



Multiple Object Tracking Systems

Javier Berneche

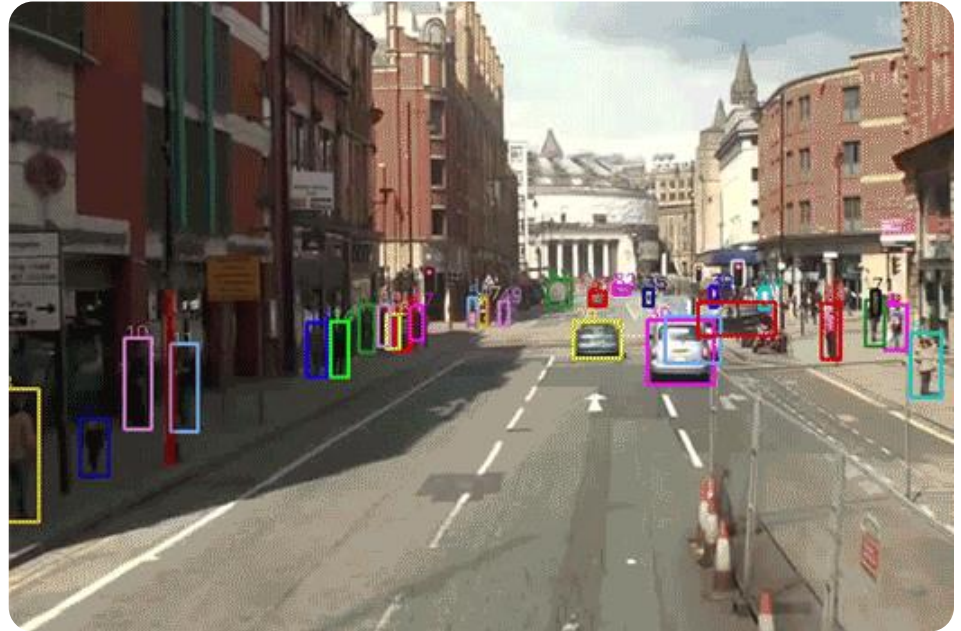
Senior Machine Learning Engineer

 tryo.labs

1. Intro to Multiple Object Tracking (MOT)
2. Building blocks
3. Challenges
4. Evaluation and promising research

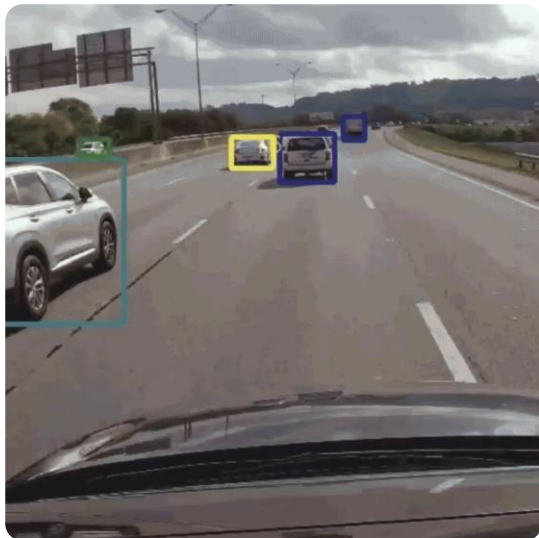
Definition

Multiple Object Tracking (MOT) is the problem of identifying **multiple objects** in a video or live feed and representing them as a set of **trajectories**



Video: MOT Challenge

Applications



Autonomous navigation systems

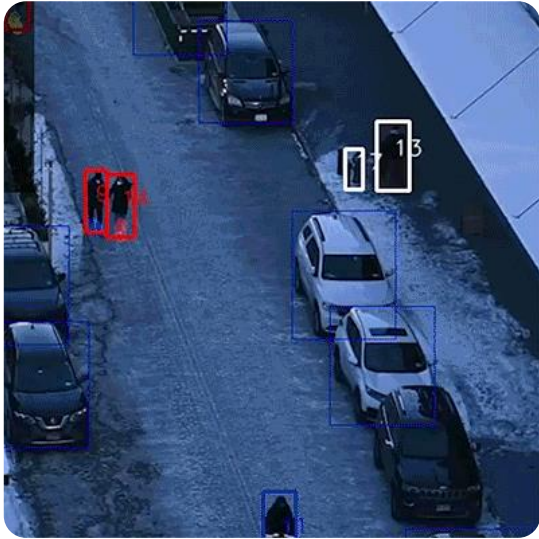


Analyze and monitor congestions



Augmented reality

Applications



Surveillance

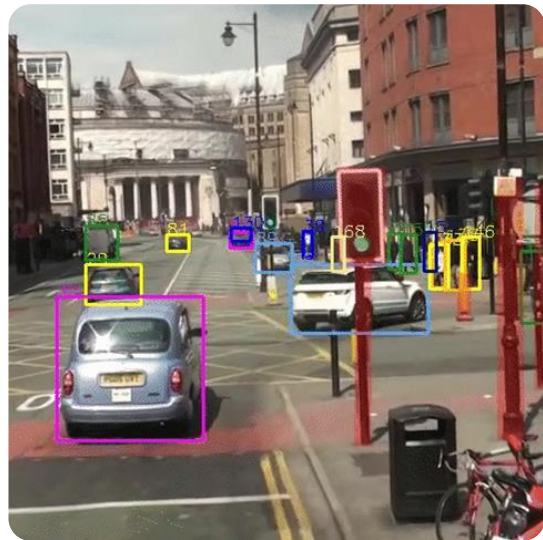


Crowd analysis

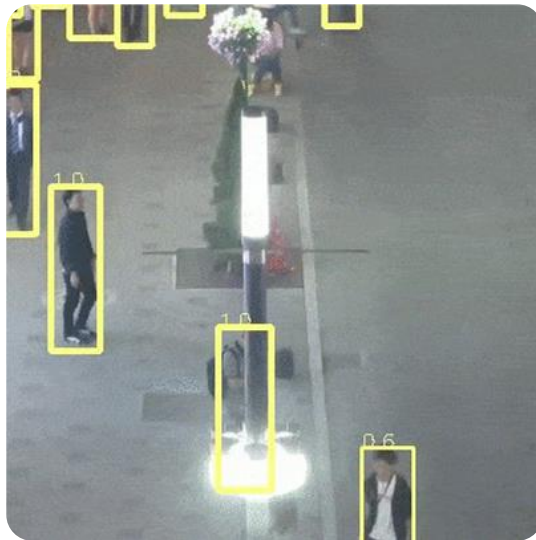


Sports analytics

Challenges



Changes in appearance



Occlusions

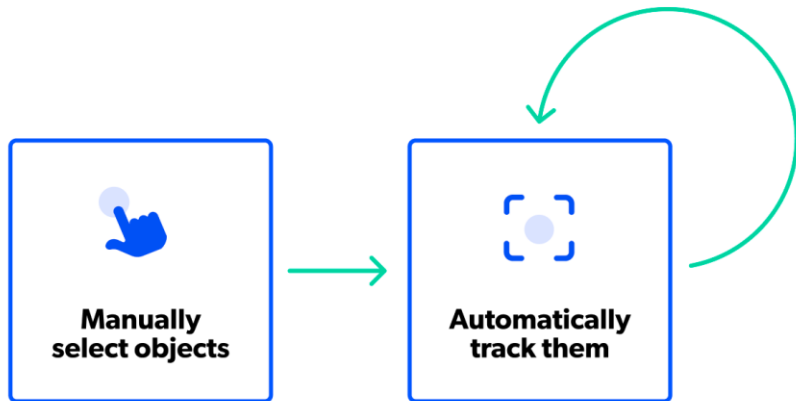


Crowded scenes

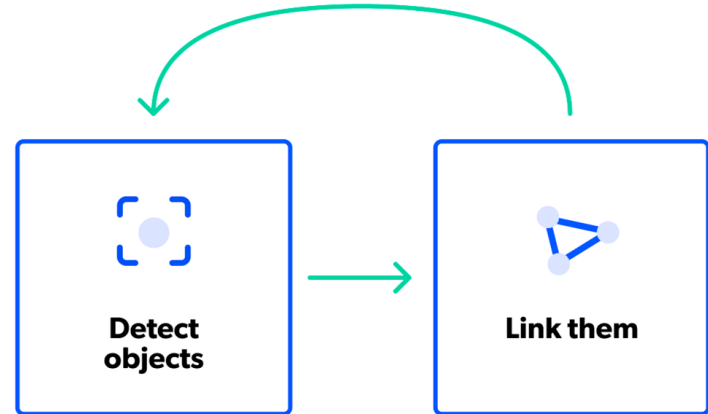
Building Blocks

Building Blocks: Initialization

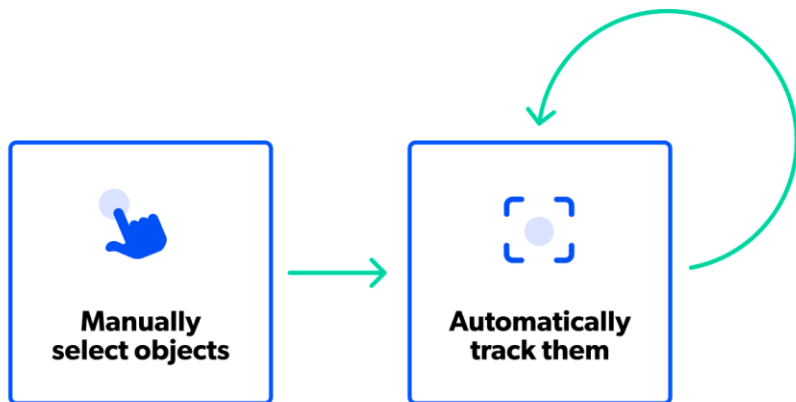
Detection Free (DFT)



Detection Based (DBT)



Detection Free (DFT)



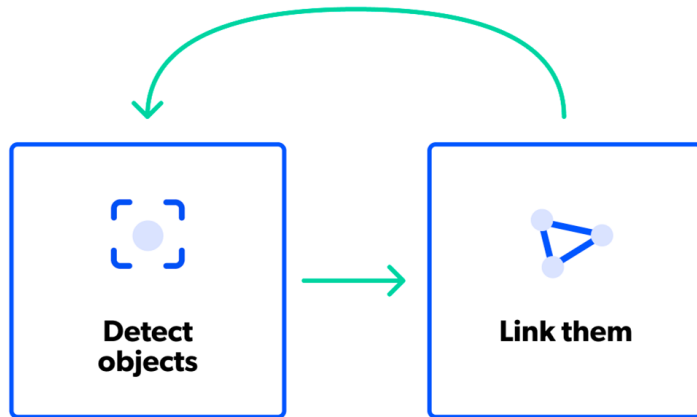
- ✓ Automatically track **manually** selected objects
- ✓ Can track **any** type of object
- ✗ **Can't** handle new objects coming into the scene

✓ Detect objects in **each** frame

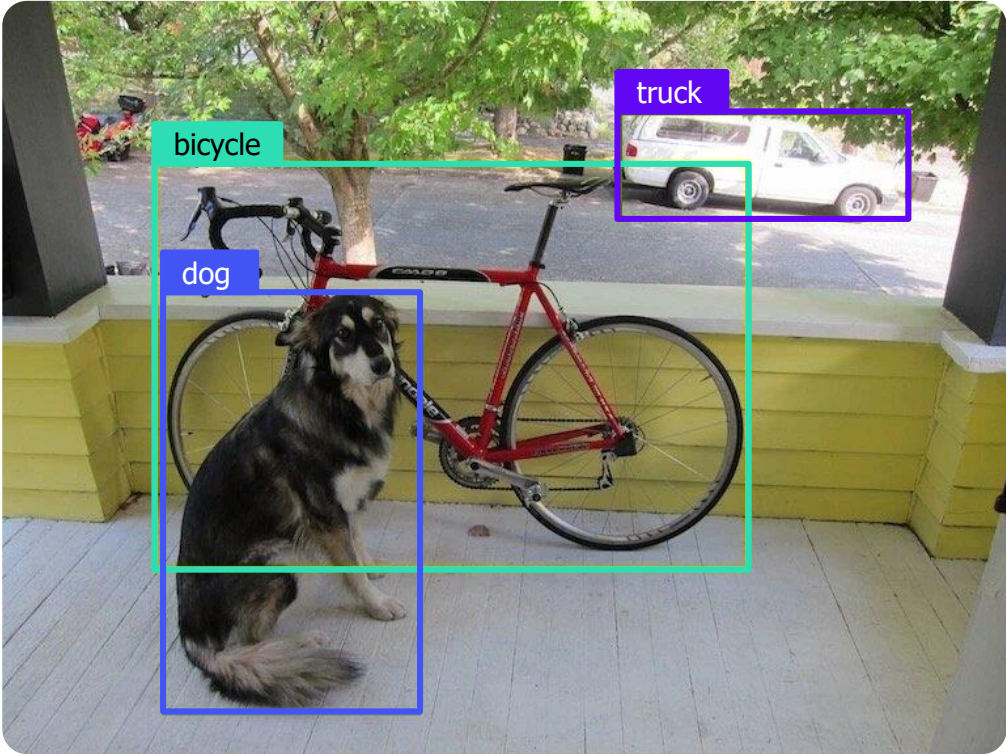
✓ New objects are **automatically** discovered

✗ Object classes need to be **predefined**

Detection Based (DBT)



Building Blocks: Detector



Building Blocks: Processing

Offline



00:15:12

Online



● LIVE

Offline



00:15:12

- ✓ Information from **future** frames can be used
- ✓ **Interpolate** trajectories
- ✓ Keep different **hypotheses**
- ✓ Can be more **accurate**

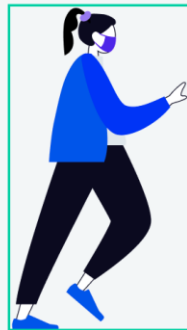
✓ Only **past** information is used

✓ **Extrapolate** trajectories

✓ **Decide** on every frame

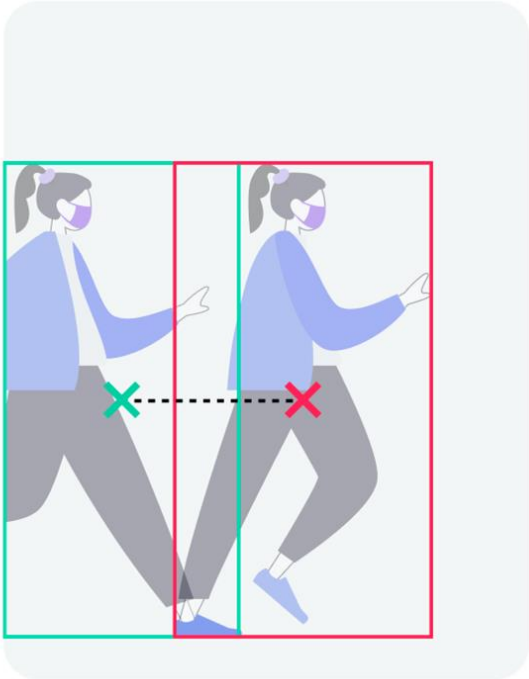
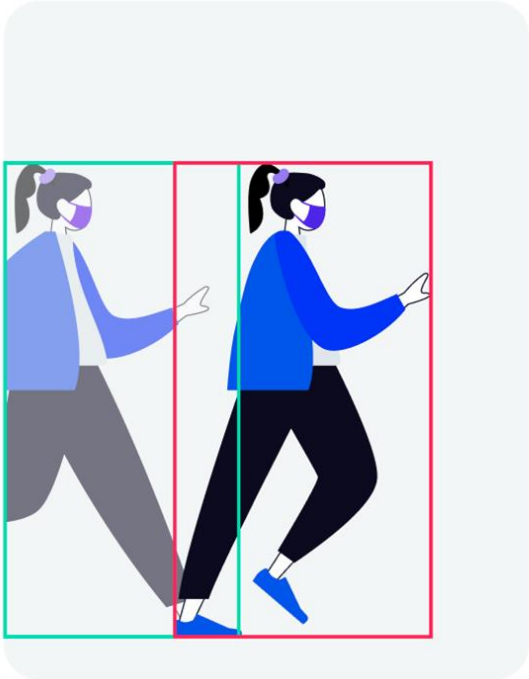
✓ Can be used in **real-time** applications

Online



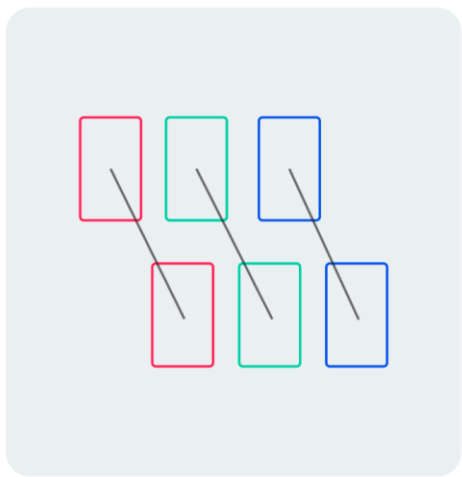
● LIVE

Building Blocks: Positional Cues

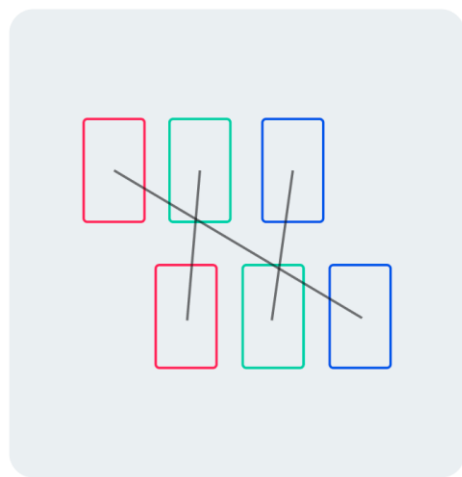


Positional Cues: Assignment

Hungarian

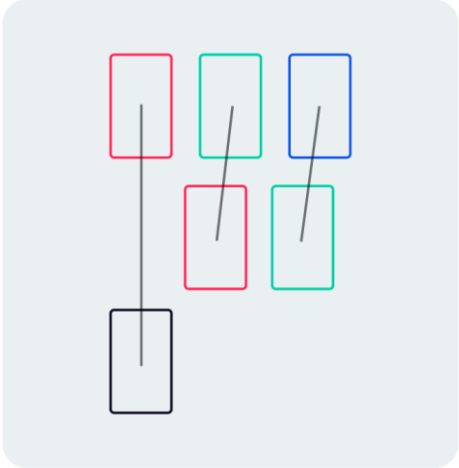


Greedy

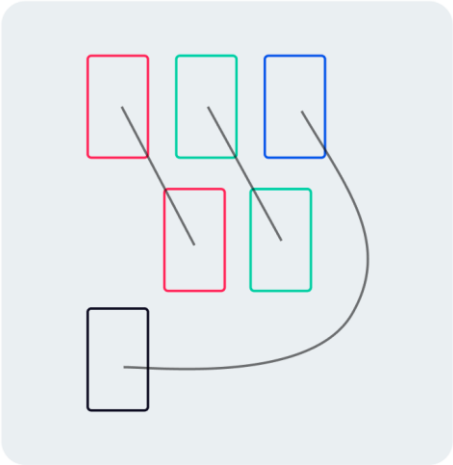


Positional Cues: Assignment

Hungarian



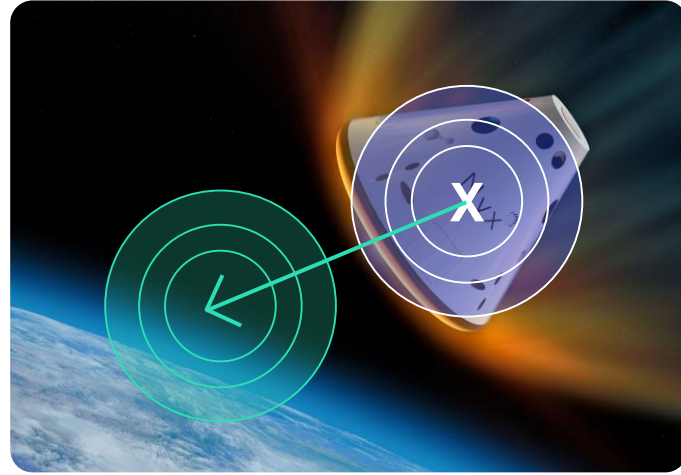
Greedy



Positional Cues: Kalman Filter

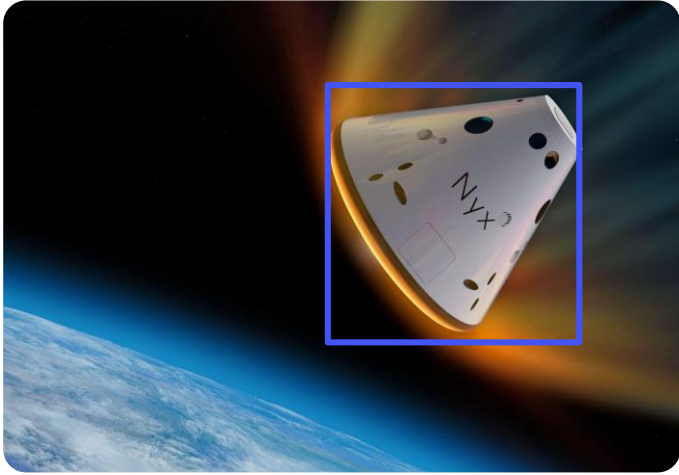


**Sensor that
measures position**



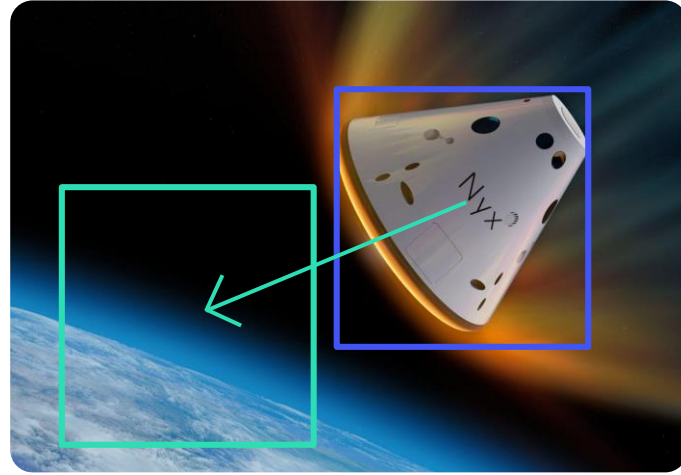
**Model that predicts
movement**

Positional Cues: Kalman Filter



Object detector

YOLO



Constant velocity

$a=0$



Important for **recovering** from occlusions and collisions



Helps with **long-term** tracking



Generally used to **complement** the positional cues



Classical approaches can be used



Usually deep-learning **embeddings** are used



Cosine distance is usually used to compare the embeddings

Visual cues: History



Decide how to represent an object's embedding considering all **past embeddings**



Rolling **averages**



Clustering to maintaining different versions of the object



Memory usage and computational **cost** of comparison

A complete tracking system



Object
Detector



Positional cues
for short-term
tracking



Visual cues for
long-term
tracking

Challenges

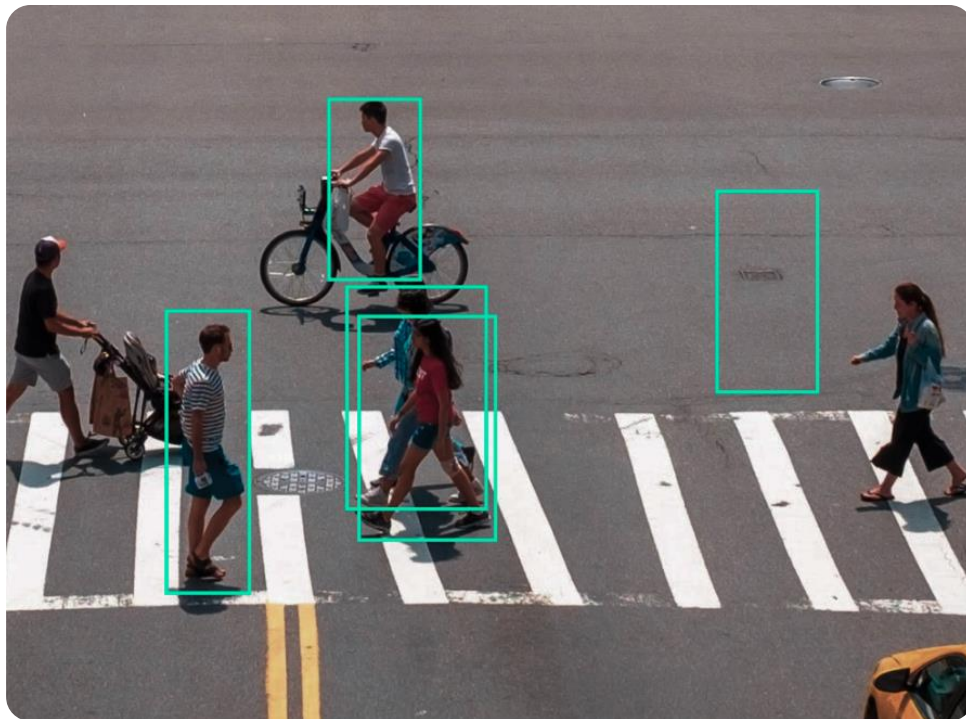
Challenges: Movement



Erratic movement of the objects

Camera movement

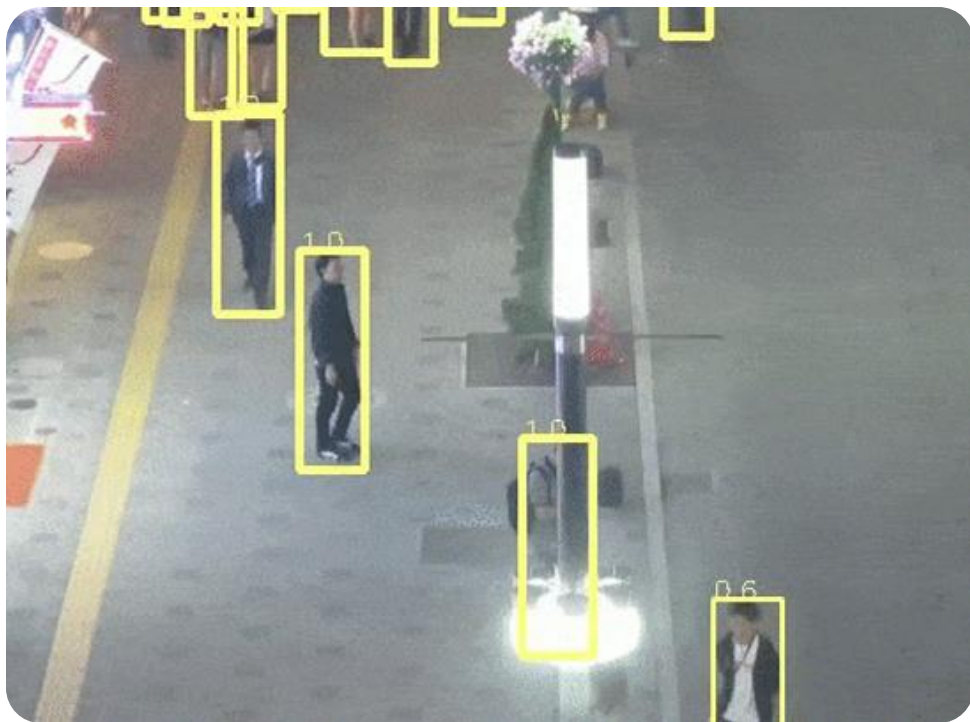
Challenges: Detection quality



False positives

False negatives

Challenges: Occlusions



Causes more False Negatives

Positional tracking can fall apart

Embeddings of partially occluded objects can be bad

Challenges: Embeddings

Object detectors usually do not yield good **embeddings**

Need to add a **second model** for embeddings

Partial occlusions

No obvious model to start with

Evaluation & Research



Improvements to embeddings



Models that combine ReID and Detection



One-Stage models



More Datasets

MOT Challenge

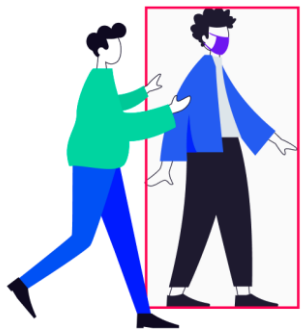
Go-to benchmark for MOT

More datasets are added periodically

Few classes

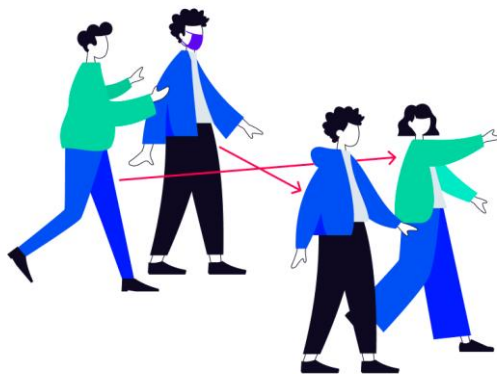
Videos are short

Detection error



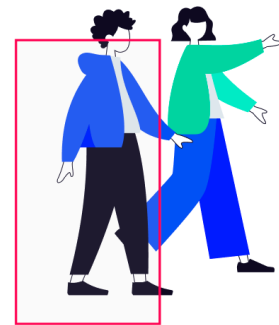
Multi-Object Tracking Accuracy (MOTA)

Association error



Identity F1 (IDF1)

Localization error



High Order Tracking Accuracy (HOTA)

Conclusions

MOT has a huge variety of **applications**

The problem is **challenging**

Solutions involve a number of **components**

Lots of promising **research**

Trackers

- [ByteTrack](#)
- [Norfair](#)
- [SORT](#)
- [DeepSORT](#)

Tools

- [MOTMetrics](#)
- [YOLO](#)
- [OpenMMLab](#)

- [MOTChallenge](#)
- [Luo et al. 2021 Literature review](#)
- [Laura Leal-Taixe](#)
- [Object detection](#)
- [Hungarian method](#)
- [Greedy matching](#)
- [Embeddings](#)
- [Kalman filter](#)

Thank you!

jberneche@tryolabs.com