.
  + Motion bracnh:
    - Thuật toán Kalman filter được sử dụng để dự đoán vị trí của quỹ đạo trong khung hiện tại. Hoạt động trên 2 giai đoạn:
      * State-prediction:

A white background with black text

Description automatically generated

* + - * State\_update:

A white paper with black text

Description automatically generated

* + - Mahalanobis distance được sử dụng để đo sự khác biệt về không gian thời gian giữa trạng thái chuyển động của quỹ đạo và detections mới.
    - Cuối cùng, thuậ toán matching cascade được sử dụng để giải quyết association task. Thuật toán này giúp các vật thể được thấy thường xuyên có khả năng được matching nhiều hơn.

# Kalman Filter algorithm:

A screenshot of a video game

Description automatically generated

* Kalman Filter algorithm là thuật toán dùng để ước lượng trạng thái/ vị trí của hệ thống sử dụng:
  + Quan sát quá khử (past observations).
  + Đo đạc hiện tại (current measurements).

A group of white letters

Description automatically generated

* 2 bước của Kalman Filter:
  + Prediction step:
    - prior estimate -> idea version (prediction)
    - Noise covariance matrix -> ?? cần chú ý

A black circle with white lines

Description automatically generated

* + Update Step:
    - Lấy kết quả từ Prediction Step kết hợp trạng thái đo đạc hiện tại của hệ thống và cùng 1 estimate
    - Cần tính toán Kalman Ratio:

A purple background with yellow letters

Description automatically generated

* + - * Kết quả giữa 0 và 1: sigma xx là độ không chắc chắn về dự đoán và Qxx là độ không chắc chắn về đo đạc hiện tại.
        + Nếu bằng 0 thì rất chắc chắn với dự đoán.
        + Nếu bằng 1 thì chắc chắn là đo đạc hiện tại.

A purple background with white text

Description automatically generated

* + - Update giá trị dự đoán dựa trên Kalman ratio, surprise term deltaX và estimate results

A black circle with white arrows and white text

Description automatically generated