ETH zürich VIS Visual Intelligence and Systems

Tracking

CV HS 2021

Fisher Yu

What's Tracking

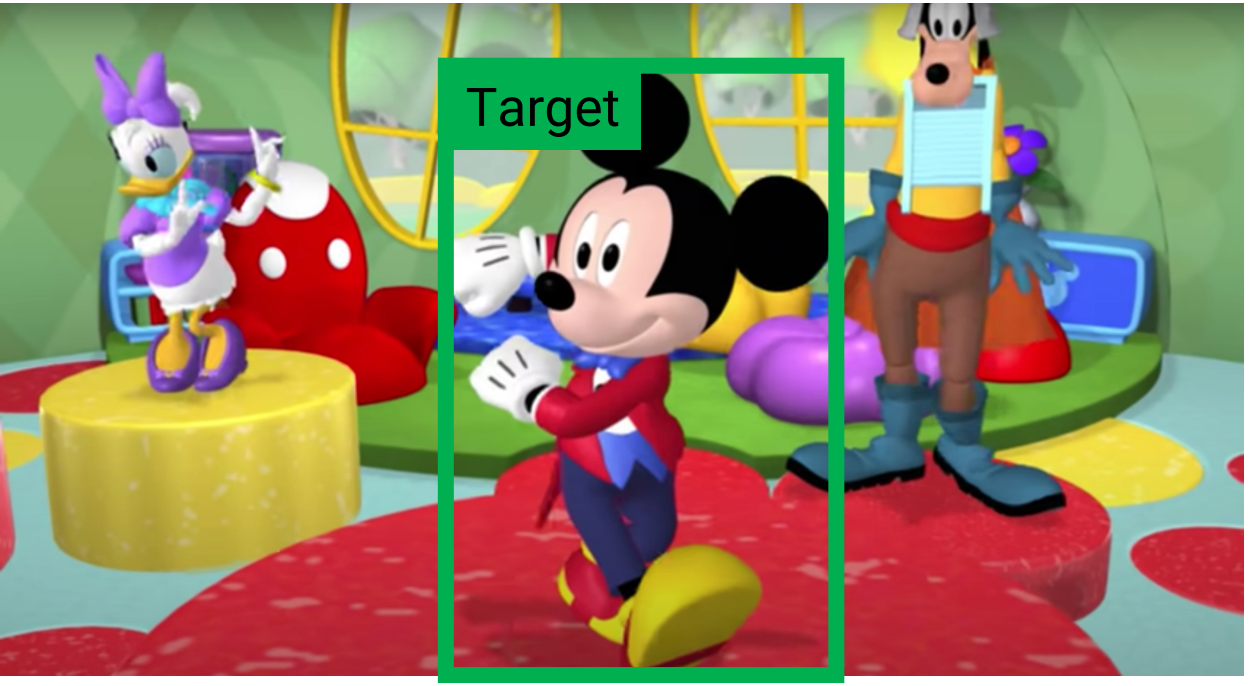
- "follow the movements of somebody/something" – Oxford Dictionary
- Something
  - Point
  - Region
  - Template
- Somebody
  - An object with known identity
  - A person, a vehicle, a face, etc.

What's "Follow"

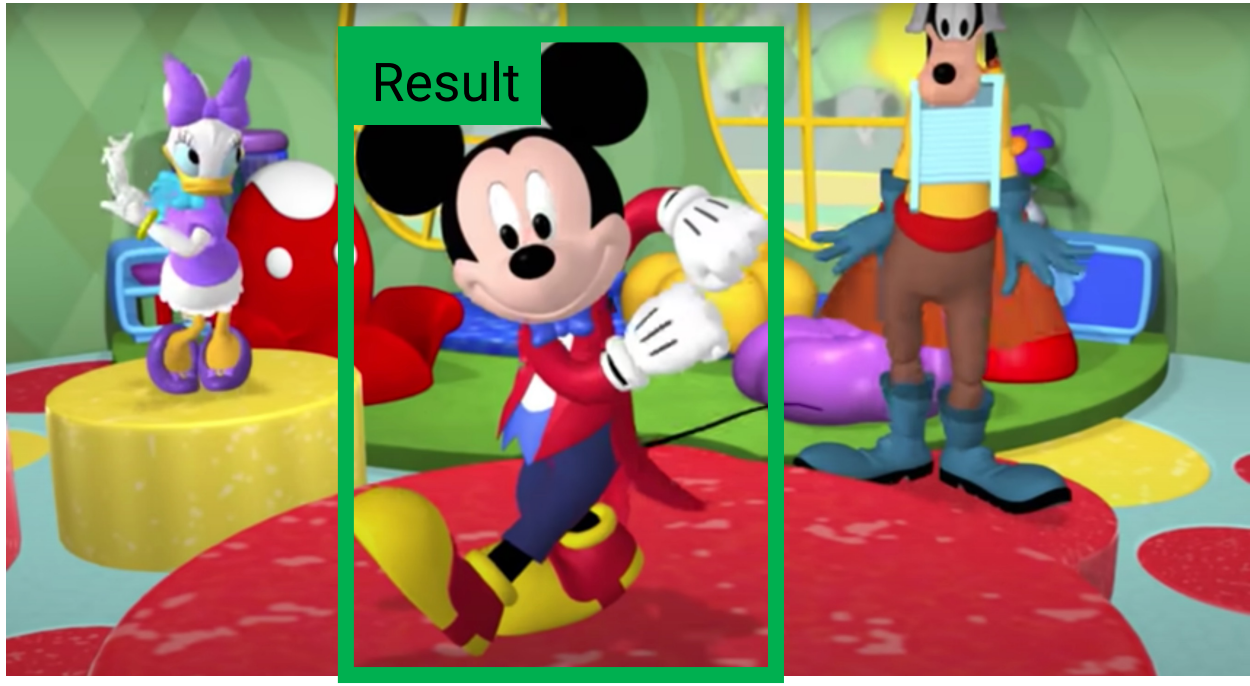
By Eadweard Muybridge, 1878, First Film Ever

ETH zürich VIS Visual Intelligence and Systems

What's "Follow"



Frame T



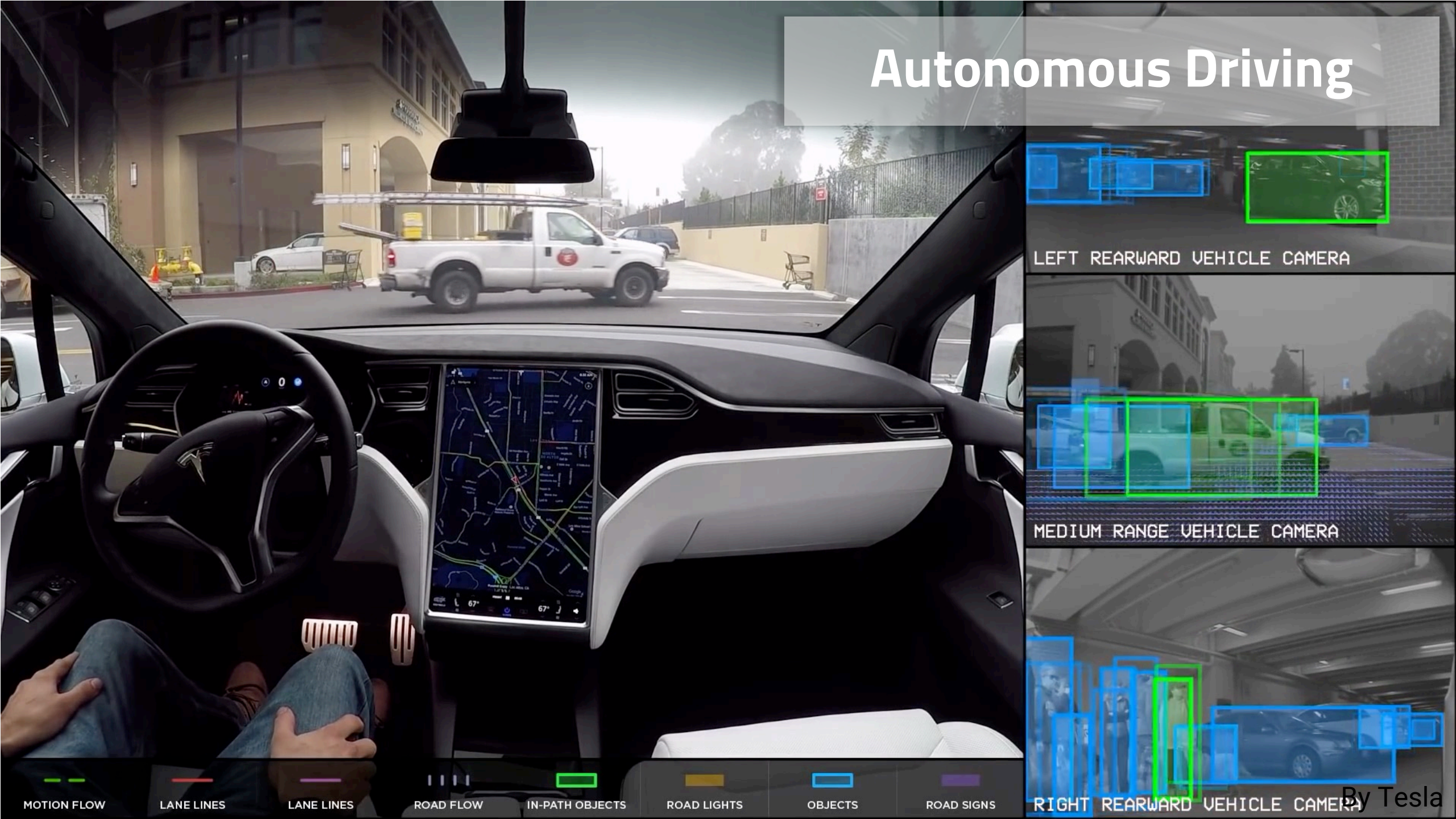
Frame T+1

Tracking Applications

Many!

- Autonomous Driving
- Image Editing
- Safety Monitoring
- Sports
- AR/VR
- Space Management
- You can easily name more!

Autonomous Driving

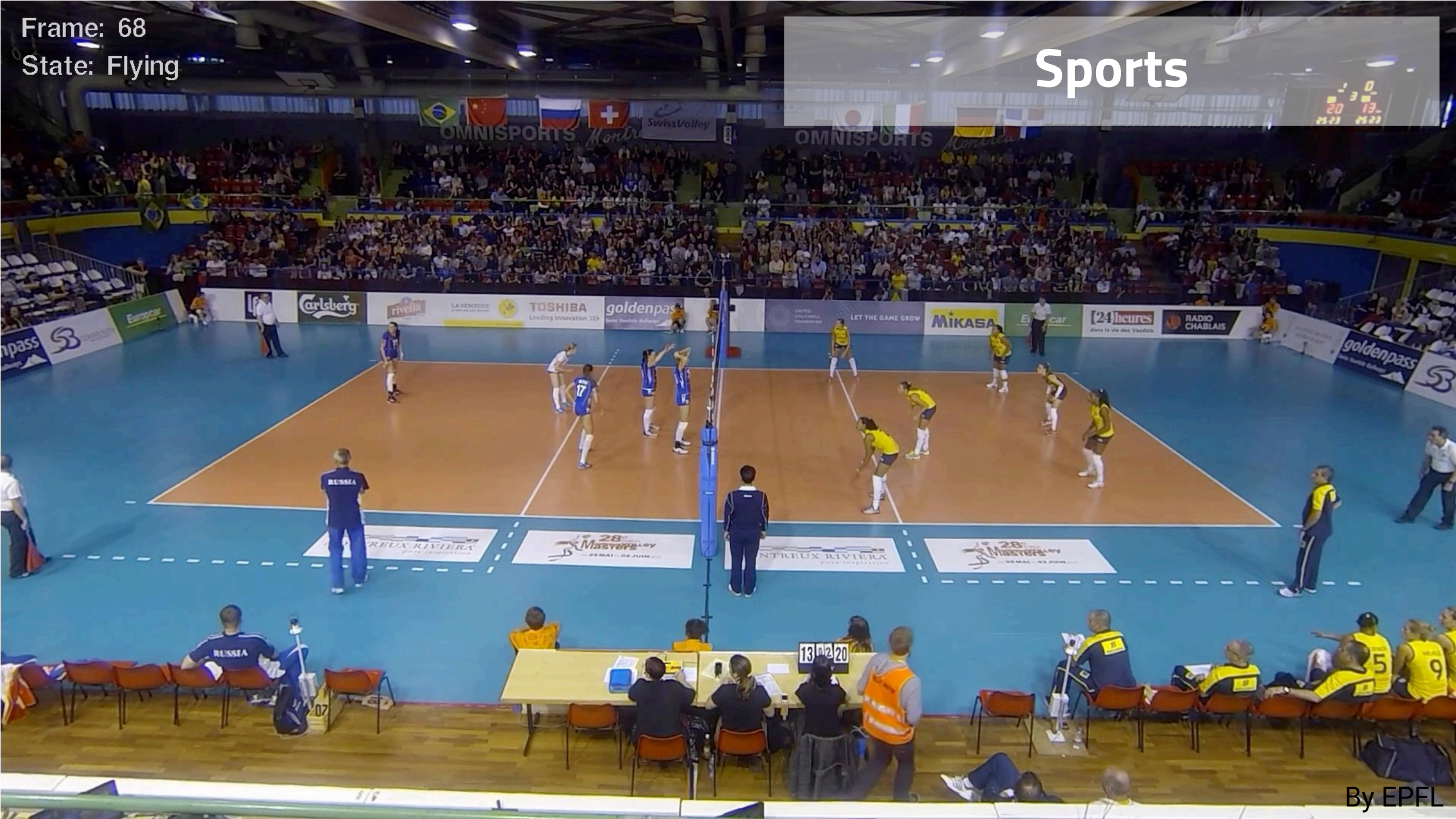


By Tesla

Image Editing

Safety Monitoring

By Microsoft


Frame: 68
State: Flying
Sports
OMNISPORTS
SwissVolley
Carlsberg
TOSHIBA
Leading Innovation
goldenpass
MIKASA
LET THE GAME GROW
24 heures
RADIO CHABLAIS
RUSSIA
By EPFL

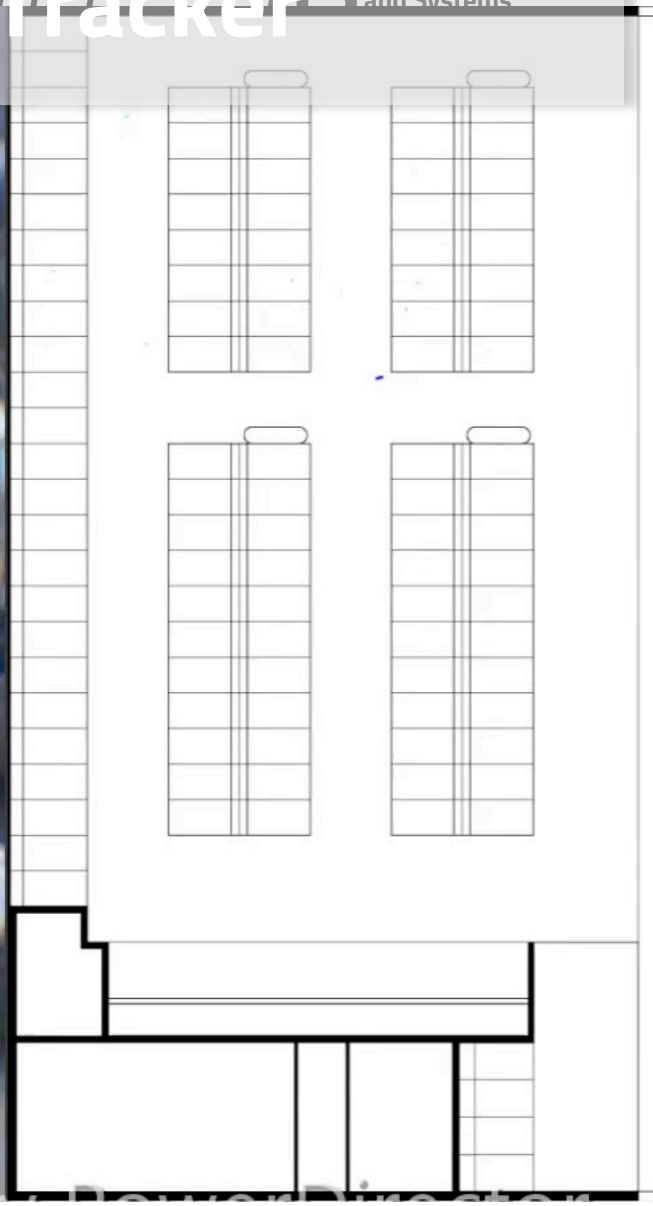
AR/VR

By Microsoft

Customer Tracking

By MEGACOUNT

Cow Tracker




FPS: 0.000000
타는 곳
Track
乗車のりば
나가는 곳
Way Out
Counting
By VCA Technology

Tracking

> “When the pandemic stopped fans attending matches, the club announced it would live stream its games, using an automatic camera system with “in-built, AI, ball-tracking technology” to make sure people always get the best view of the action.”

Football Tracking

THE VERGE TECH REVIEWS SCIENCE CREATORS ENTERTAINMENT VIDEO MORE

TECH \ ARTIFICIAL INTELLIGENCE

TL;DR

AI camera operator repeatedly confuses bald head for soccer ball during live stream

8

Like a distracted AI with a crush

By James Vincent | Nov 3, 2020, 8:07am EST

SHARE

Where’s the ball? There’s the ball, get it! | GIF: YouTube / Chuckiehands

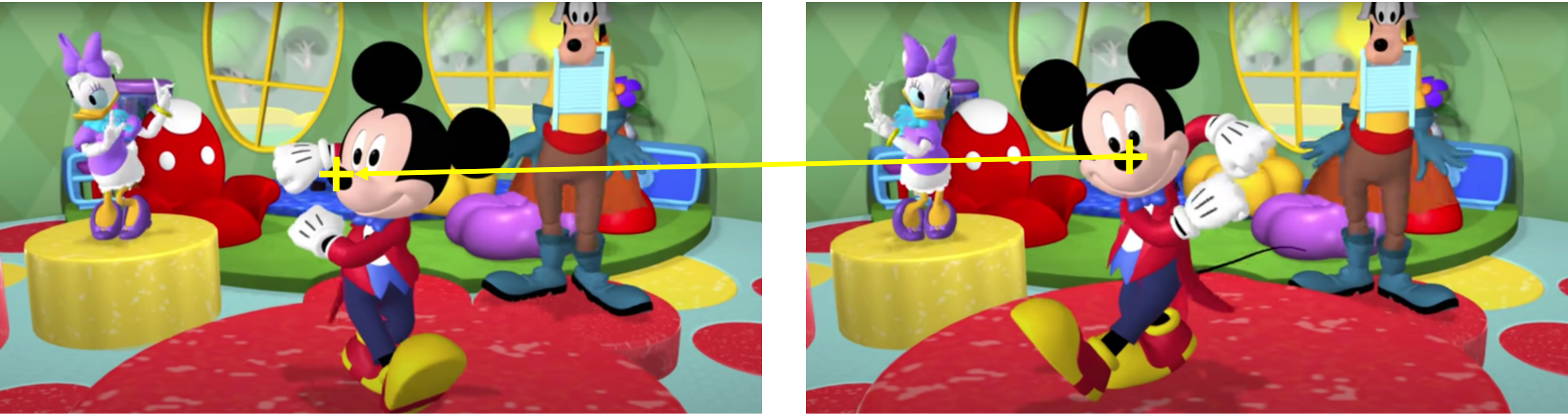
VERGE DEALS

Apple’s AirPods are $60 off this weekend

Tracking

- **Track a point**
- Track a bigger box
- Track by detection
- Online learning
- Motion
- Multiple object tracking
- 3D object tracking

Track a Point



Track a Point

Easy! Let's find the point with the same color!

$$E(h) = [I_0(x + h) - I_1(x)]^2$$

Optimize displacement for this energy so
the two pixels have the same color

Track a Point

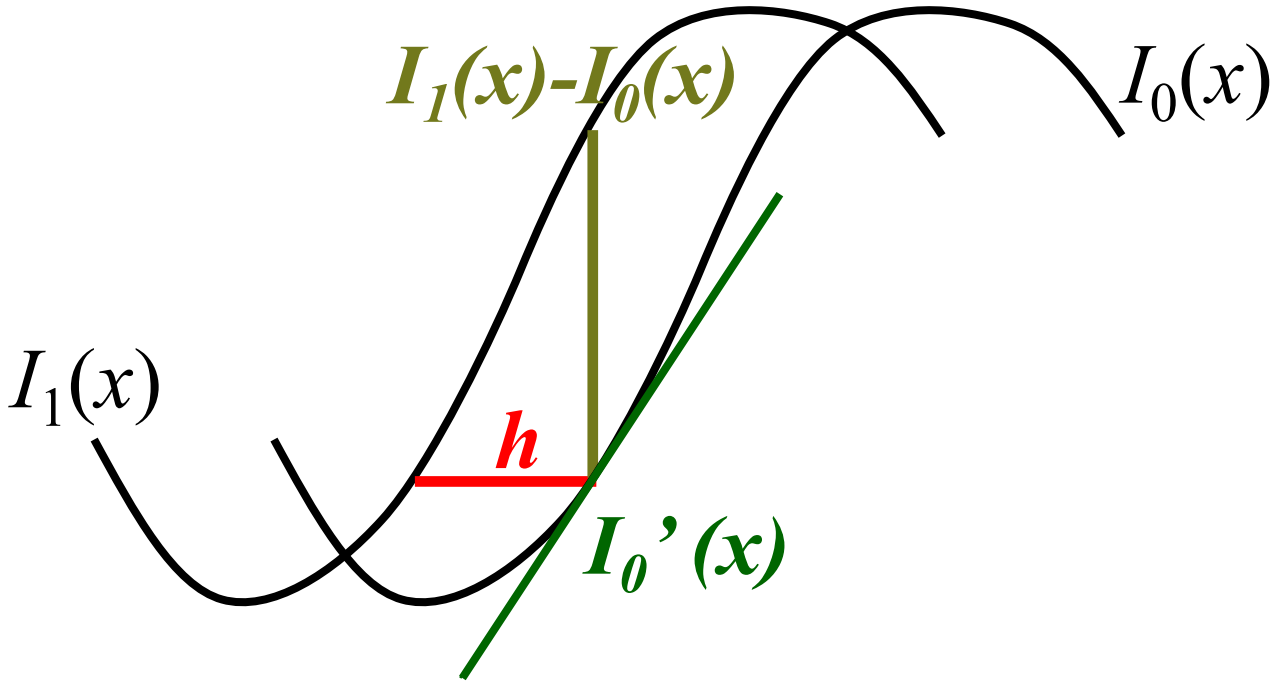
$$E(h) = [I_0(x + h) - I_1(x)]^2$$

$$E(h) \approx [I_0(x) + h\, I_0'(x) - I_1(x)]^2$$

$$\frac{\partial E}{\partial h} = 2I_0'(x)[I_0(x) + hI_0'(x) - I_1(x)]$$

$$\frac{\partial E}{\partial h} = 0 \quad \Rightarrow \quad h \approx \frac{I_1(x) - I_0(x)}{I_0'(x)}$$

Intuition

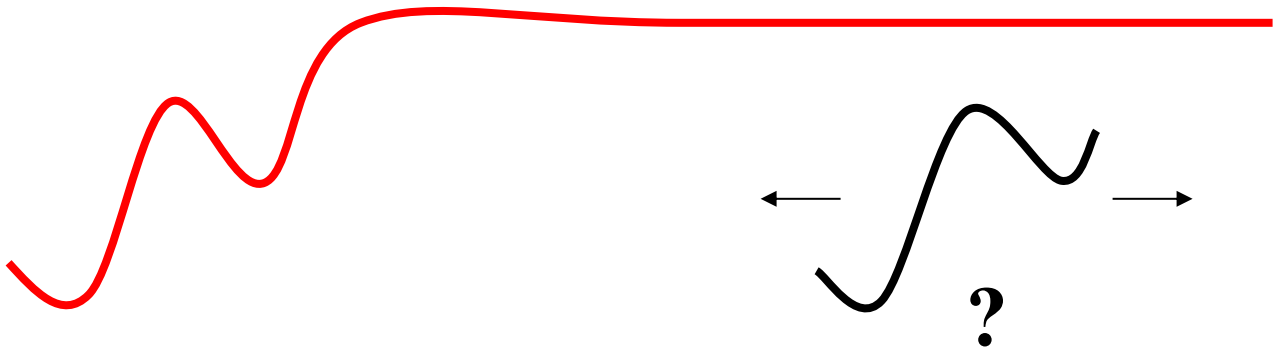


Problem 1 Zero Gradient

What if the image gradient is 0?

$$h \approx \frac{I_1(x) - I_0(x)}{I_0'(x)}$$

a.k.a. The image region is flat!

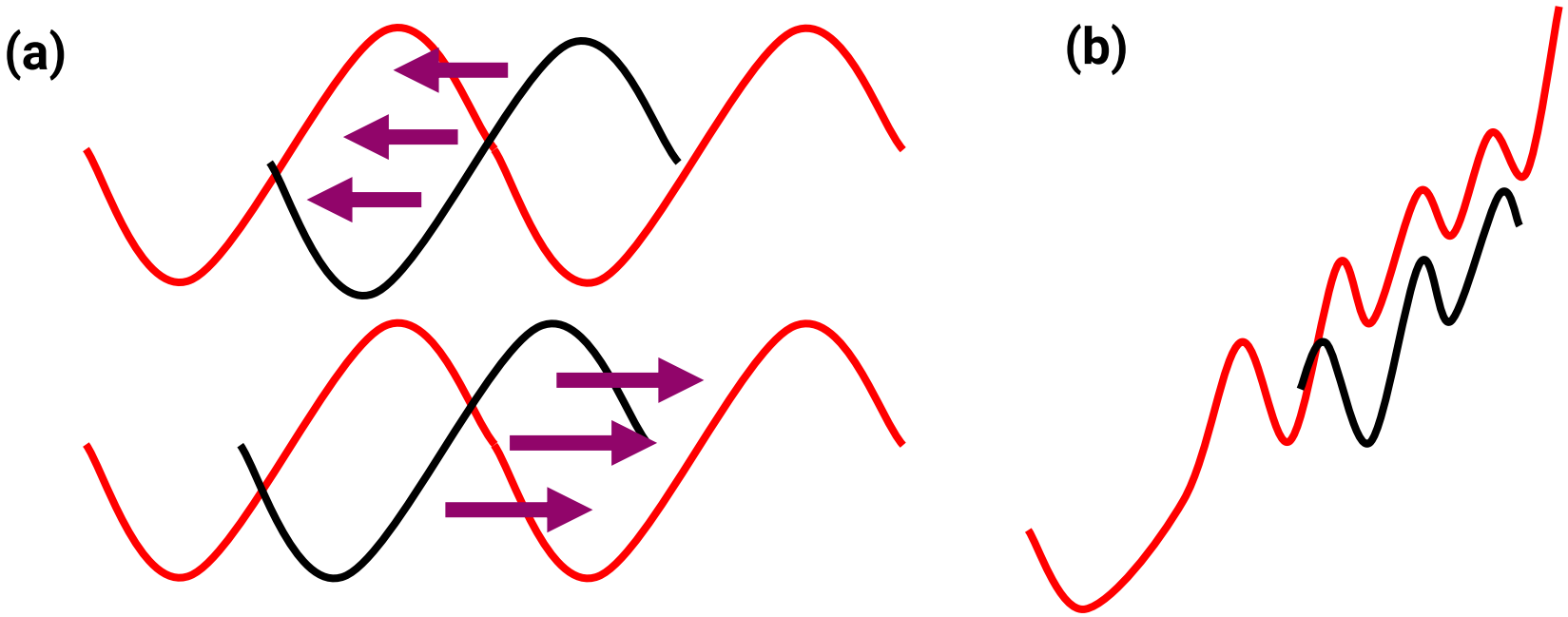


Problem 1 “Aperture problem”

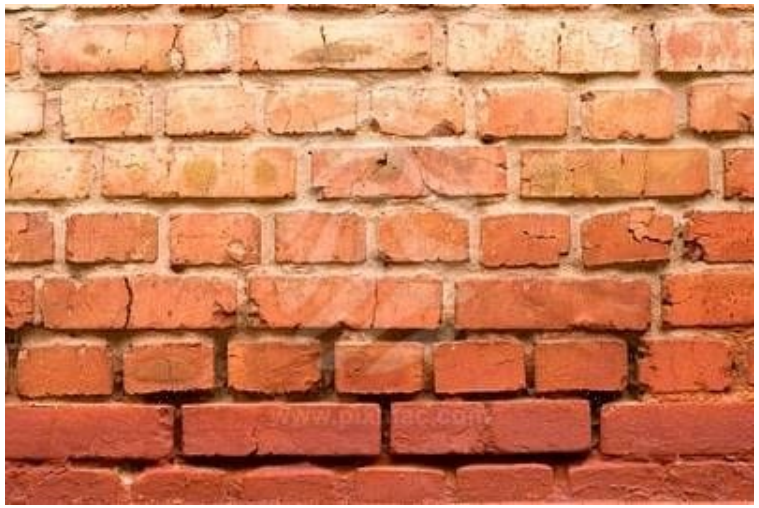
- For tracking to be well defined, nonzero gradients in all possible directions are needed
- If no gradient along one direction, we cannot determine relative motion in that axis

Problem 2 Local Minima

- Motion to closest minimum has to be assumed
- Indirect result: Frame-rate should be faster than motion “of half-wavelength” (Nyquist rate)
- Nonconvex regions may indicate multiple solutions



Problem 2: Local Minima



Recall: Optical Flow in Motion Estimation

- Optical Flow recovers (smooth) motion everywhere
- Least-squares regularization: Horn-Schunk makes smooth spatial change assumption

Recall: Optical Flow

$$I_x u + I_y v + I_t = 0$$

Assuming displacements are small and using Taylor expansion

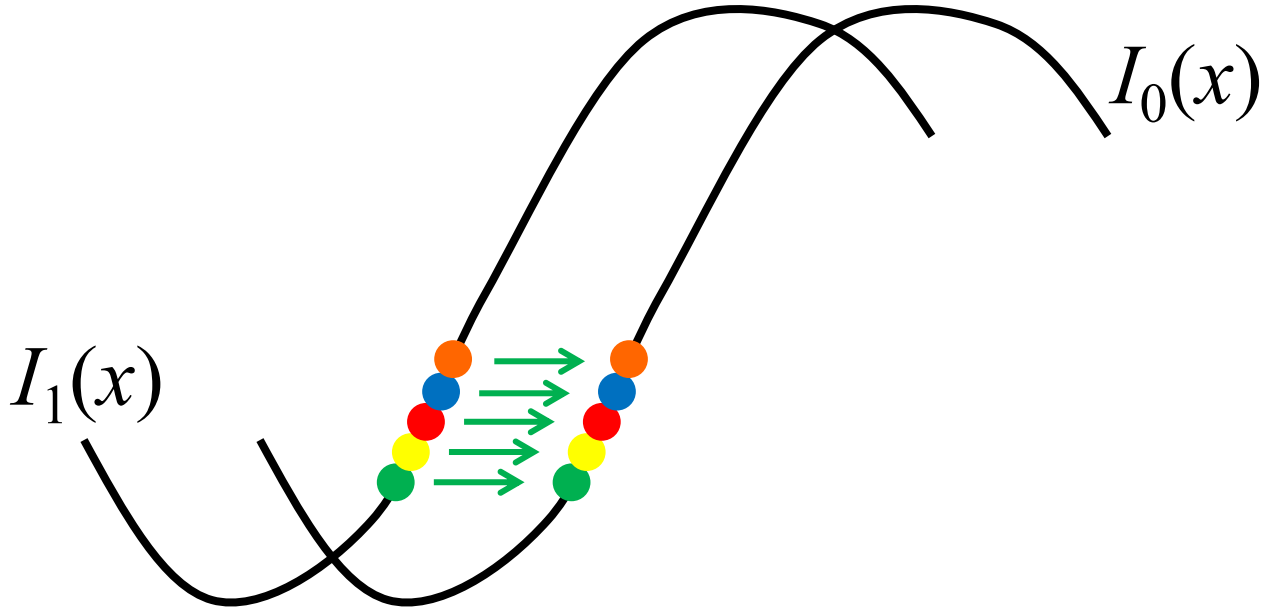
$$I_x = \frac{\partial I}{\partial x}, \quad I_y = \frac{\partial I}{\partial y}, \quad I_t = \frac{\partial I}{\partial t}$$

$$u = \frac{\partial x}{\partial t}, \quad v = \frac{\partial y}{\partial t}$$

1 equation in 2 unknowns

Treating Aperture Problem in Tracking

- Get additional info to constrain motion:
  - Optical Flow: Smoothly regularize in space
  - Tracking: Assume single motion for a region
- Spatial coherence constraint: “A pixel’s neighbors all move together”



Least Squares Problem

$$\begin{bmatrix} I_x(\mathbf{p}_1) & I_y(\mathbf{p}_1) \\ I_x(\mathbf{p}_2) & I_y(\mathbf{p}_2) \\ \vdots & \vdots \\ I_x(\mathbf{p}_n) & I_y(\mathbf{p}_n) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_t(\mathbf{p}_1) \\ I_t(\mathbf{p}_2) \\ \vdots \\ I_t(\mathbf{p}_n) \end{bmatrix}$$

Over determined System of Equations

$$A\,\mathbf{d} = \mathbf{b}$$

Pseudo Inverse

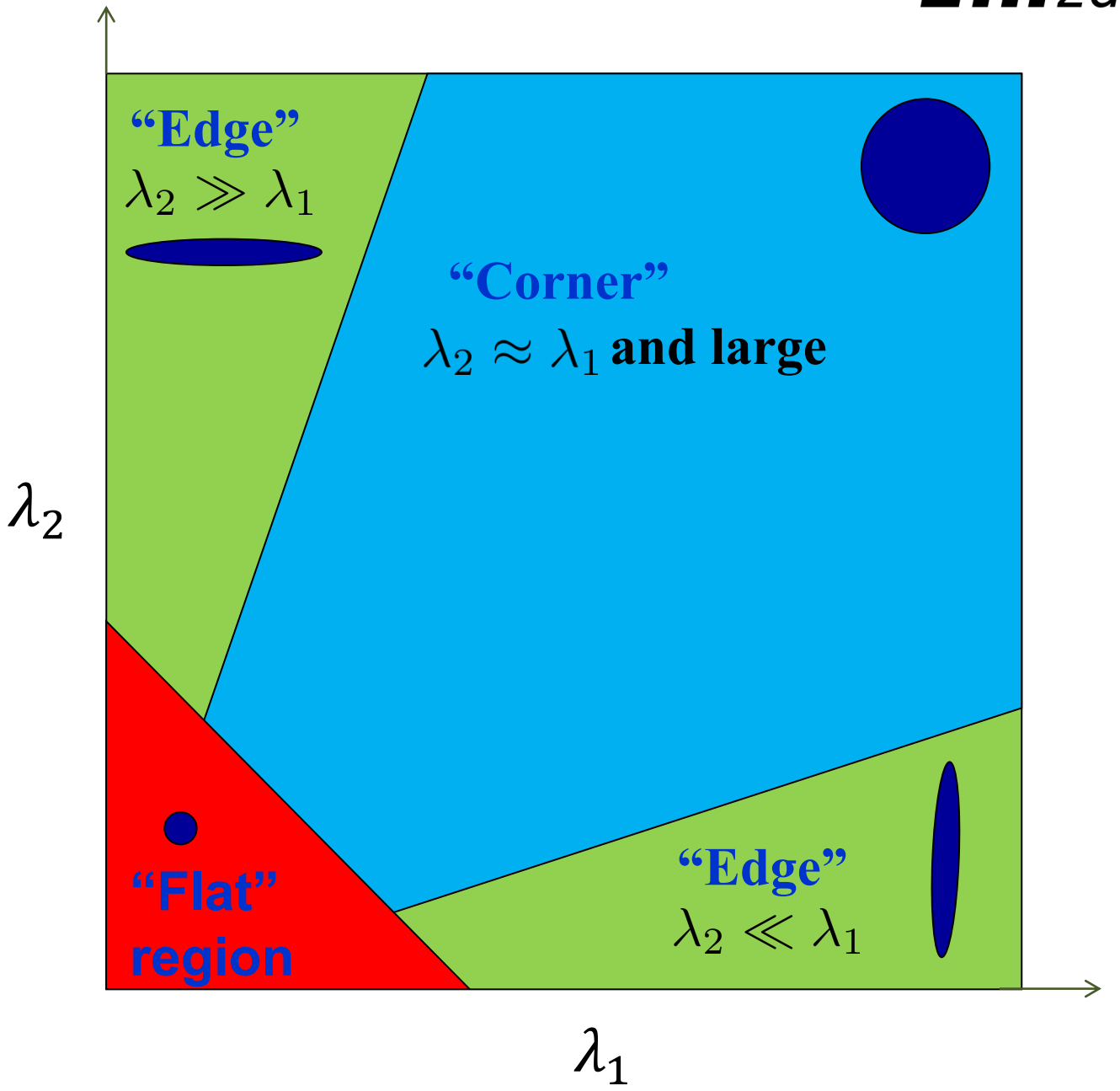
$$(A^T A)\,\mathbf{d} = A^T \mathbf{b}$$

$$\begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix}$$

Least Squares Problem

$$\begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix}$$

- (u,v) can only be found, if this is solvable
  - 2x2 image structure matrix is invertible
  - with no small eigenvalue
- This matrix and the requirement sound familiar – have we seen these before?
- Recall Harris corner detector!
- Thus, good image features (with large structural eigenvalues) are also good for tracking (with which we can find motion)


"Edge"
$\lambda_2 \gg \lambda_1$
"Corner"
$\lambda_2 \approx \lambda_1$ and large
"Flat" region
"Edge"
$\lambda_2 \ll \lambda_1$
λ_2
λ_1

A "Classical" Example

Tracking

- Track a point
- **Track a bigger box**
- Track by detection
- Online learning
- Motion
- Multiple object tracking
- 3D object tracking

Template Tracking

- Keep a template image to compare with each frame
- This is typically applied for small patches, e.g. 5x5
- Why not run it for the entire object (for a larger window)

Template Tracking

- What if the template takes some transformation
  - Translation
  - Rotation
  - Scaling
  - Projective

Lucas-Kanade Template Tracker

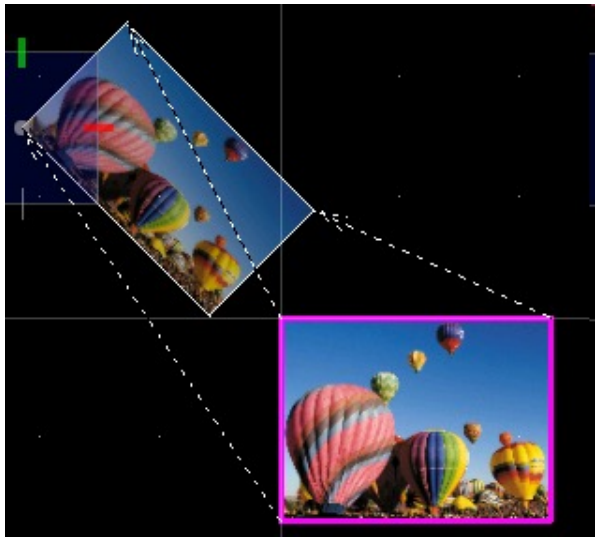
- We can easily generalize the motion model to other parametric models!

$$E(u,v) = \sum_{x,y} [I(x+u, y+v) - T(x,y)]^2$$

$$E(p) = \sum_{x,y} [I(W(x;p)) - T(x,y)]^2$$

Lucas-Kanade Template Tracker

At an iterative step, assuming we know the “optimal” warp *W(x,p)*,



$$\sum_{\mathbf{x}} [I(\mathbf{x} + \Delta\,\mathbf{x}) - T(\mathbf{x})]^2$$

$$\sum_{\mathbf{x}} [I(\mathbf{W}(\mathbf{x}; \mathbf{p} + \Delta\,\mathbf{p})) - T(\mathbf{x})]^2$$

$$\sum_{\mathbf{x}} [I(\mathbf{W}(\mathbf{x};\mathbf{p}+\Delta\,\mathbf{p})) - T(\mathbf{x})]^2$$

- $\sum_{\mathbf{x}} [I(\mathbf{W}(\mathbf{x};\mathbf{p})) + \nabla I \frac{\partial \mathbf{W}}{\partial \mathbf{p}} \Delta\mathbf{p} - T(\mathbf{x})]^2$ — By Taylor expansion
- $\sum_{\mathbf{x}} \left[\nabla I \frac{\partial \mathbf{W}}{\partial \mathbf{p}}\right]^{\mathrm{T}} [I(\mathbf{W}(\mathbf{x};\mathbf{p})) + \nabla I \frac{\partial \mathbf{W}}{\partial \mathbf{p}} \Delta\mathbf{p} - T(\mathbf{x})]$ — Derivative of $\Delta\mathbf{p}$
- $\Delta\mathbf{p} = H^{-1} \sum_{\mathbf{x}} \left[\nabla I \frac{\partial \mathbf{W}}{\partial \mathbf{p}}\right]^{\mathrm{T}} [T(\mathbf{x}) - I(\mathbf{W}(\mathbf{x};\mathbf{p}))]$ — Set the derivative to 0

Where $H = \sum_{x} \left[\nabla I \frac{\partial \mathbf{W}}{\partial \mathbf{p}}\right]^{\mathrm{T}} \left[\nabla I \frac{\partial \mathbf{W}}{\partial \mathbf{p}}\right]$

Lucas-Kanade Template Tracker

Warp I to obtain $I(W([x\ y];P))$

Compute the error image $T(x) - I(W([x\ y];\ P))$

Warp the gradient ∇I with $W([x\ y];\ P)$

Evaluate $\frac{\partial W}{\partial P}$ at $([x\ y];\ P)$ (Jacobian)

Compute steepest descent images $\nabla I \frac{\partial W}{\partial P}$

Compute Hessian matrix $\sum (\nabla I \frac{\partial W}{\partial P})^T (\nabla I \frac{\partial W}{\partial P})$

Compute $\sum (\nabla I \frac{\partial W}{\partial P})^T (T(x,y) - I(W([x,y];P)))$

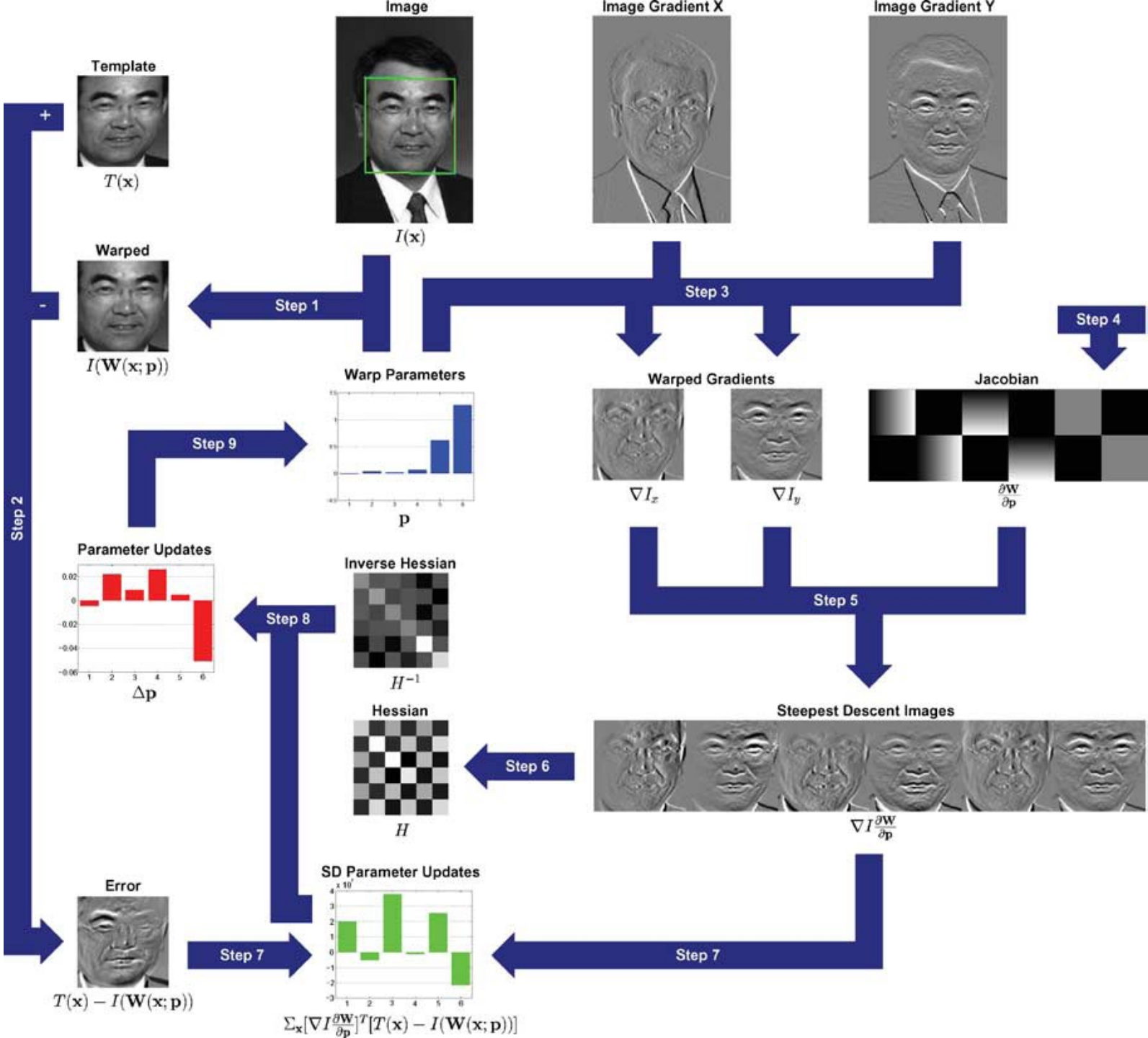
Compute ΔP

Update $P \longleftarrow P + \Delta P$

Iterative update

Baker & Matthews, IJCV'04

Lucas-Kanade 20 Years On:
A Unifying Framework



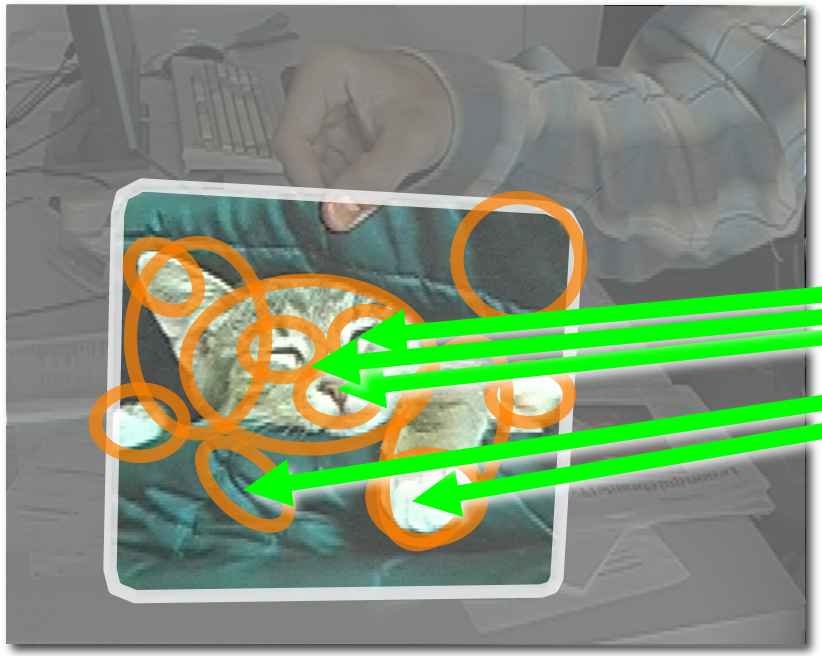
Lucas-Kanade Template Tracker

- Good
  - It is a beautiful framework
  - It can handle different parameter space
  - It can converge fast in a high-frame rate video
- Bad
  - Not robust to image noise or large displacement
  - Some transformations are impossible to parameterize

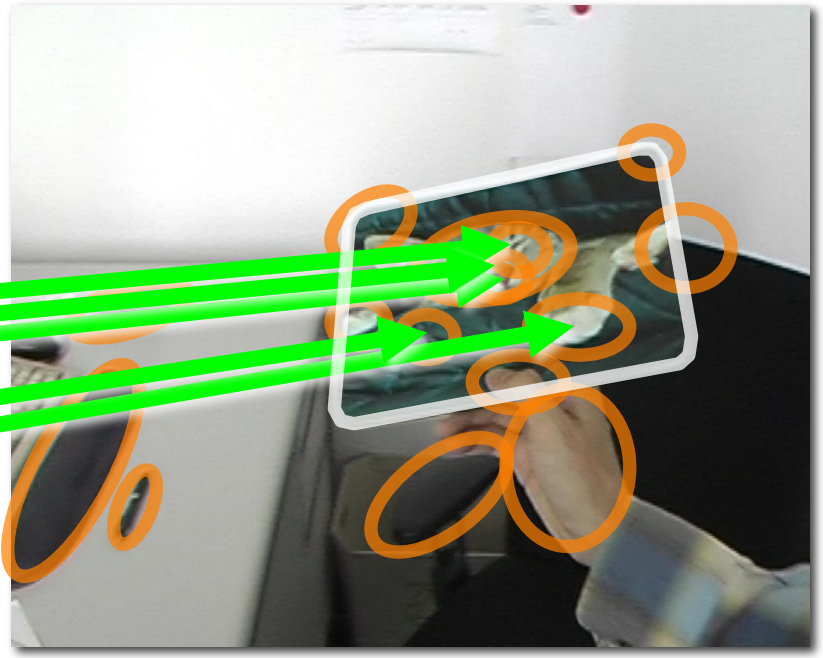
Tracking

- Track a point
- Track a bigger box
- Track by detection
  - **Use features to search for the object**
  - Known object category
- Online learning
- Motion
- Multiple object tracking
- 3D object tracking

Tracking by Features



Reference image(s) of the object to detect



Test image

Template Detection

Reference image(s) of the object to detect

MathWorks

Test image

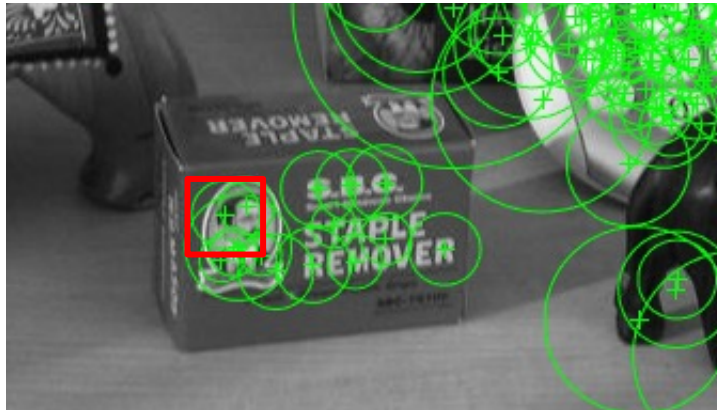
Detect Key Points

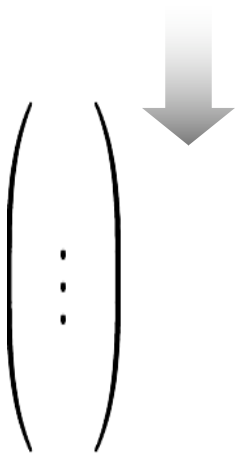
Invariant to scale, rotation, or perspective

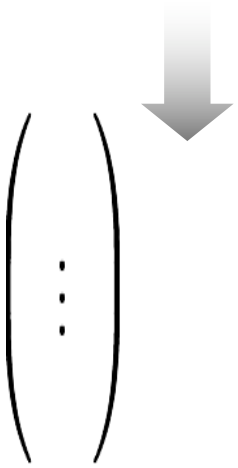
100 strongest feature points
in the reference image

300 strongest feature points
in the test image

Build Feature Descriptors

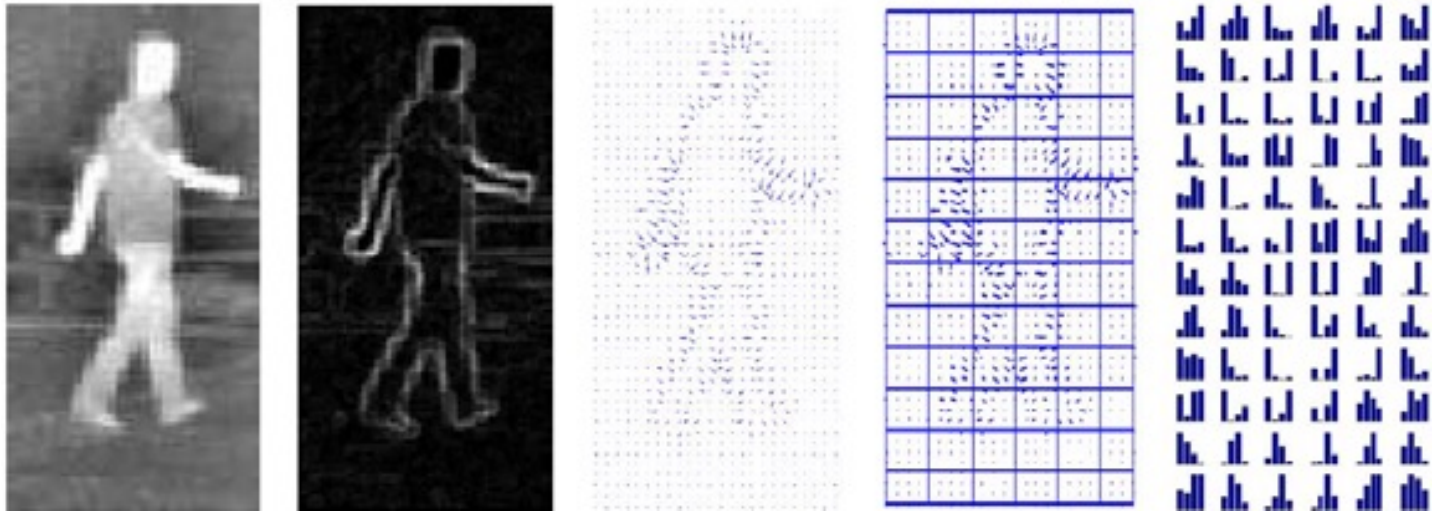


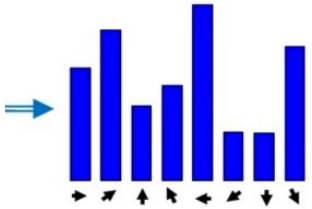




Example: Histogram of Oriented Gradients

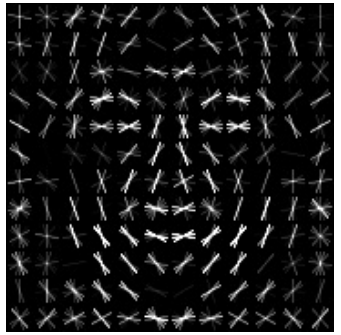
HOG is a (rotation invariant) feature descriptor





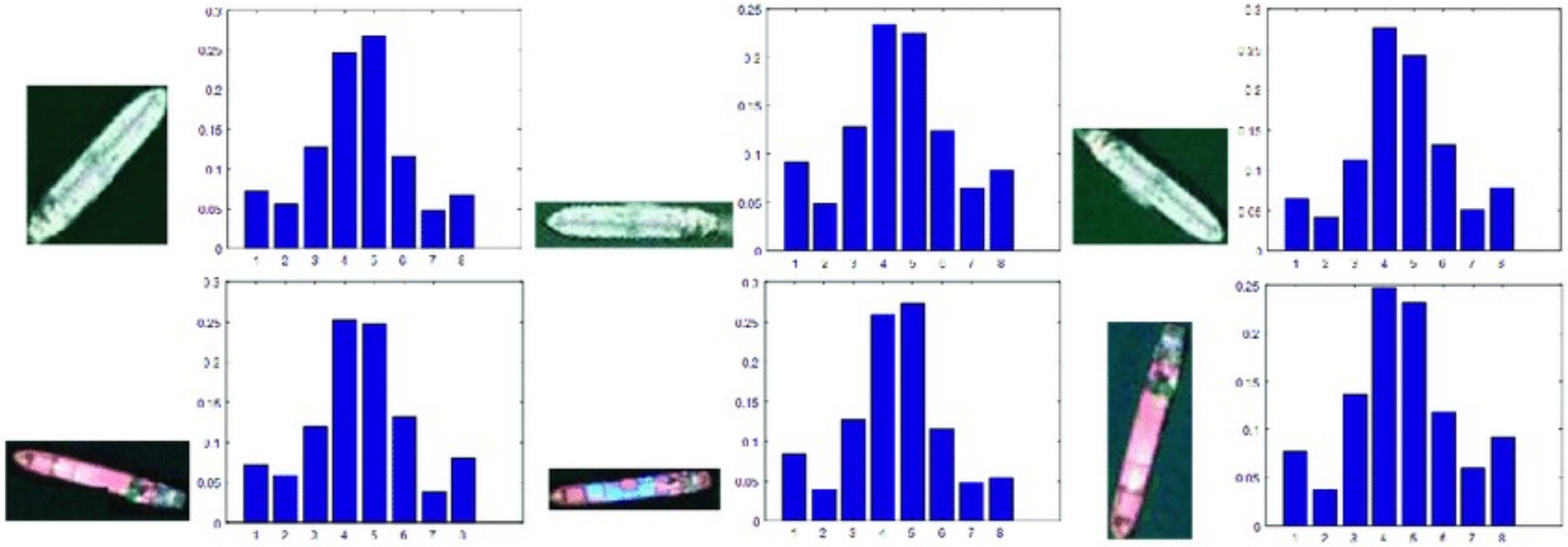
Bin magnitudes of gradients as a histogram

Track specific objects

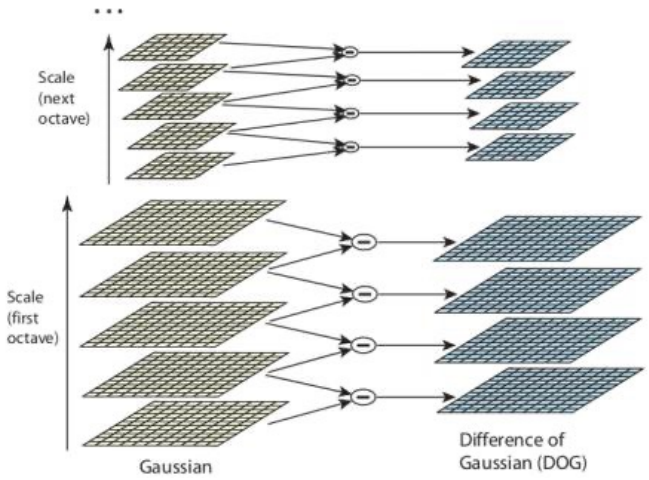


Example: Histogram of Oriented Gradients

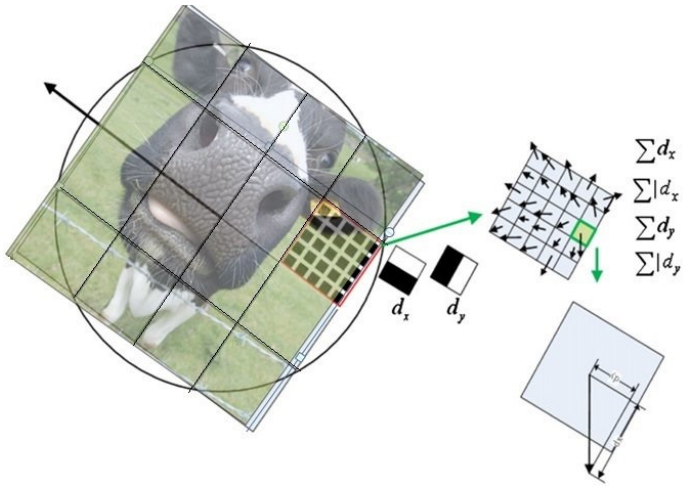
Object shapes defined by edges, thus HOG over entire objects can be descriptive



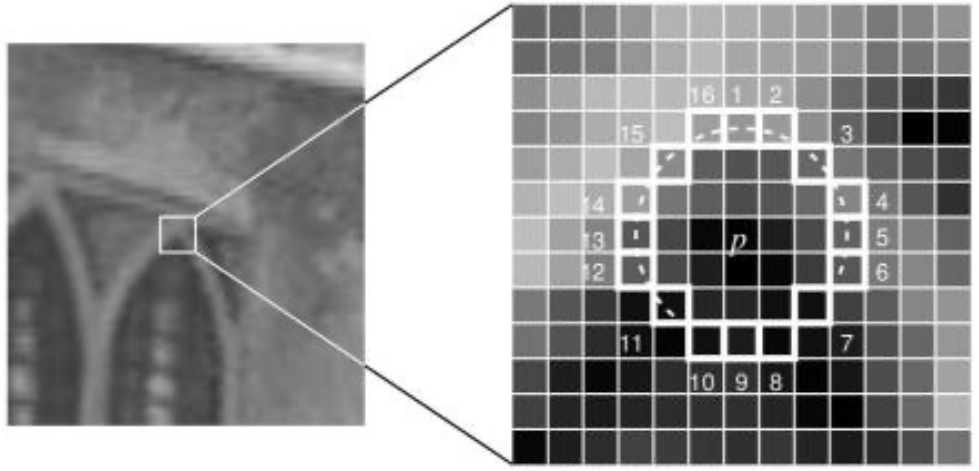
Other Features



SIFT

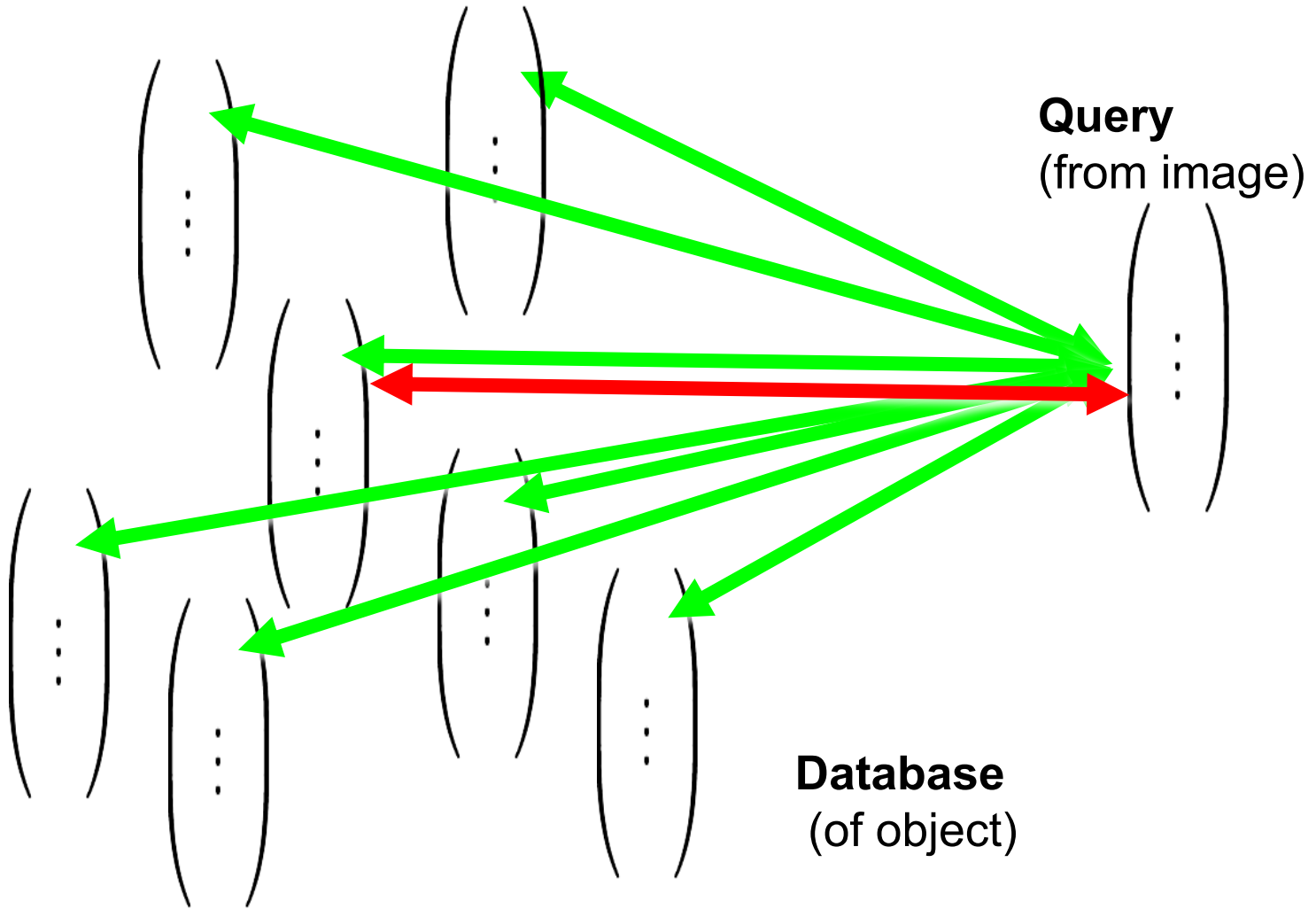


SURF



FAST

Match Keypoint Descriptors

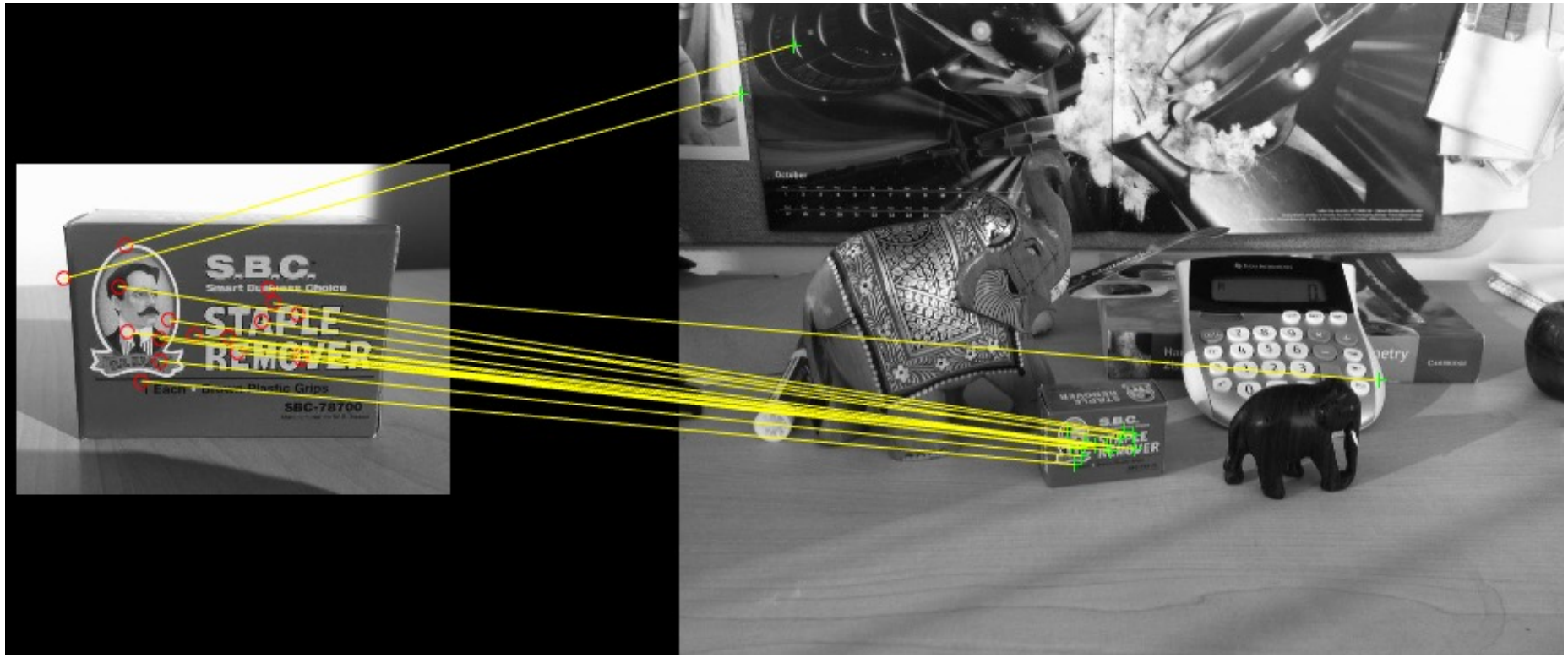


Match Keypoint Descriptors

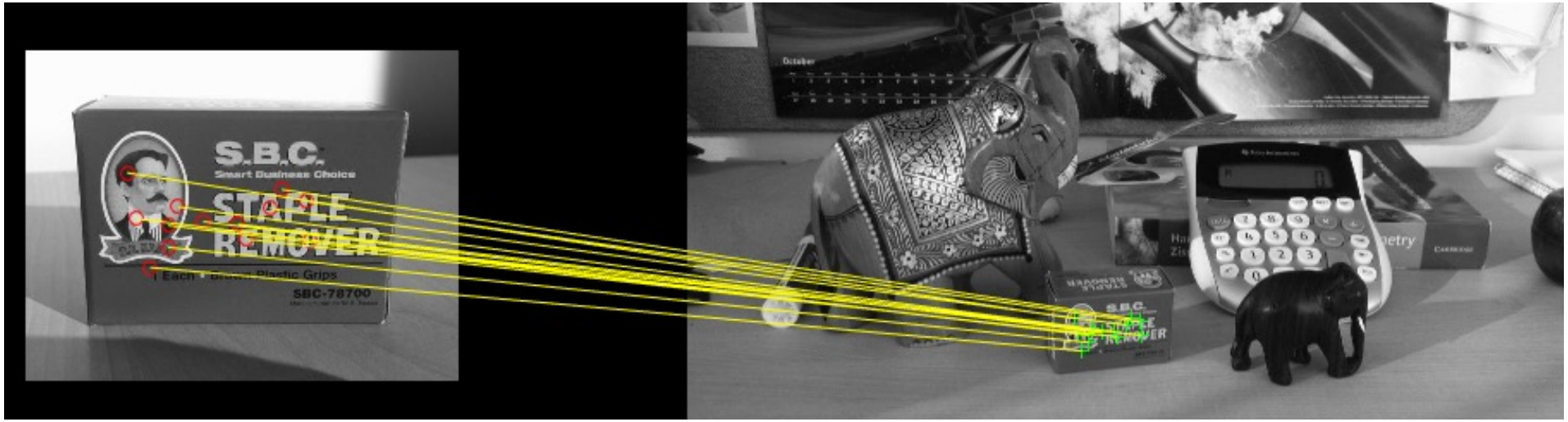
This can be accelerated in multiple ways

- Tree structure (KD-tree)
- Approximate nearest neighbors
- Hashing
- Parallel implementation

Match Keypoint Descriptors



Outlier Elimination



Summary

Keypoint Detection

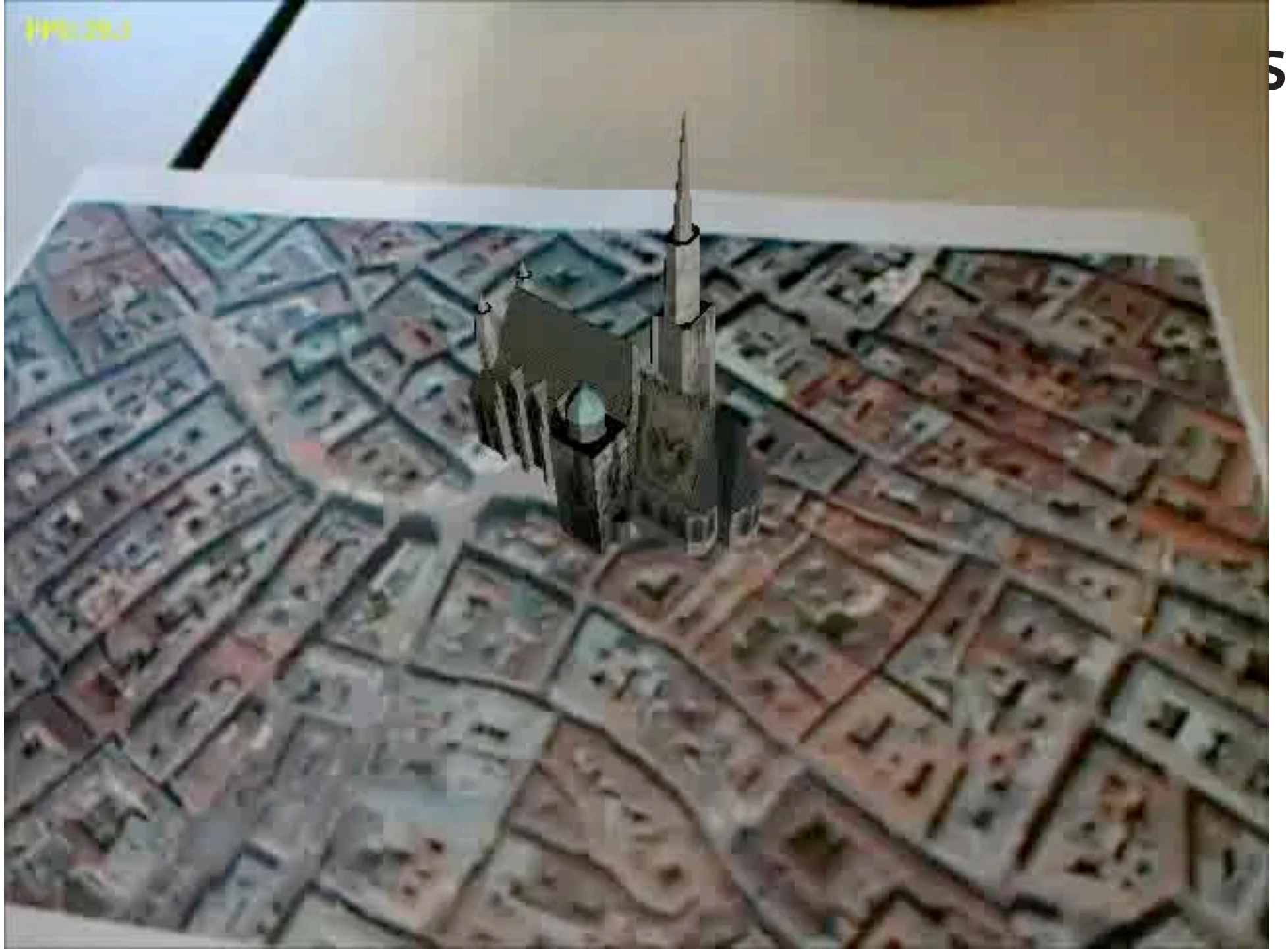
Keypoint Recognition

Geometric verification

Search in the Database

Robust 3D Pose Calculation (RANSAC)


Overall: 3.42 ms
Find Pts: 1.25 ms
Track Pts: 0.32 ms
Features: 1.16 ms
Outliers: 0.50 ms
Pose: 0.19 ms
Corners: 166
Matched Features: 29
Wrong Rotation: 0
Bad Linetest: 0
Bad Homographytest: 0
Correct: 29
From Cache: 0
From ActiveSearch: 20
Levels: 0 0 0 0 0 0 0 0 0 0
Rotation: 6
Avg. Reproj. Err: 1.31
www.topmodelgossip.com

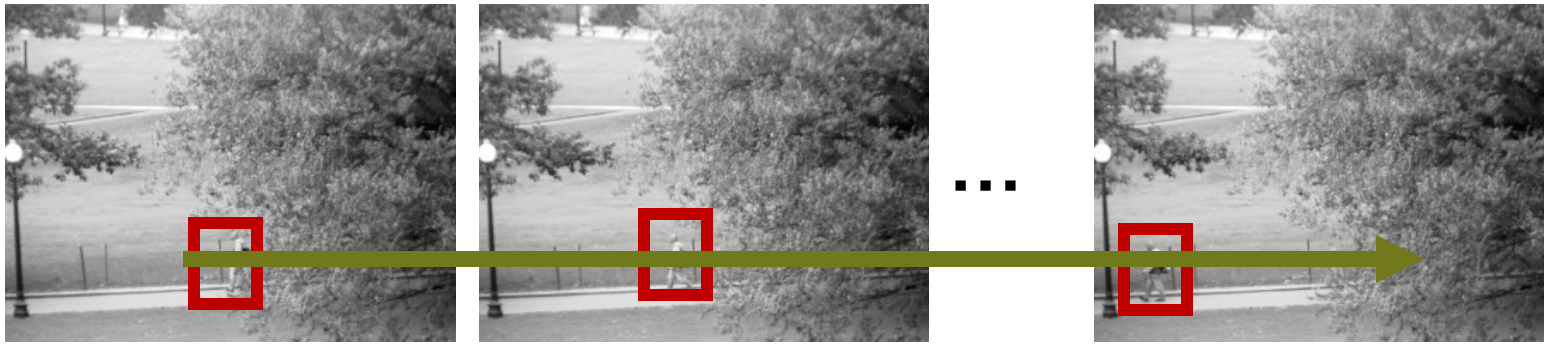


Tracking

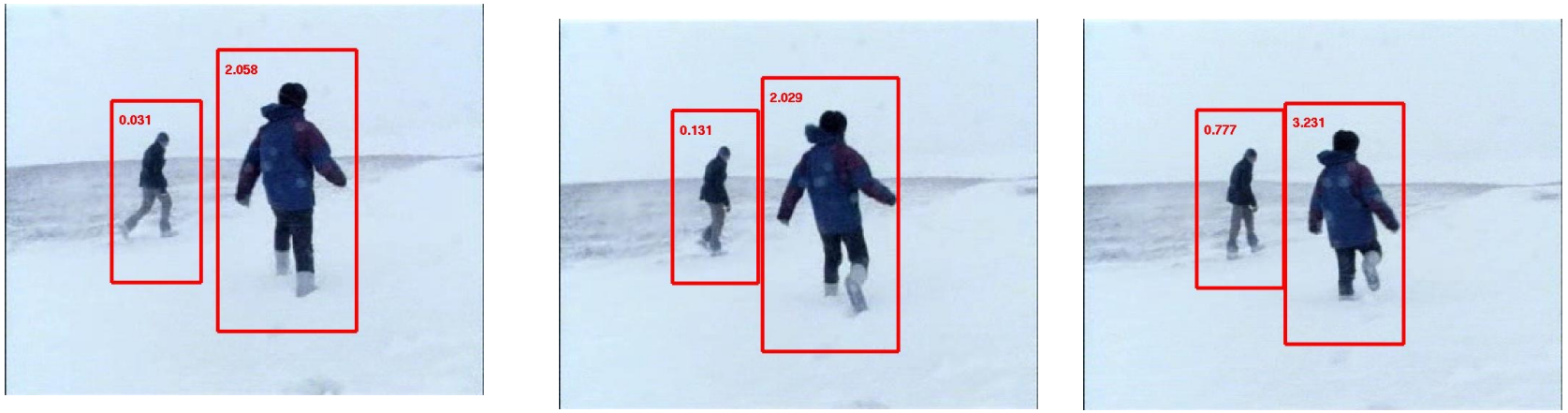
- Track a point
- Track a bigger box
- Track by detection
  - Use features to search for the object
  - **Known object category**
- Online learning
- Motion
- Multiple object tracking
- 3D object tracking

Track by Detection

- Detect object(s) independently in each frame
- Associate detections over time into *tracks*



Multiple Objects



Detecting Objects

Usually turn to supervised learning

Persons

Background

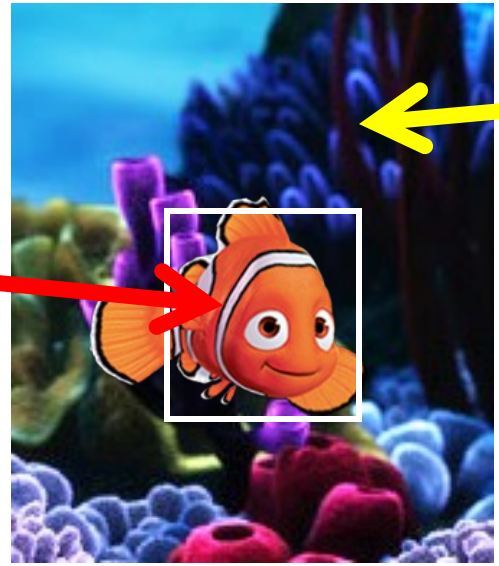
Modern detectors based on ConvNets can be both accurate and fast

Tracking

- Track a point
- Track a bigger box
- Track by detection
- **Online learning**
- Motion
- Multiple object tracking
- 3D object tracking

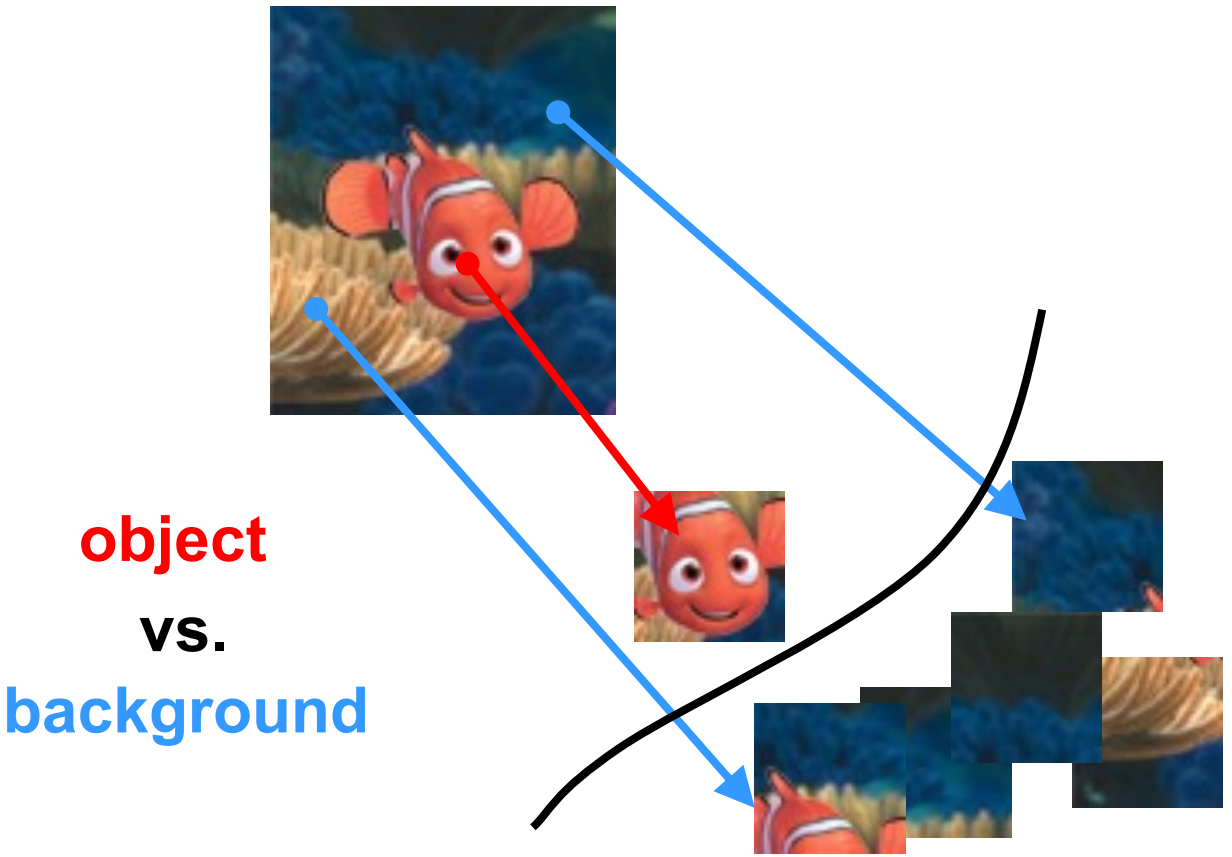
Online Learning of Appearance Model

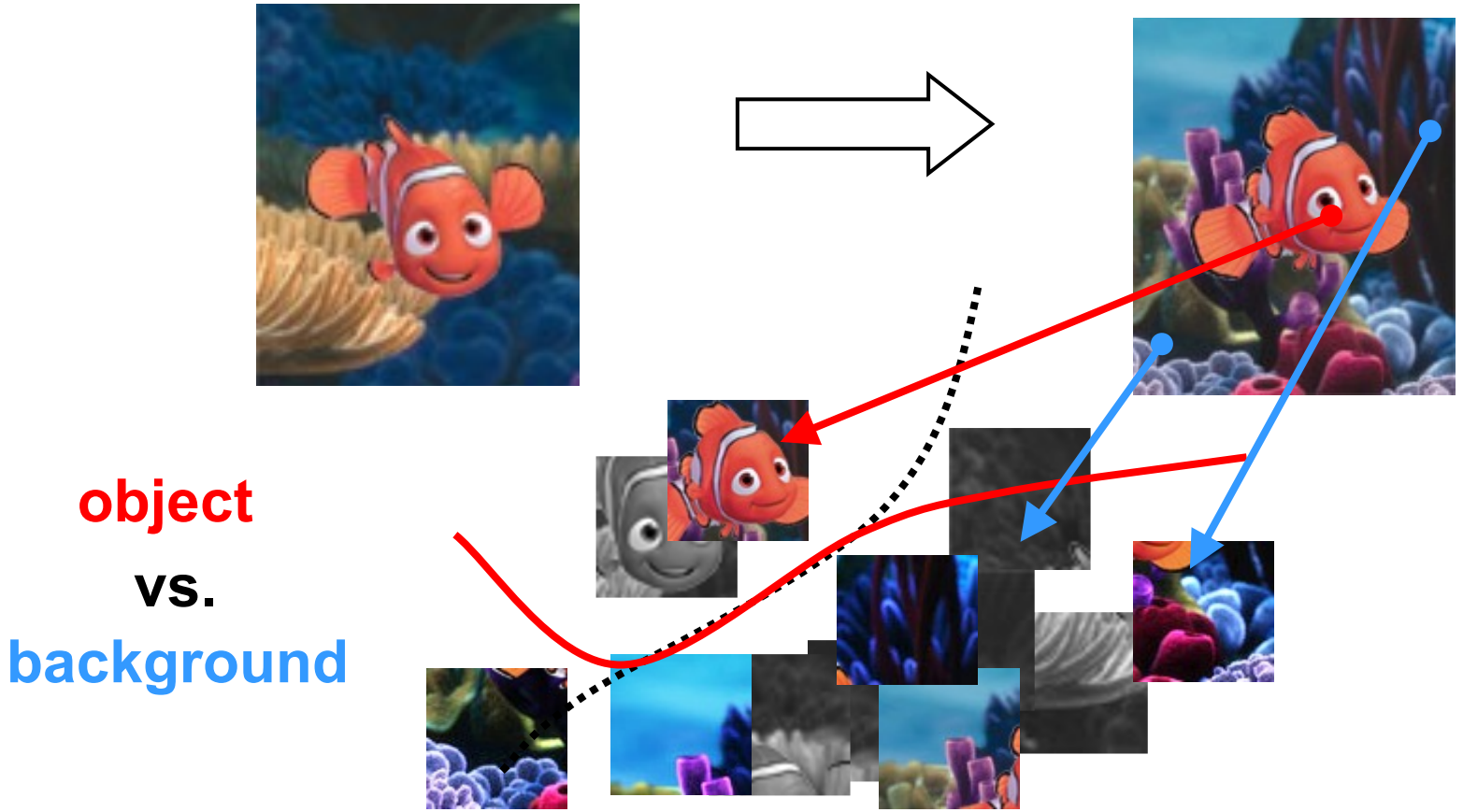
Learning current object appearance vs. local background.



Current background

Current obj. appearance


object
vs.
background

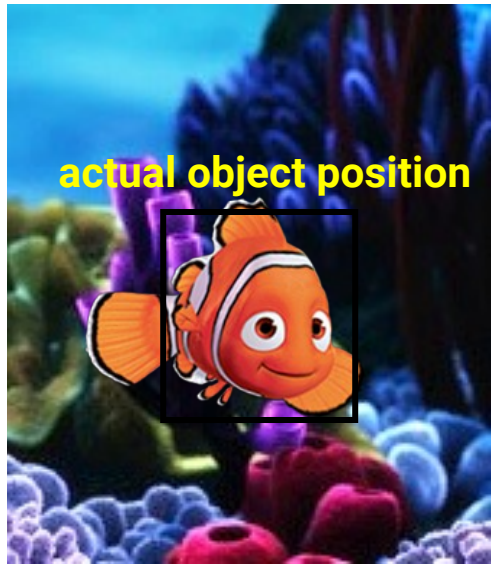


object

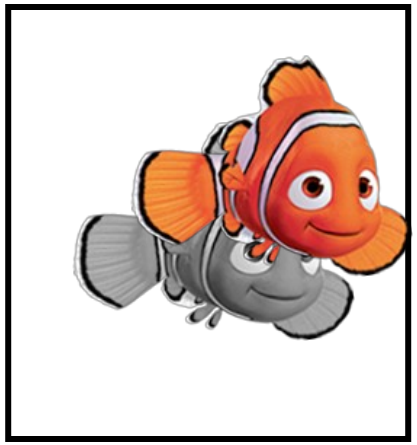
vs.

background

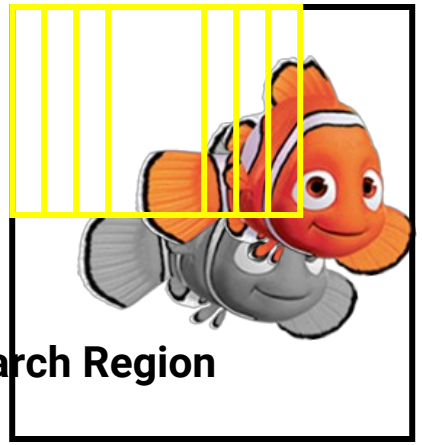
Tracking Loop



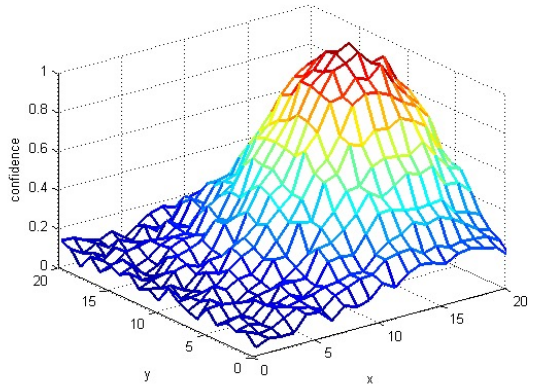
from time t to t+1



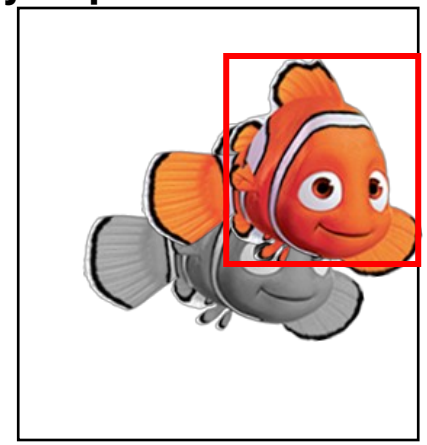
evaluate classifier on sub-patches



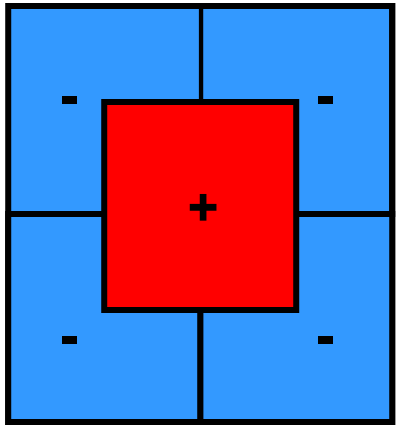
create confidence map

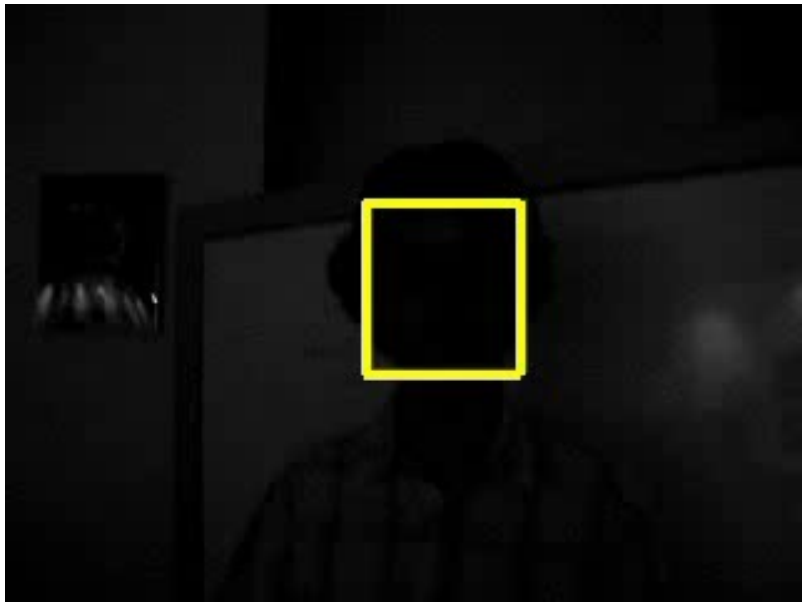


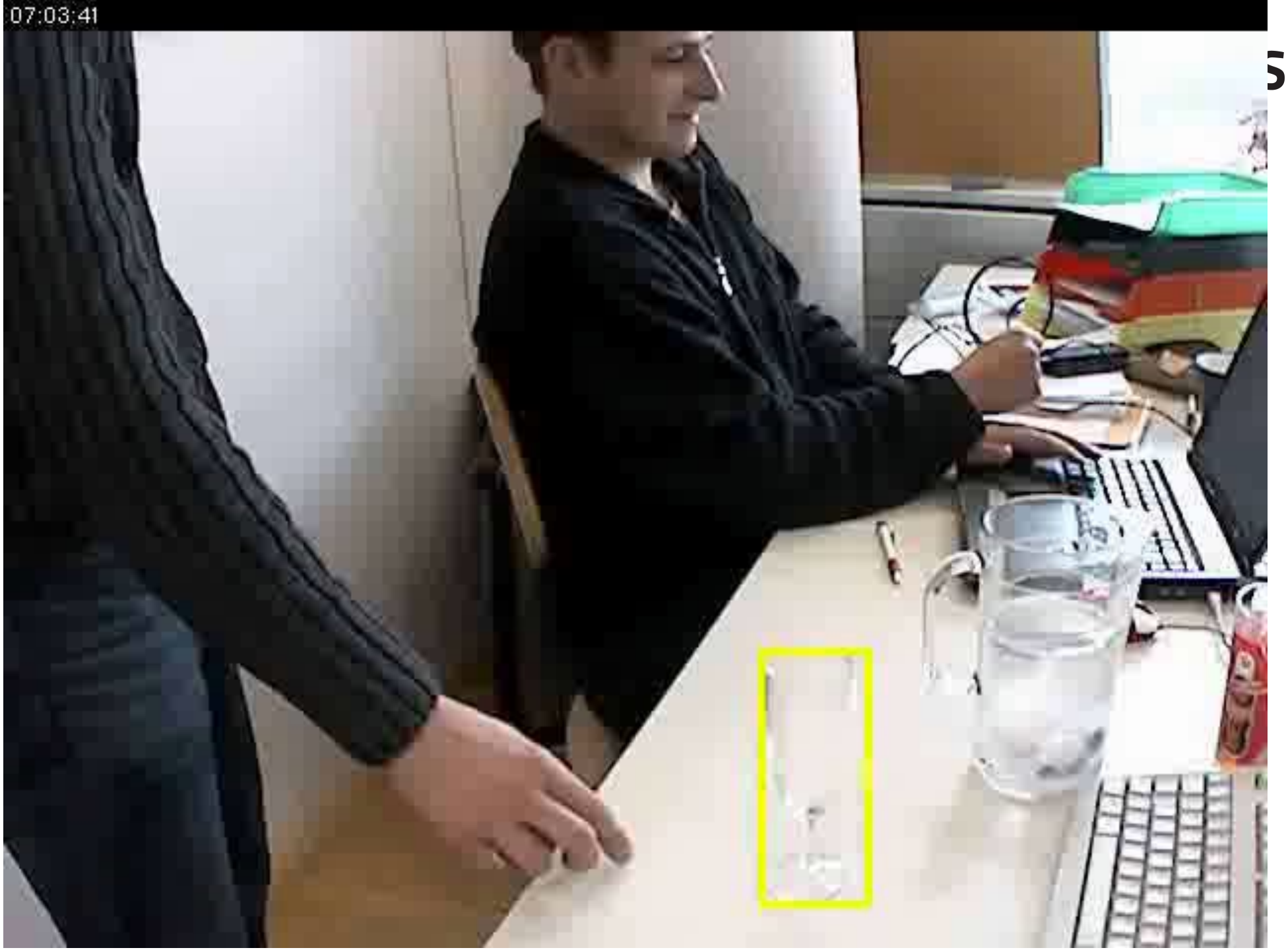
analyze map and set new object position



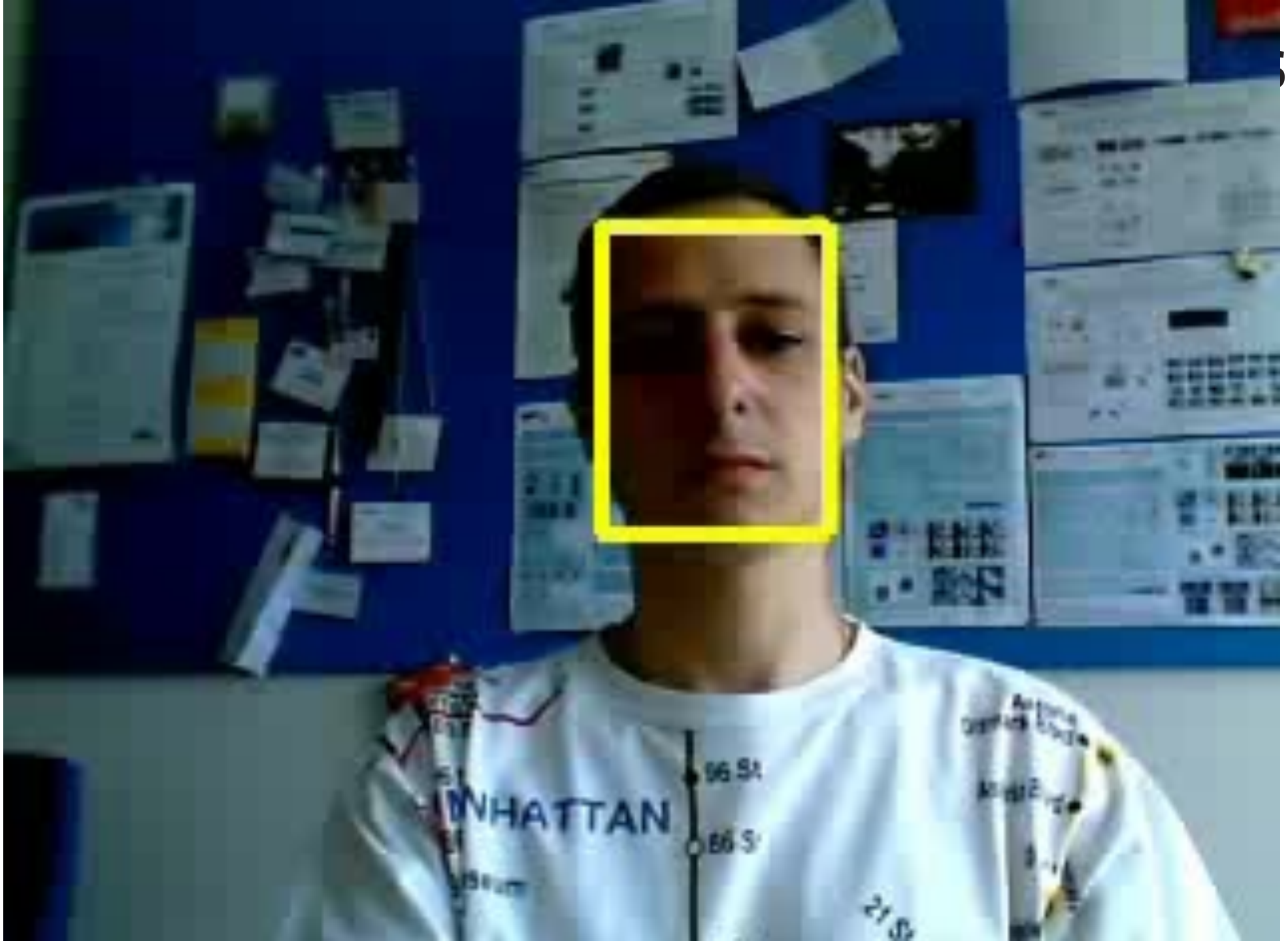
update classifier (tracker)

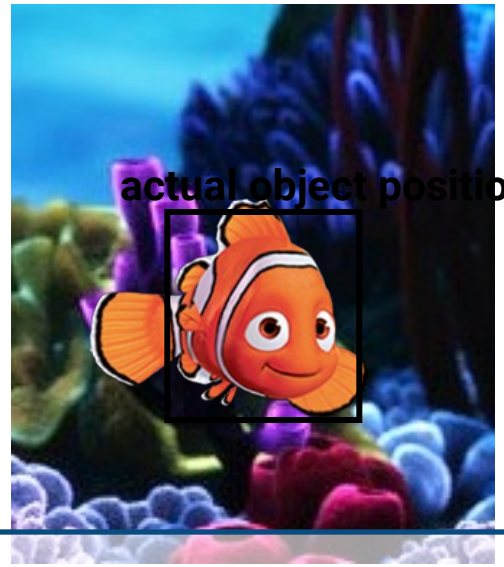




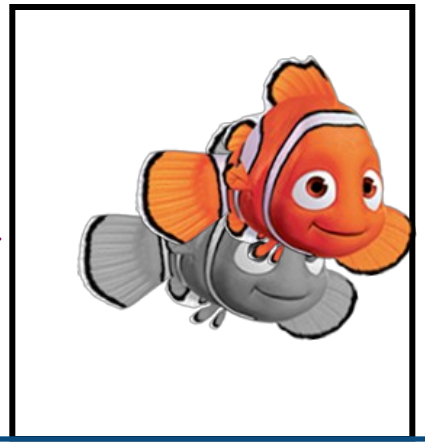


Visual Intelligence and Systems

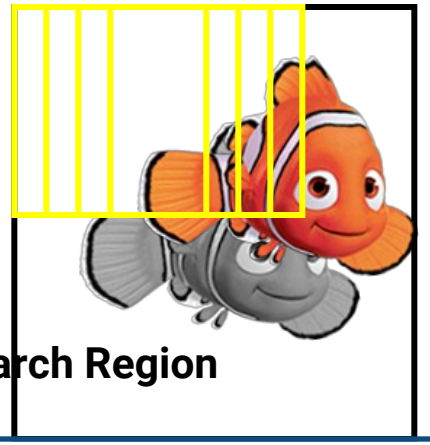

NHATTAN
96 St
86 St



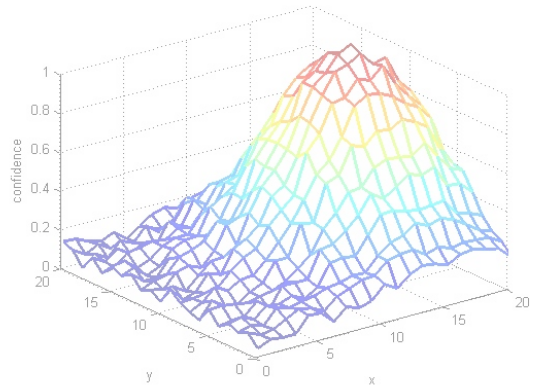
from time t to t+1



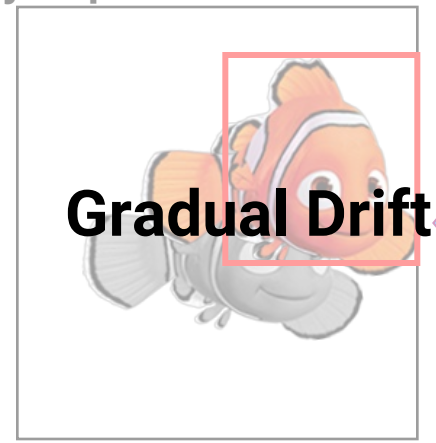
evaluate classifier on sub-patches



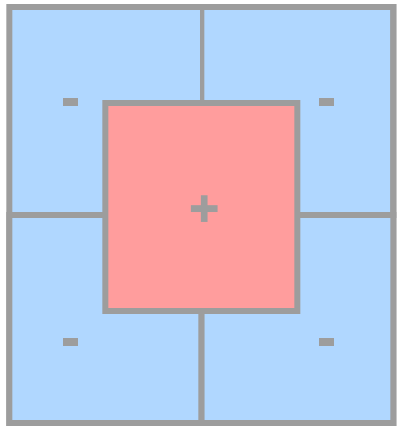
create confidence map



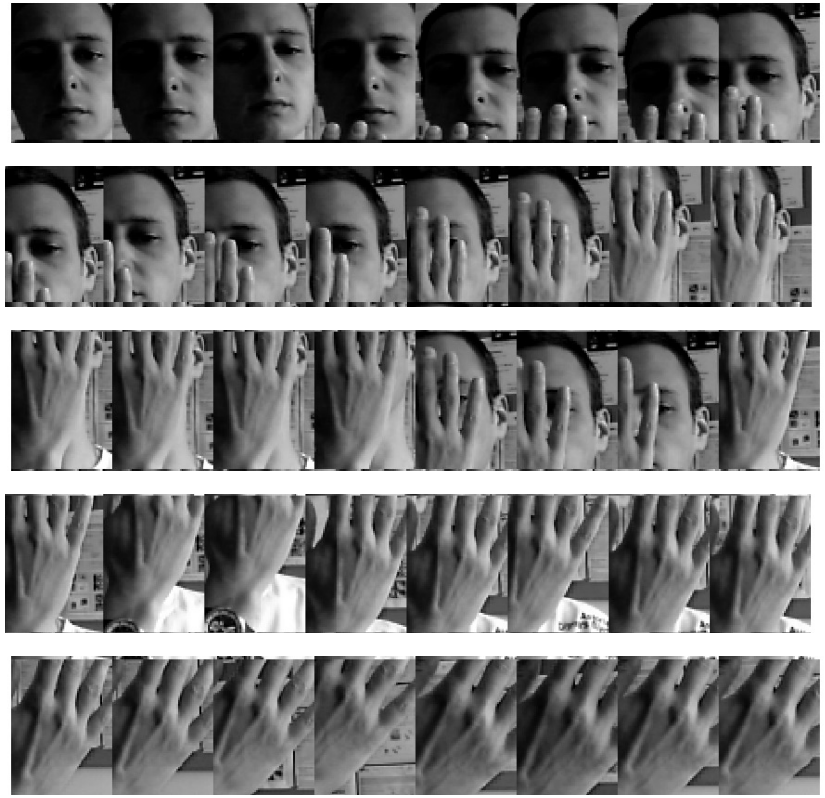
analyze map and set new object position



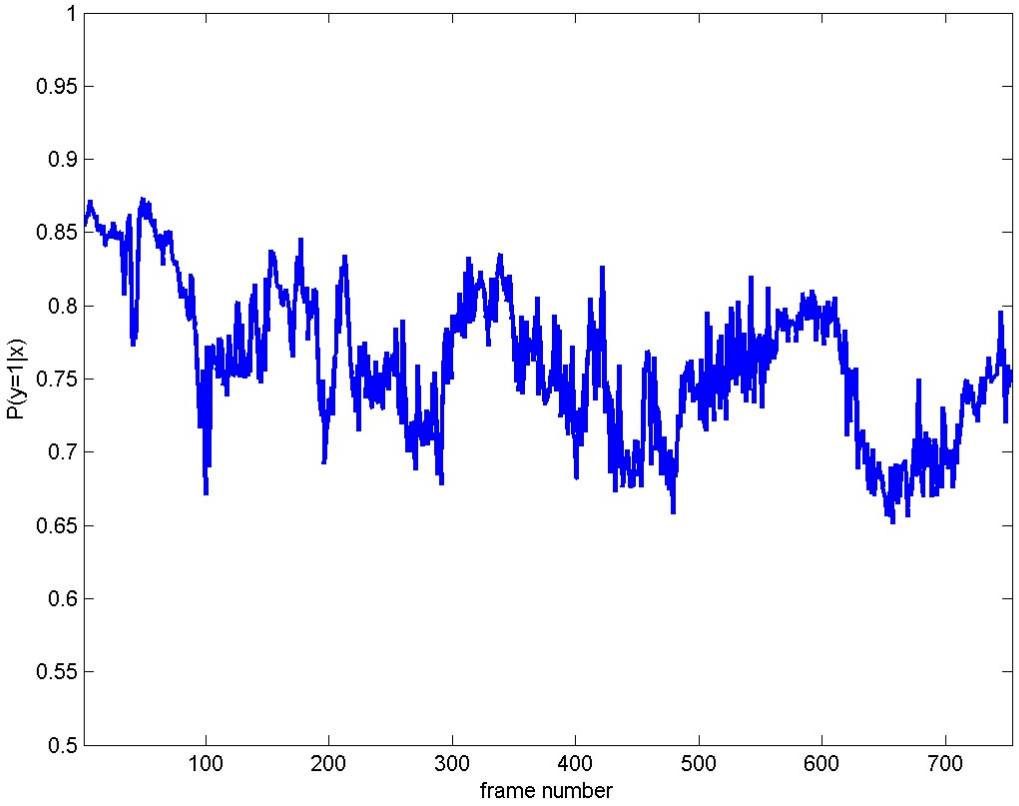
update classifier (tracker)



Tracked Patches

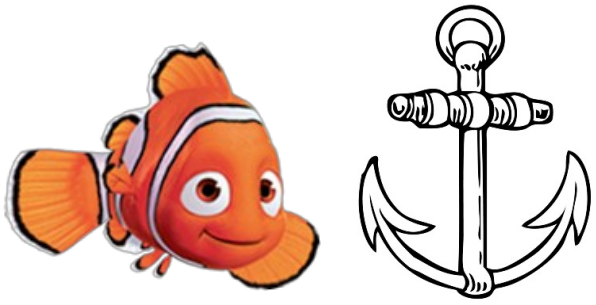


Confidence



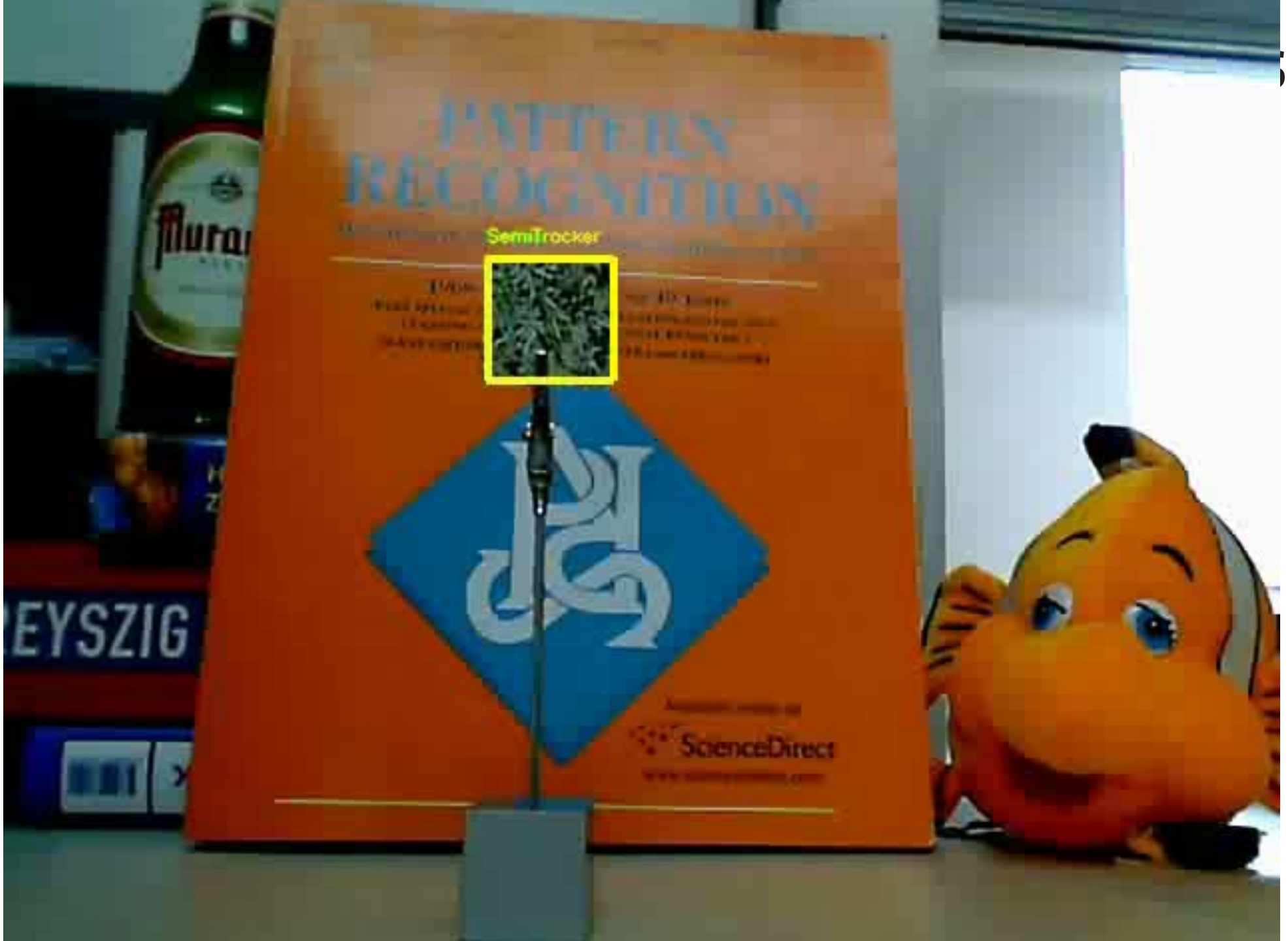
To Avoid Drift

- Only thing we are sure about the object is its initial model (e.g. appearance in first frame)
- We can “anchor” / correct our model with this information, in order to help avoid drift



Fix (initial) Model

Current Model


PATTERN
RECOGNITION
SemiTracker
ScienceDirect

Tracking

- Track a point
- Track a bigger box
- Track by detection
- Online learning
- **Motion**
- Multiple object tracking
- 3D object tracking

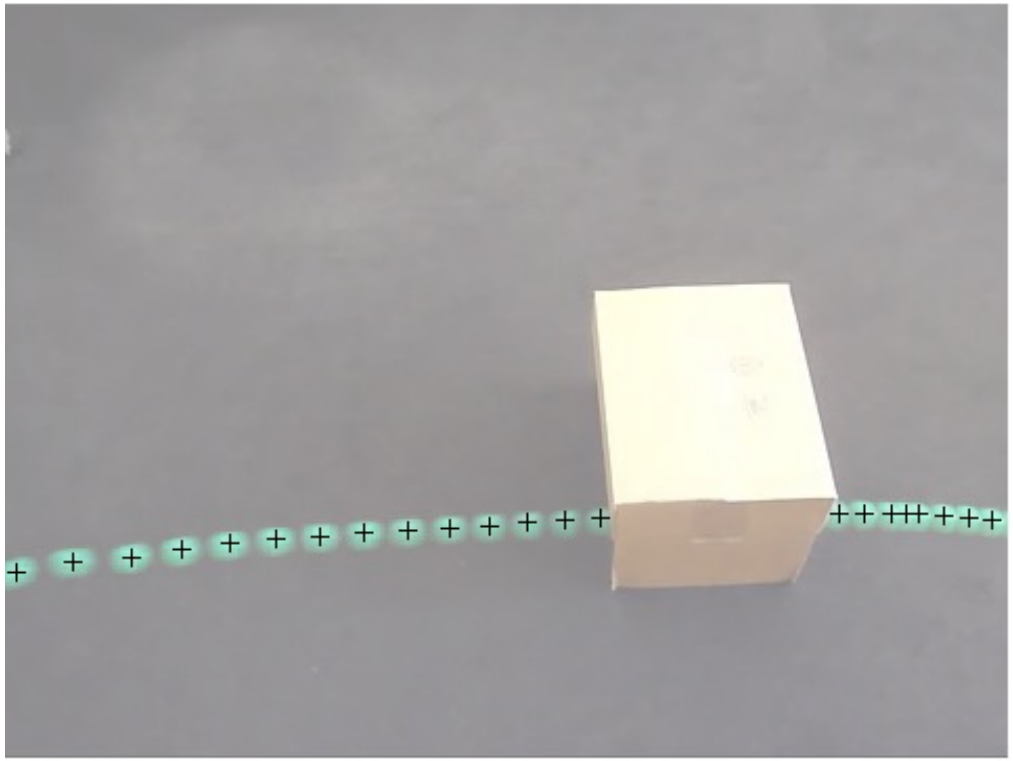
Location Information

A simple constant-velocity heuristics can go a long way, especially when the camera is stationary.

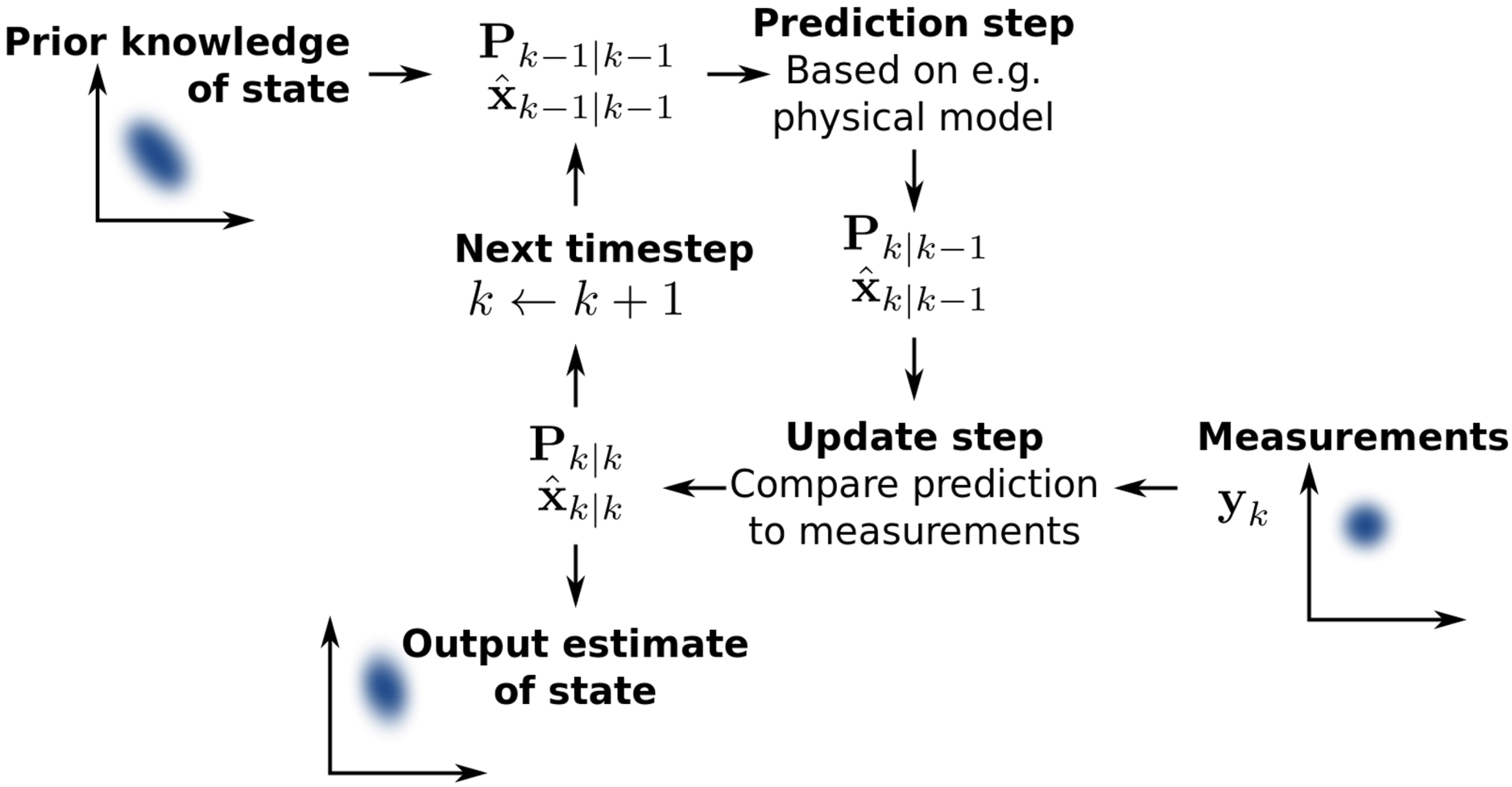
More Complicated Motion

Temporal Filtering/Predictions

- To predict location
- To reduce noise
- To disambiguate multiple objects



Kalman Filter



An ETH Legacy

http://www.ethlife.ethz.ch/archive_articles/091008_kalman_per

08.10.2009
Rudolf Kalman, ETH-Zurich emeritus professor of mathematics, is awarded the National Medal of Science by Barack Obama – one of the highest accolades for researchers in the USA.

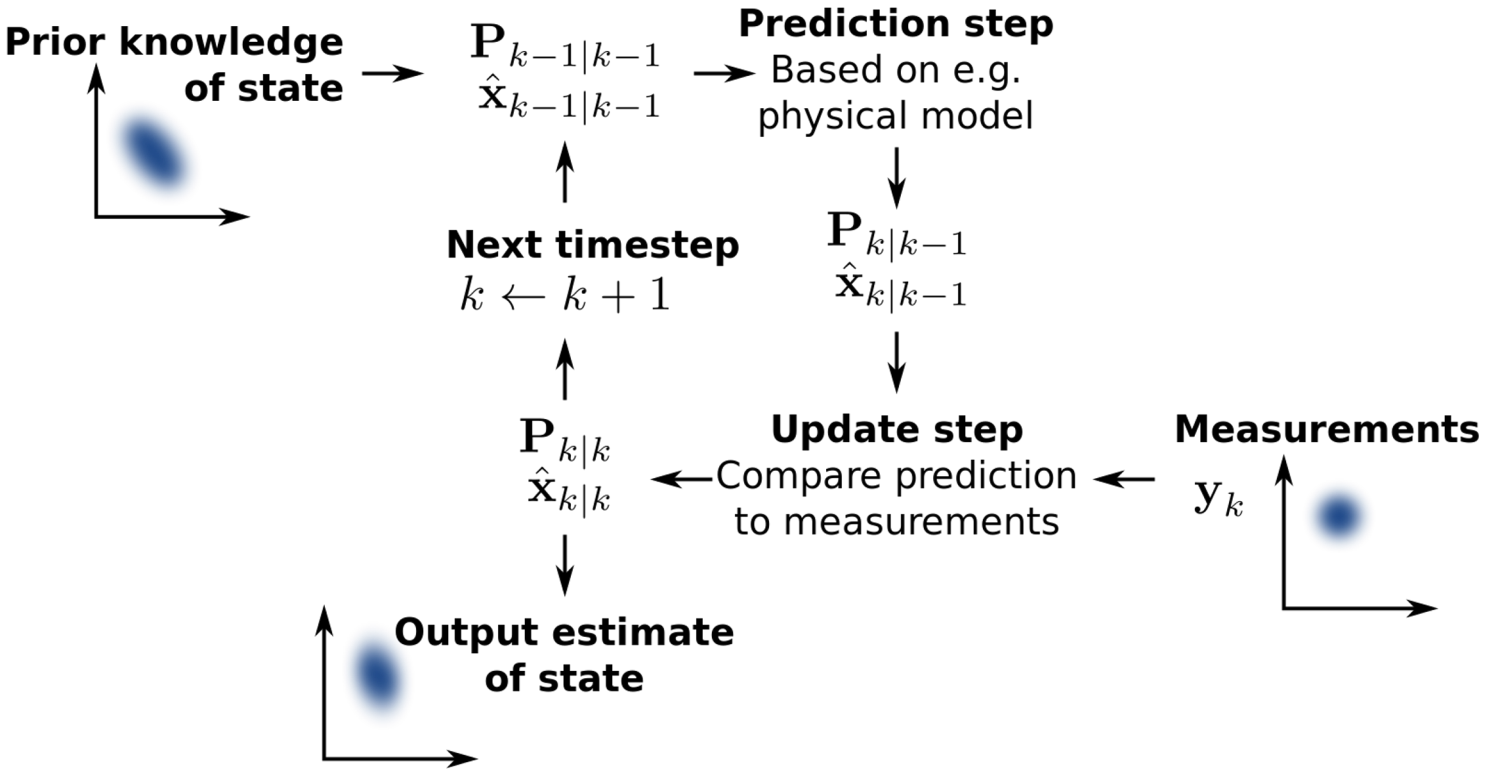
An Example

$$\mathbf{x} = [u, v, s, r, \dot{u}, \dot{v}, \dot{s}]^T$$

u, v: bounding box coordinate
s: area of the bounding box
r: aspect ratio

Simple online and realtime tracking, 2016

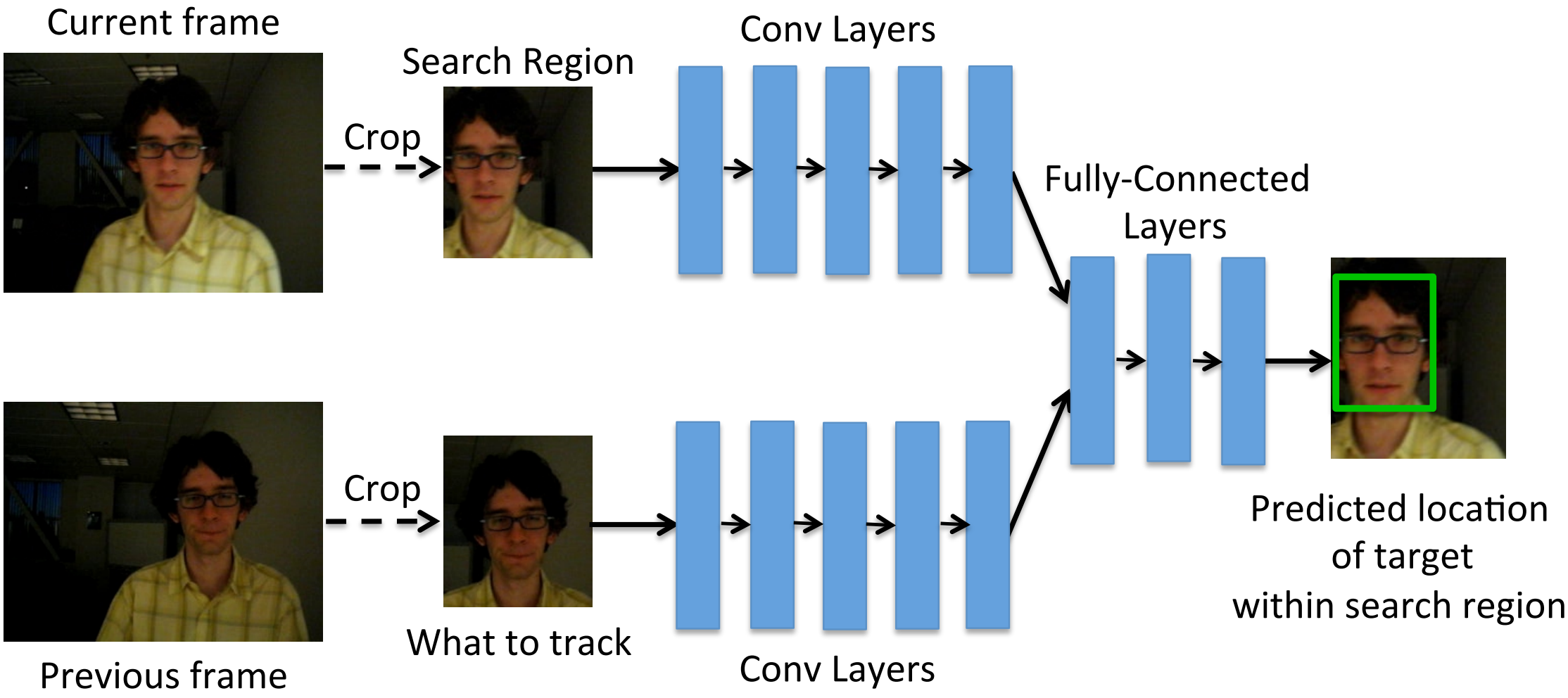
Kalman Filter



$$\mathbf{x} = [u, v, s, r, \dot{u}, \dot{v}, \dot{s}]^T$$

What can you do with a training dataset?

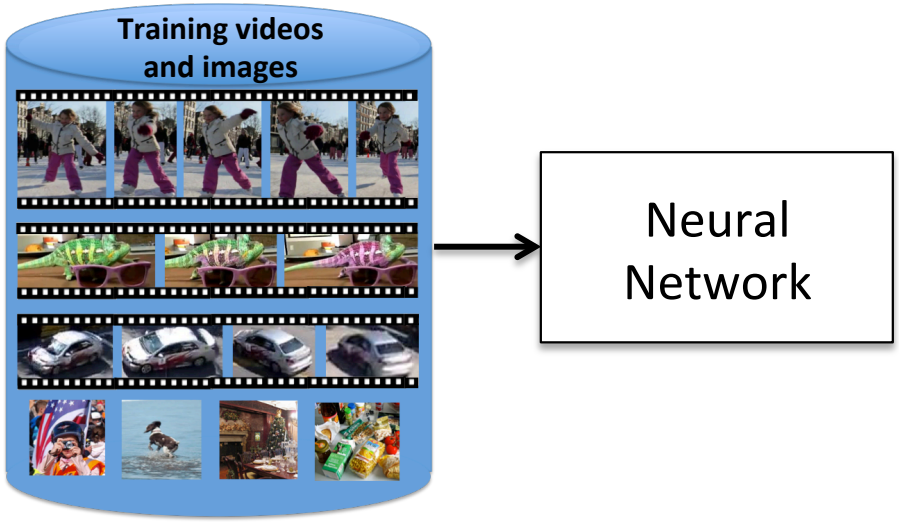
Predict the location



Learning to Track at 100 FPS with Deep Regression Networks, ECCV 2016

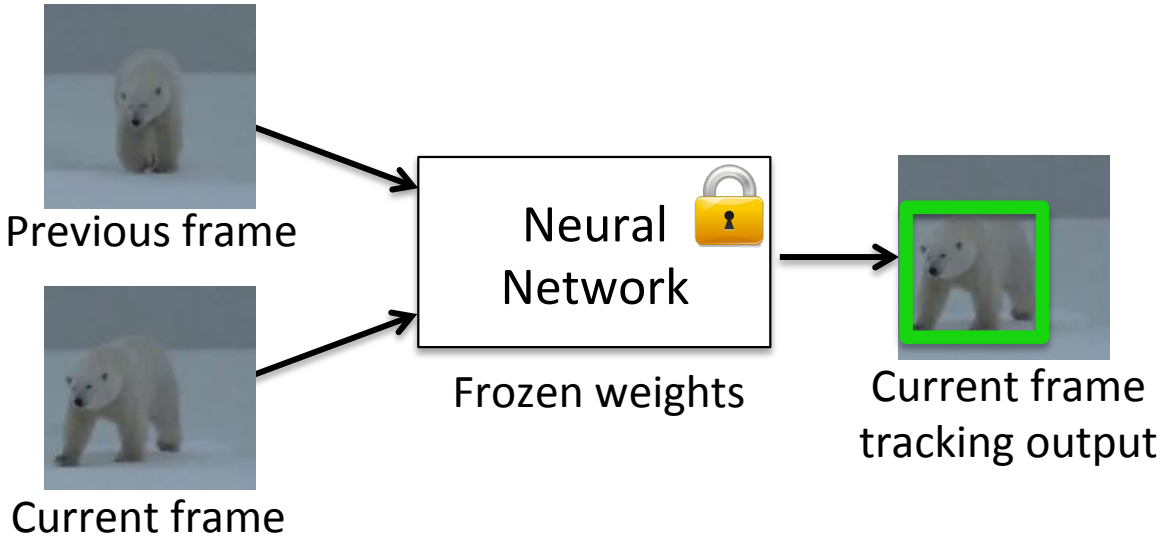
Predict the location

Training:



Network learns generic object tracking

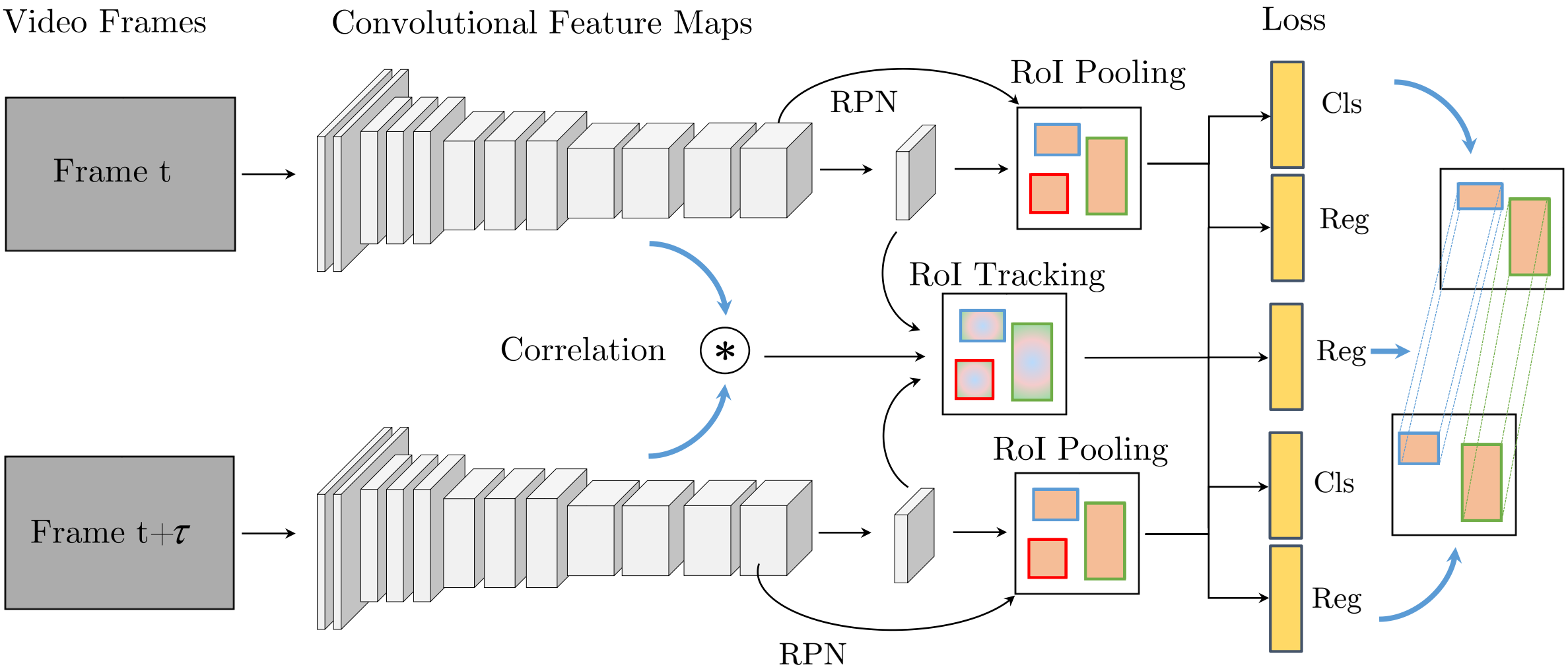
Test:



Network tracks novel objects (no finetuning)

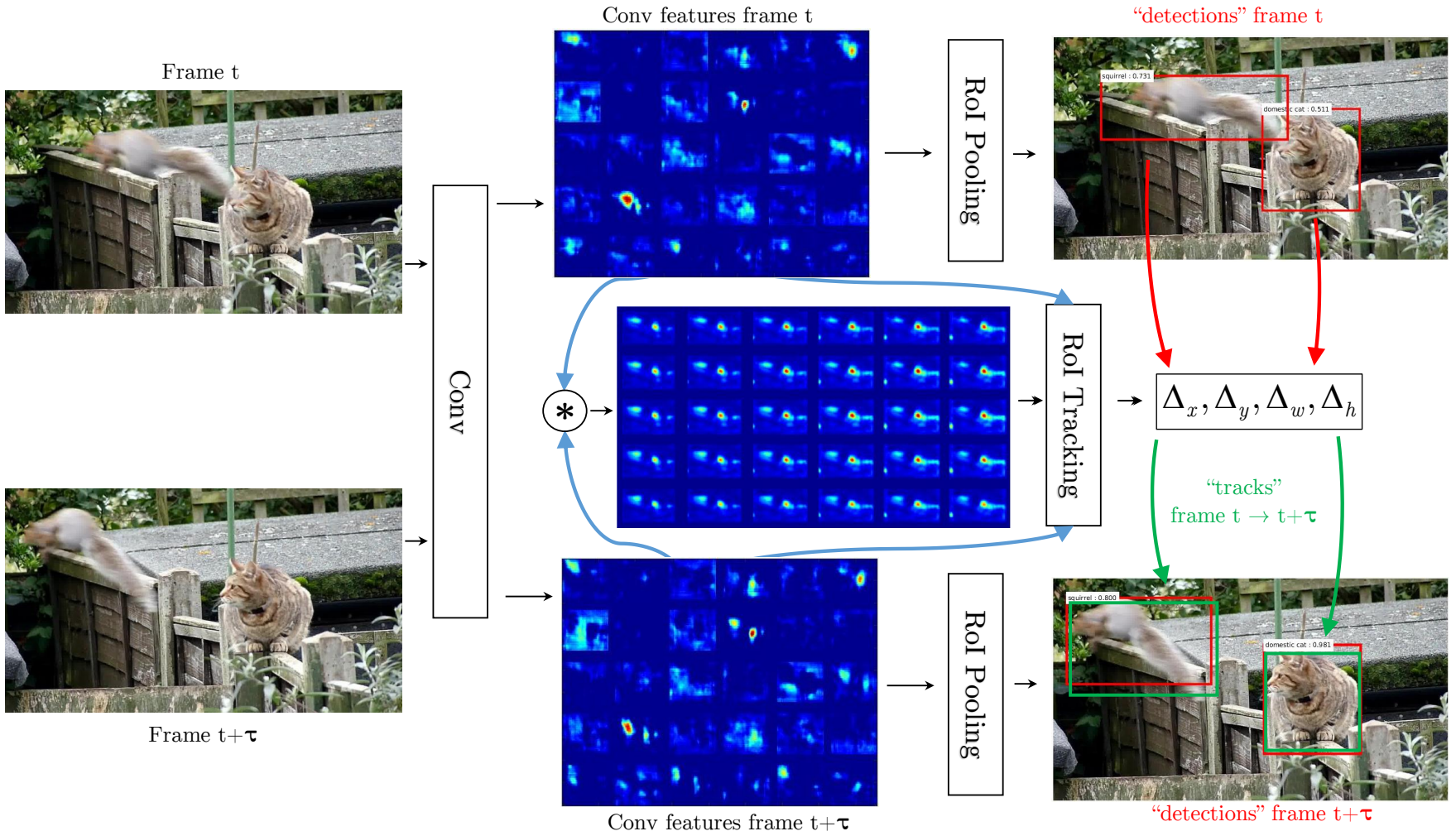
Learning to Track at 100 FPS with Deep Regression Networks, ECCV 2016

Combining Tracking and Detection



Detect to Track and Track to Detect, ICCV 2017

Combining Tracking and Detection



Detect to Track and Track to Detect, ICCV 2017

Tracking

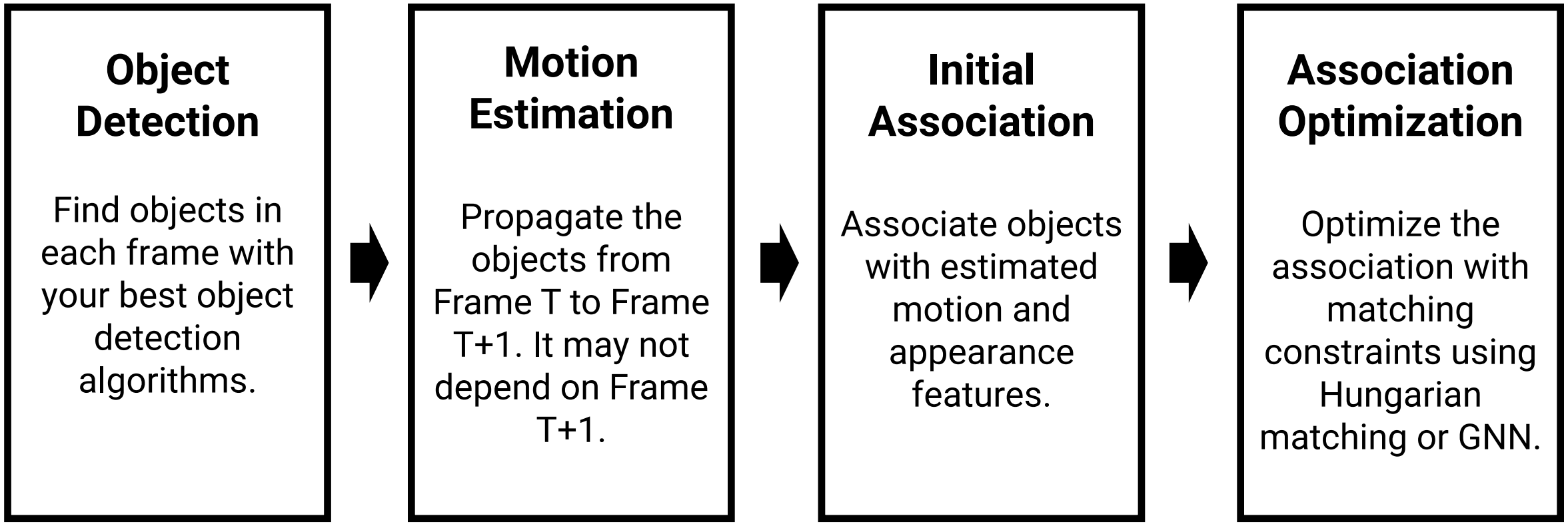
- Track a point
  - Consider its neighbors to avoid aperture problem
- Track a bigger box
  - We should consider image warping
- Track by detection
  - Use features to search for the object
  - Known object category
- Online learning
- Motion
- **Multiple object tracking**
- 3D object tracking

Multiple Object Tracking

Multiple Object Tracking

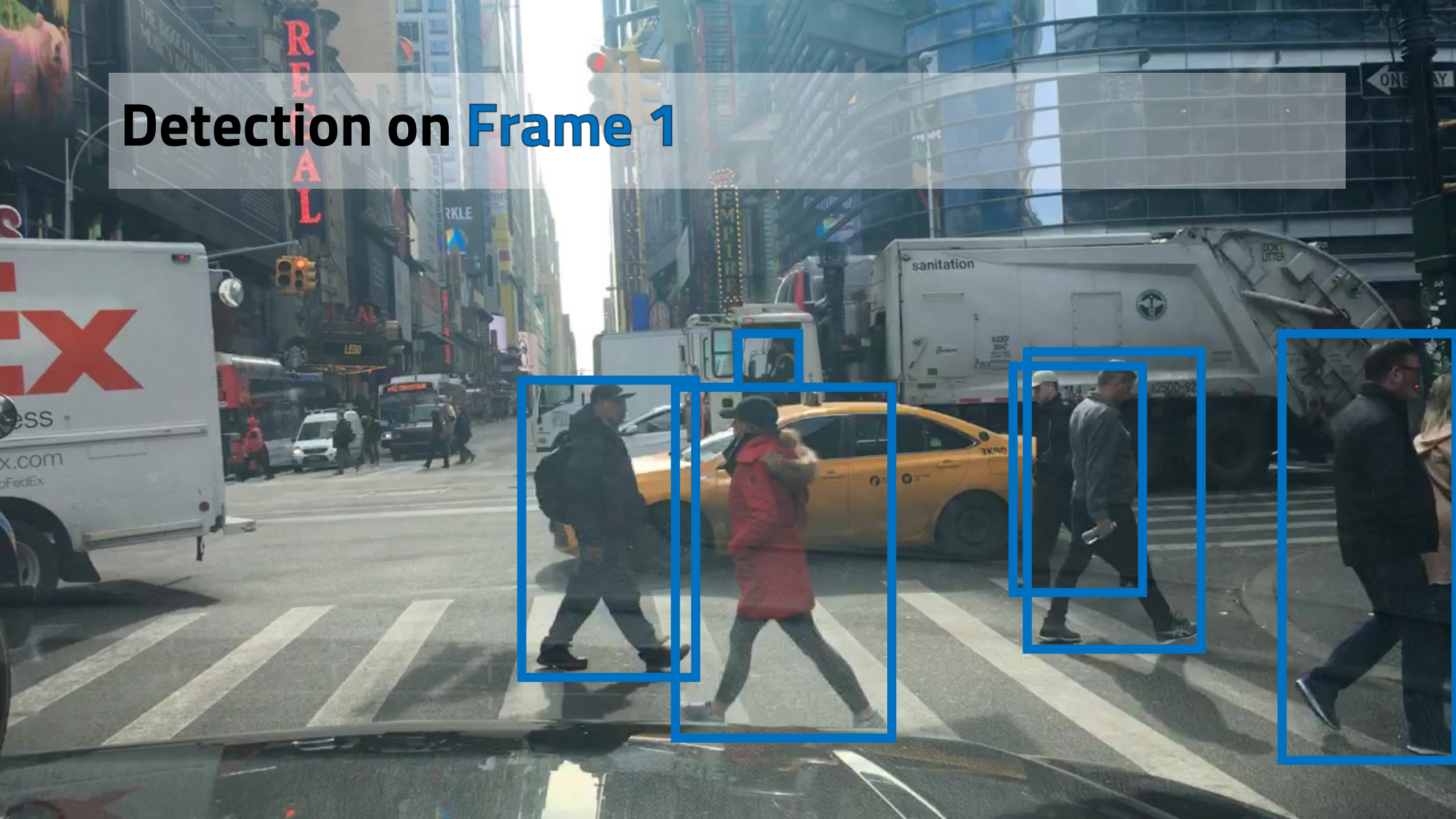
- Tracking known objects
- Many objects in the same frame
- Very challenging in the natural scenes
- Leveraging supervised learning for detection and association

Multiple Object Tracking

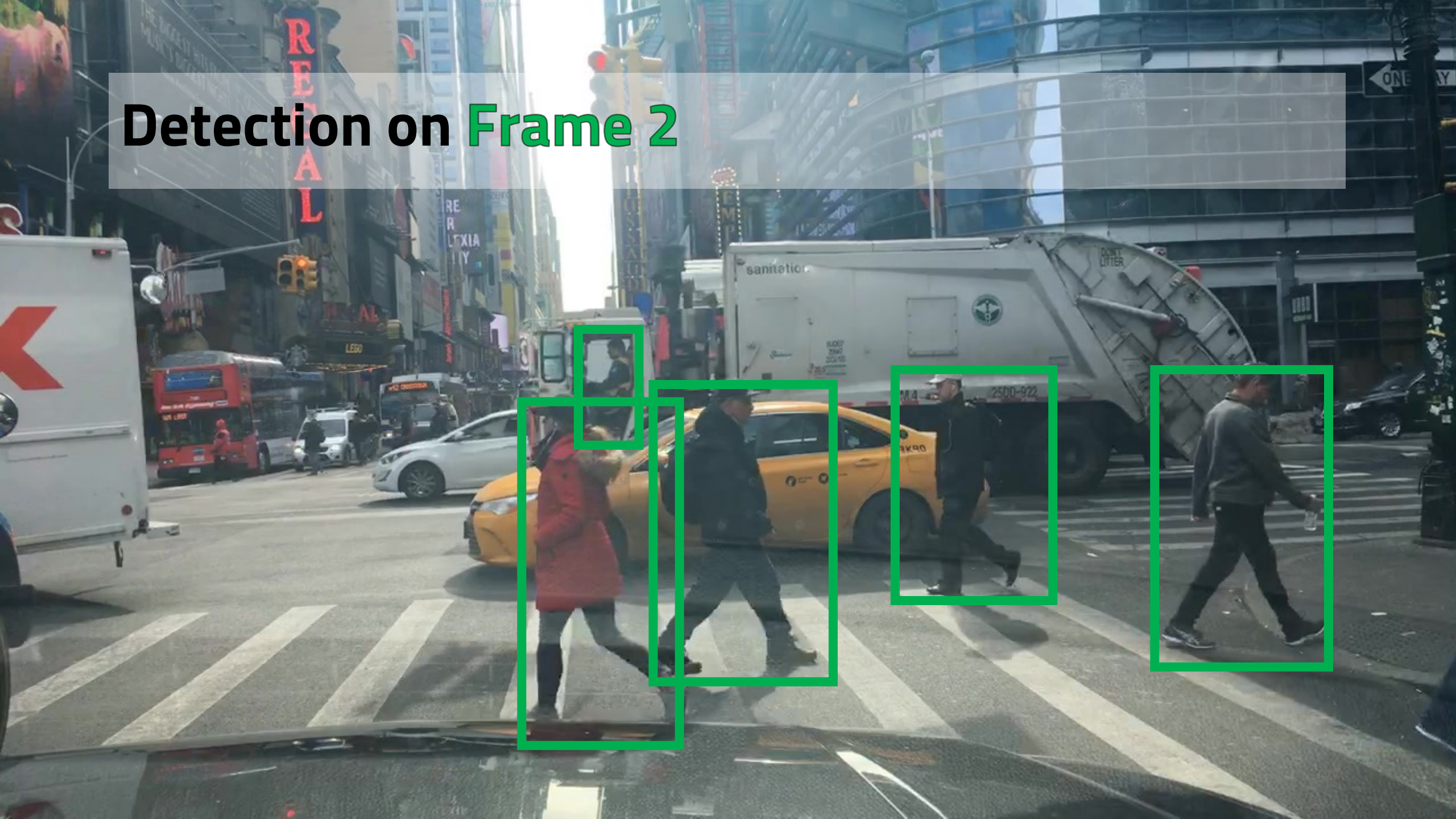


An Example

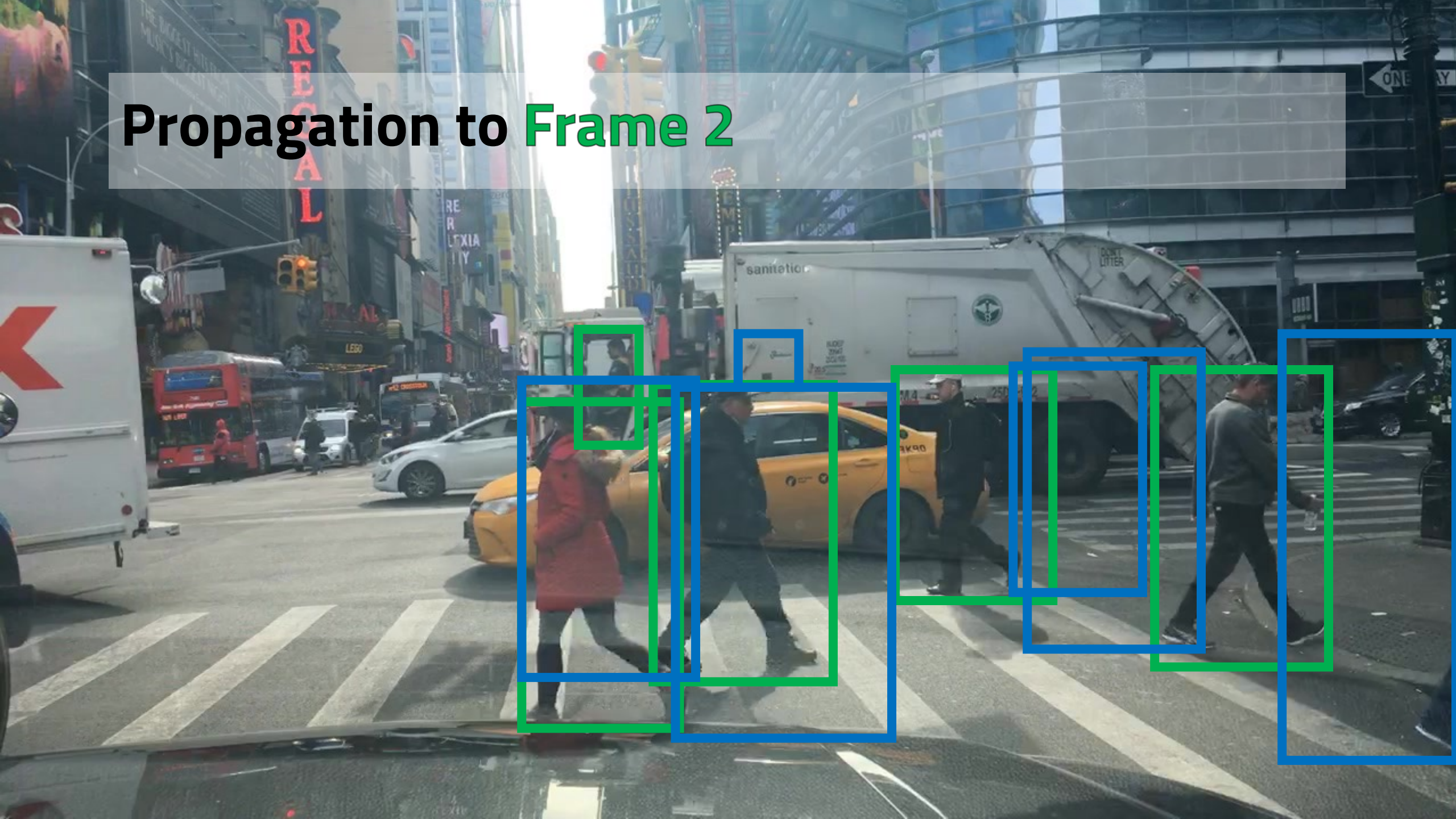
Detection on Frame 1



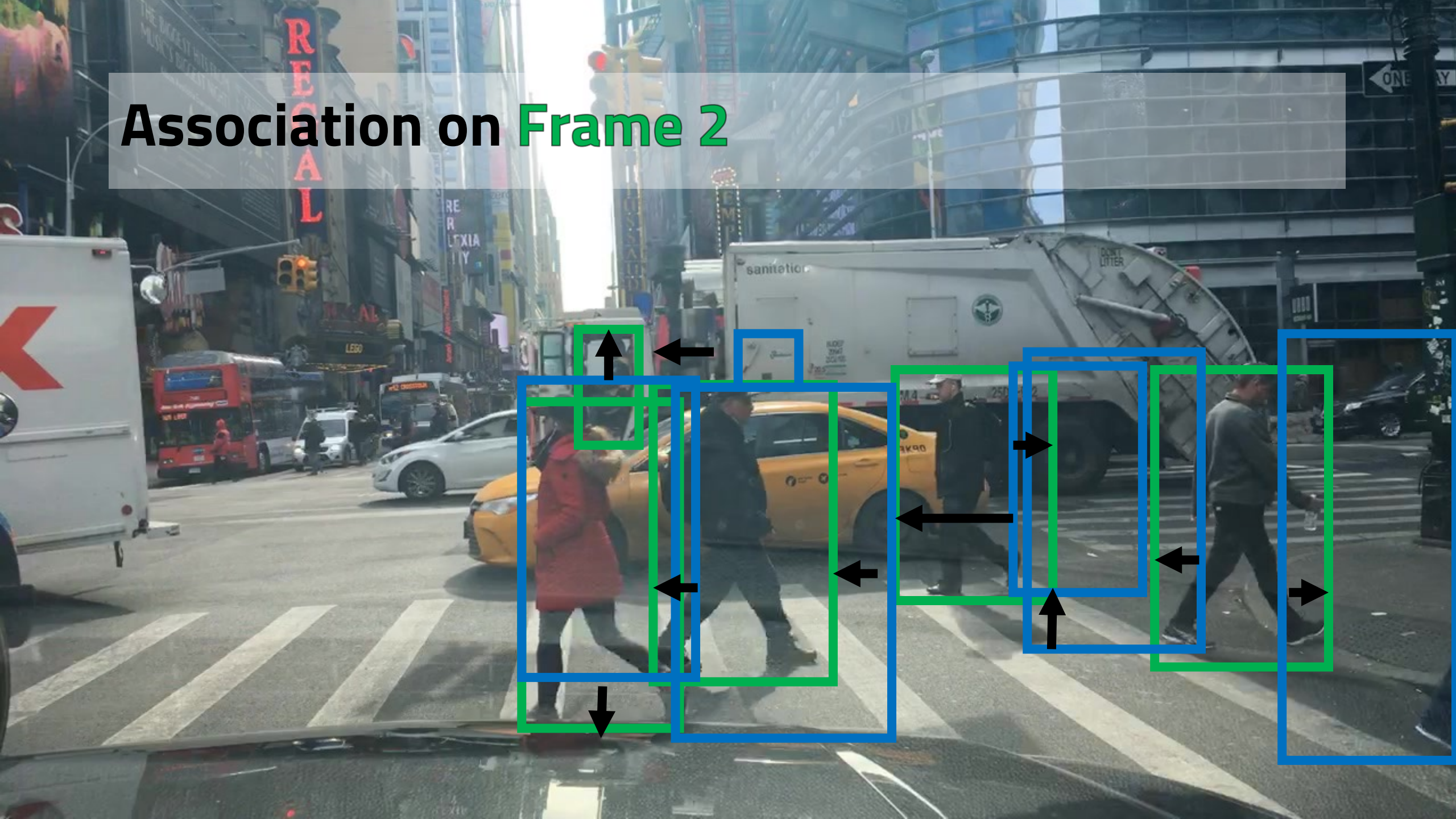
Detection on Frame 2



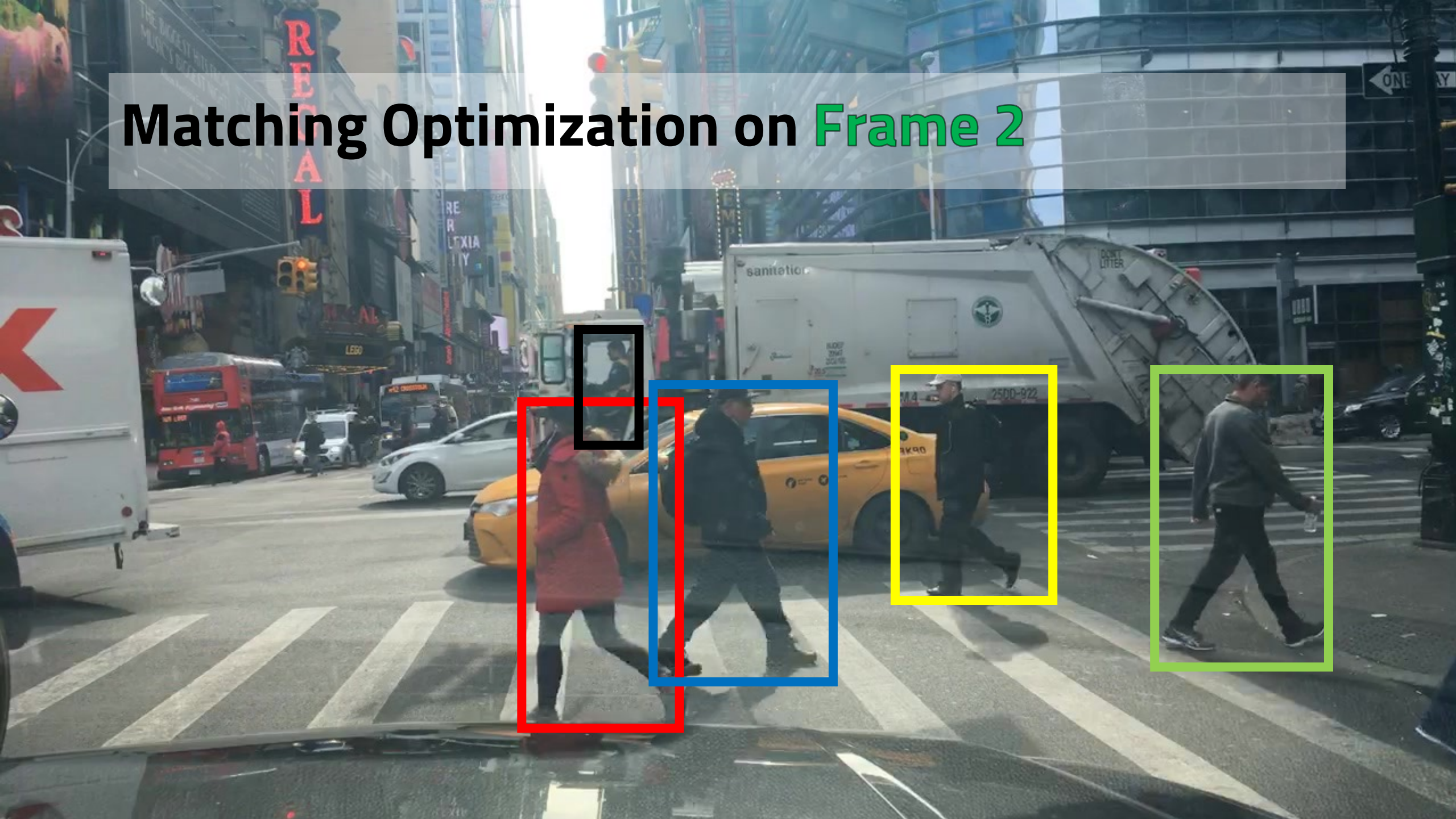
Propagation to Frame 2



Association on Frame 2



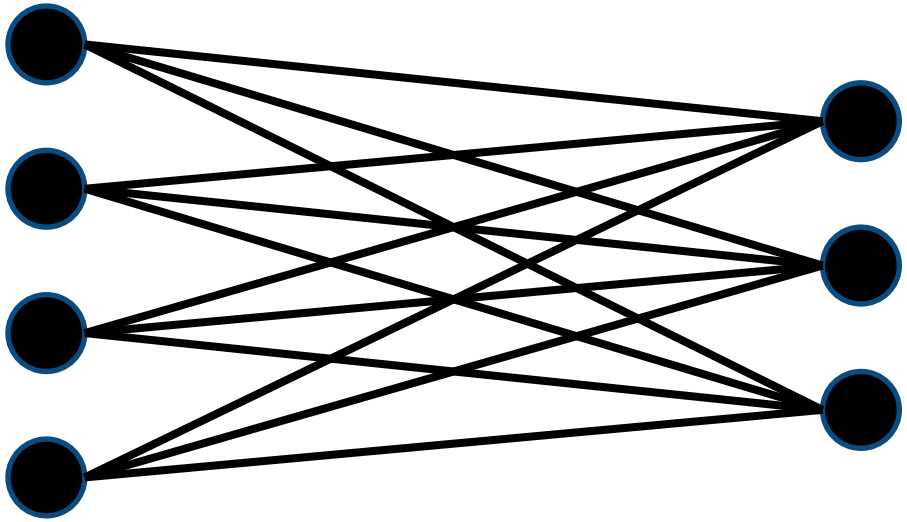
Matching Optimization on Frame 2



Appearance Modeling

- We can get a set of matching candidates from two frames
- It can be formulated as a bipartite matching problem

Similarity between each pair of candidates



Candidates on Frame 1 Candidates on Frame 2

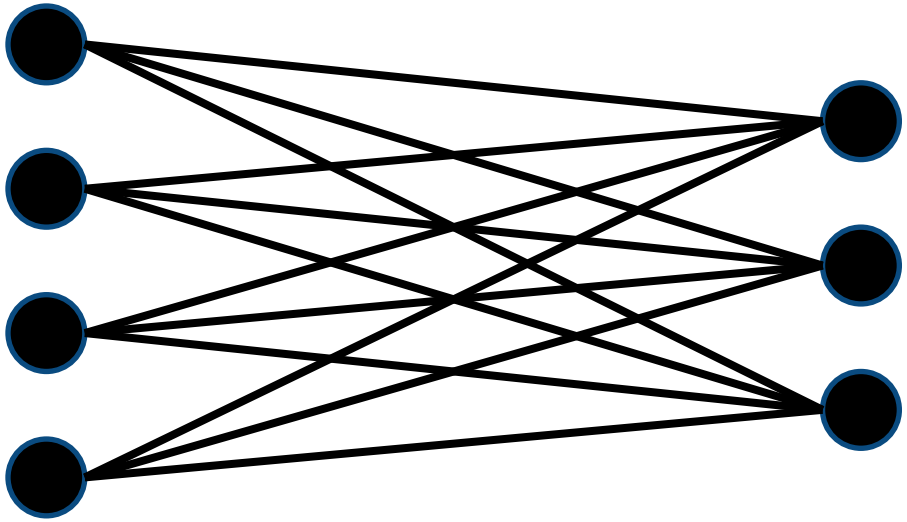
Simple online and realtime tracking with a deep association metric, ICIP 2017

Assignment Problem

Similarity / Affinity

- Box overlap, feature distance, social constraints, motion priors, etc.

Similarity between each pair of candidates

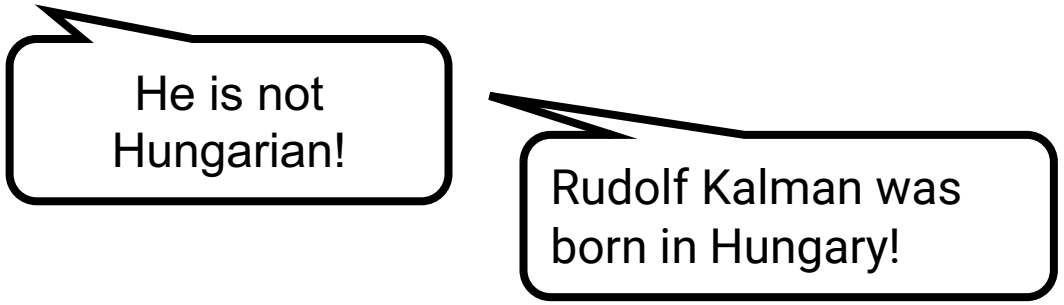


Candidates on Frame 1 Candidates on Frame 2

Simple online and realtime tracking with a deep association metric, ICIP 2017

Hungarian Algorithm

- A combinatorial optimization algorithm
- Developed and published in 1955 by Harold Kuhn
- Complexity O(n^3)

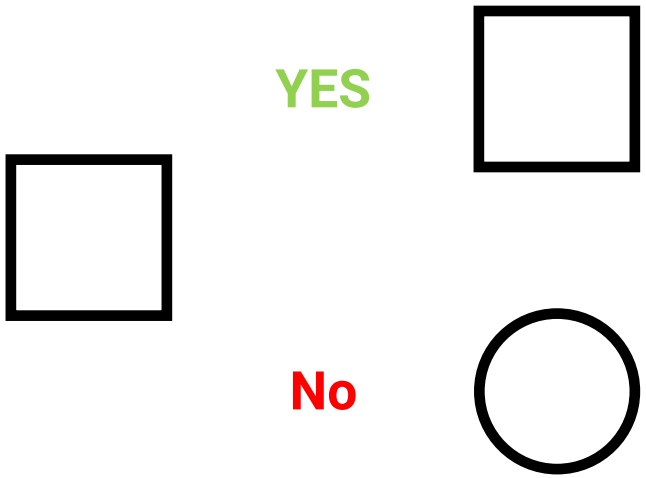


	Clean bathroom	Sweep floors	Wash windows
Paul	$2	$3	$3
Dave	$3	$2	$3
Chris	$3	$3	$2

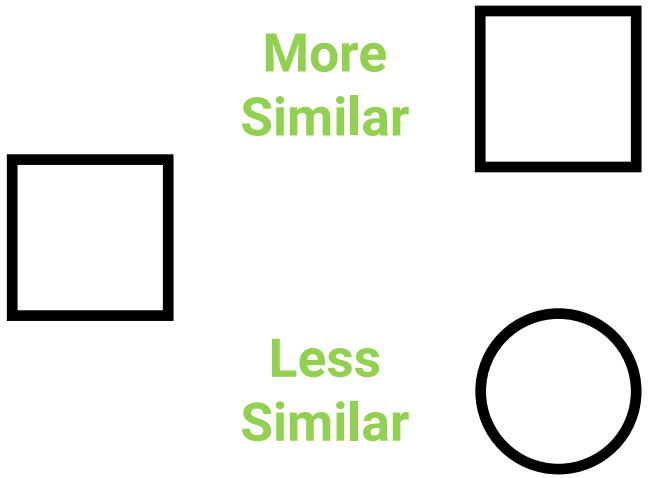
Simple online and realtime tracking with a deep association metric, ICIP 2017

Similarity Learning

- Learning appearance feature embeddings and measure for similarity
- Binary or triplet loss



Binary



Triplet

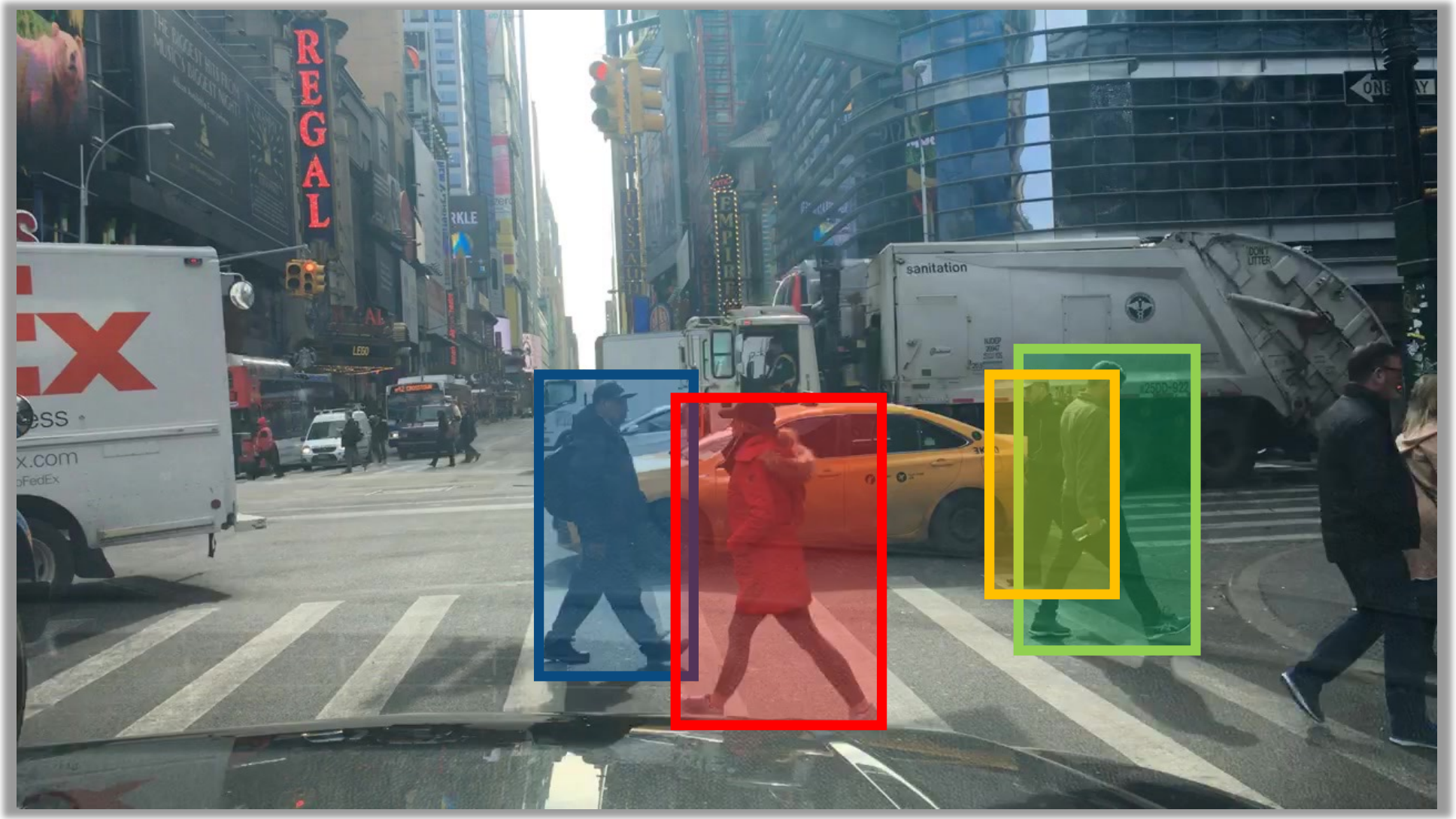
Simple online and realtime tracking with a deep association metric, ICIP 2017

DeepSort

Simple online and realtime tracking with a deep association metric, ICIP 2017

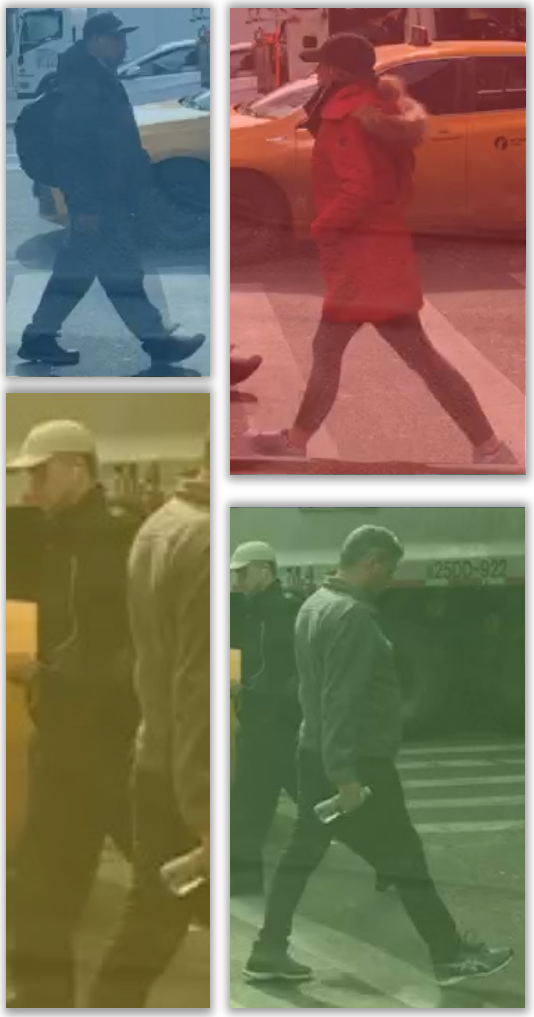
Let’s play a game

Let's play a game



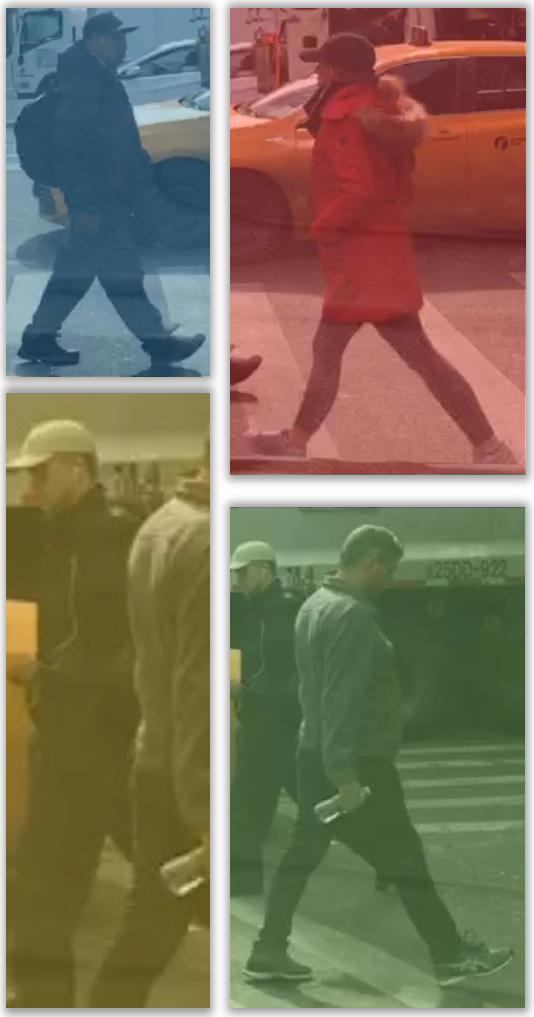
ETH zürich VIS Visual Intelligence and Systems

Let's play a game



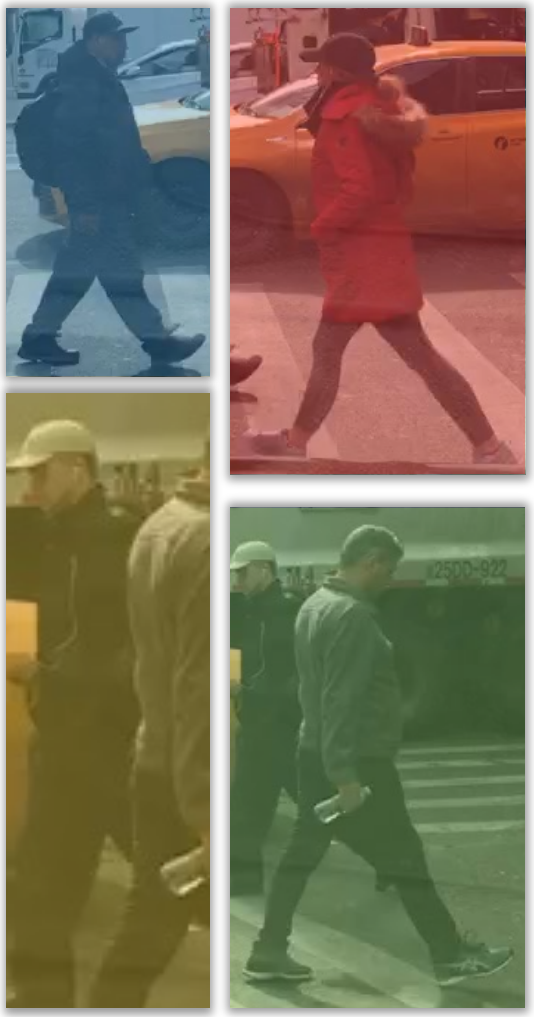
ETH zürich VIS Visual Intelligence and Systems

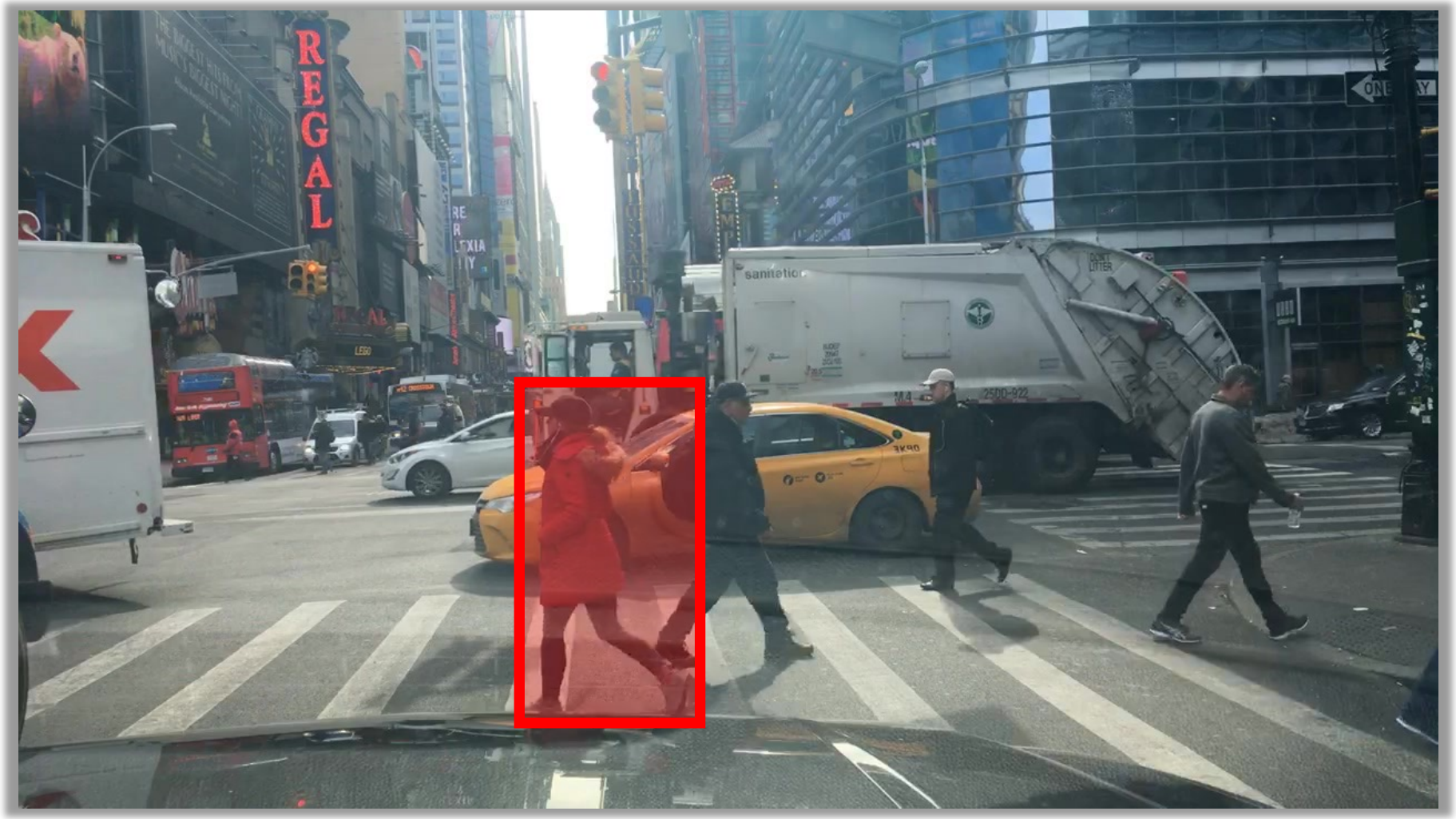
Let's play a game



ETH zürich VIS Visual Intelligence and Systems

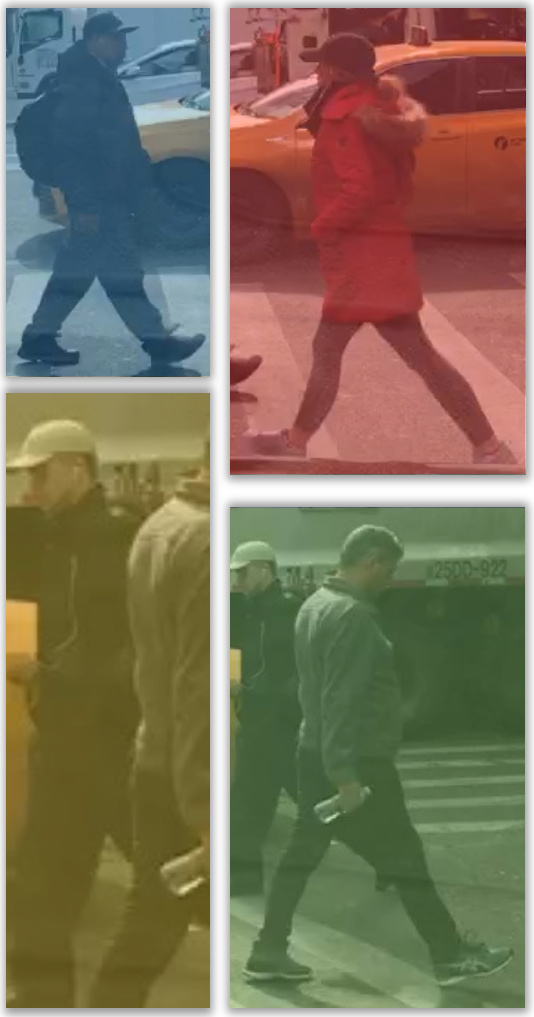
Let's play a game

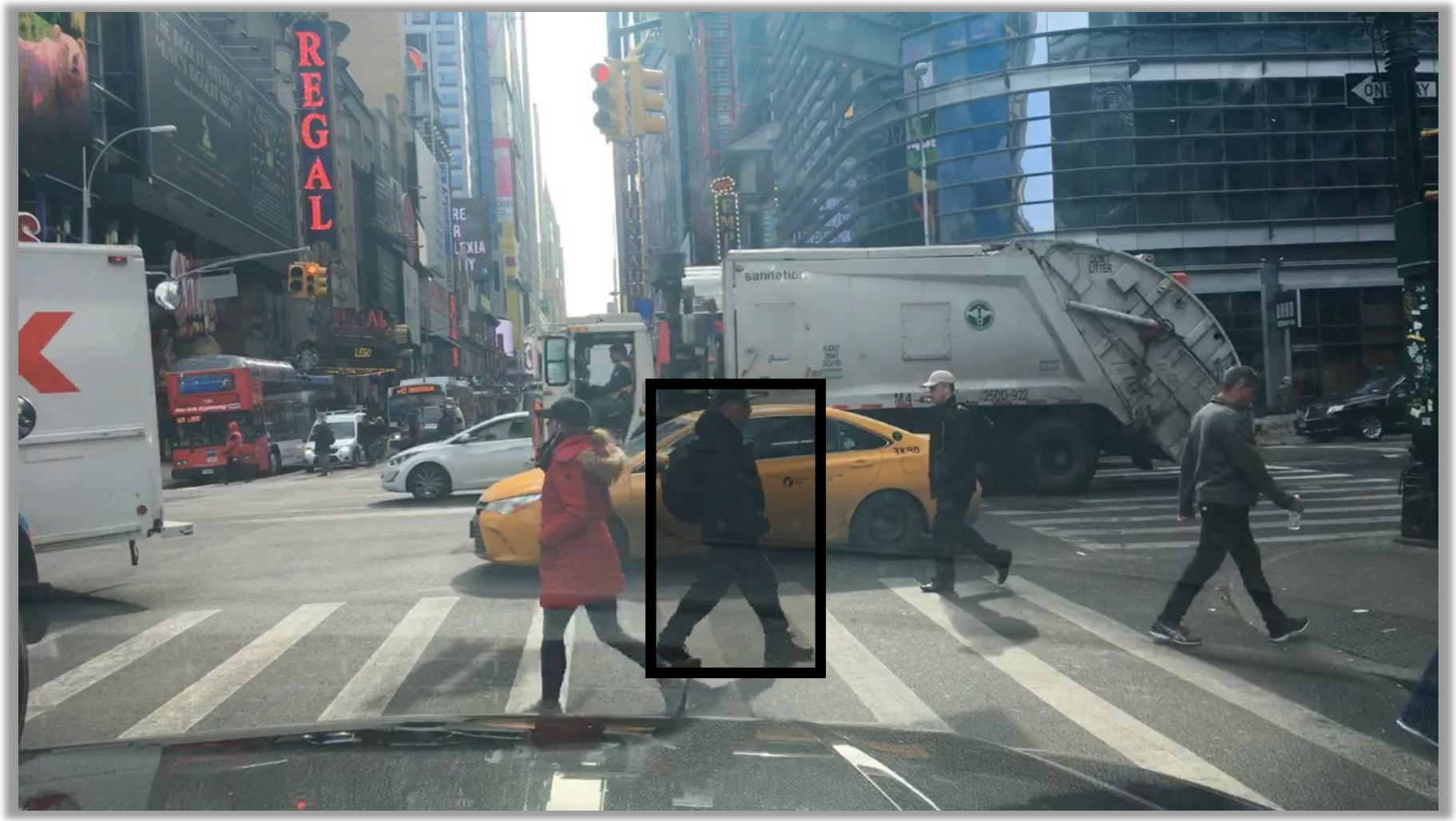




ETH zürich VIS Visual Intelligence and Systems

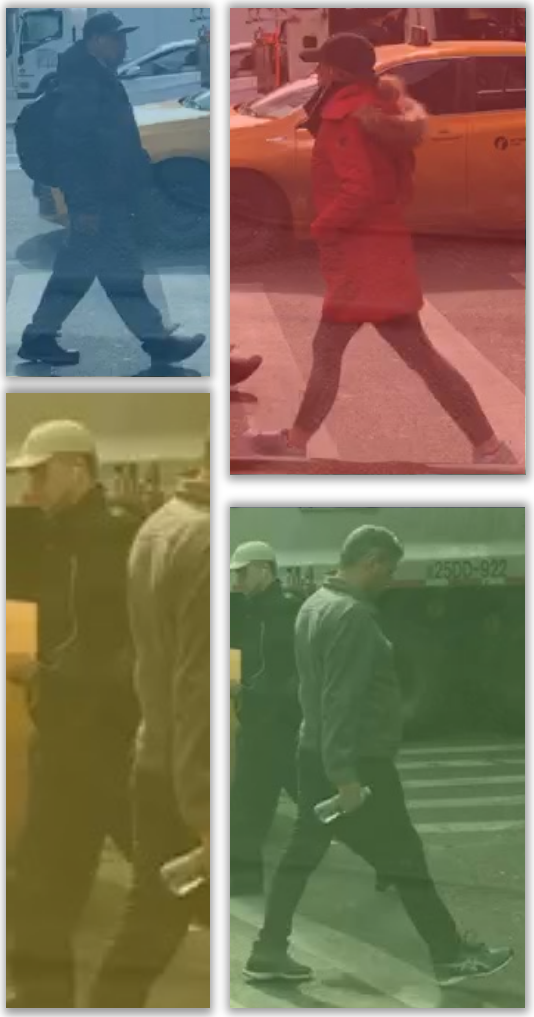
Let's play a game

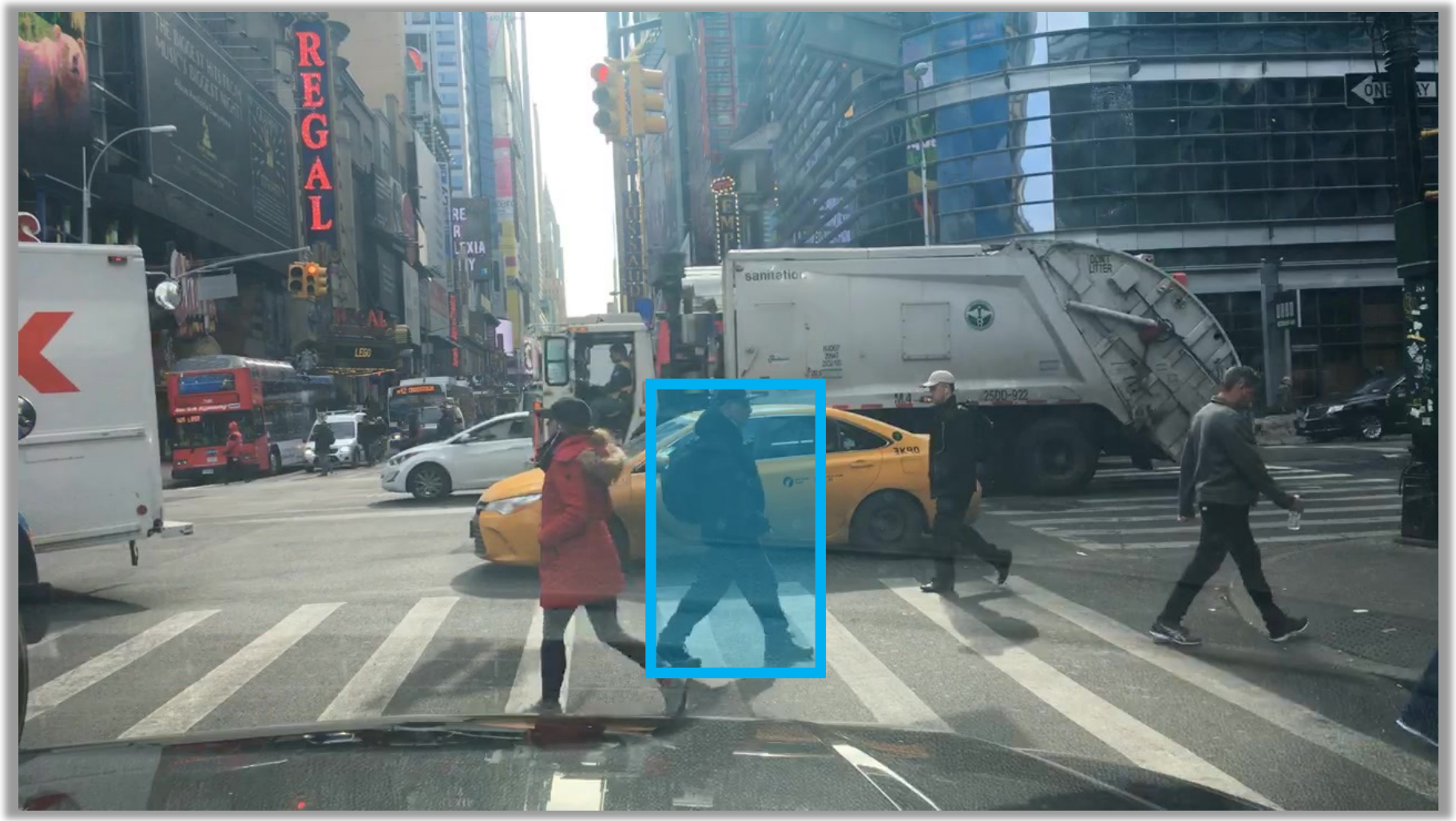




ETH zürich VIS Visual Intelligence and Systems

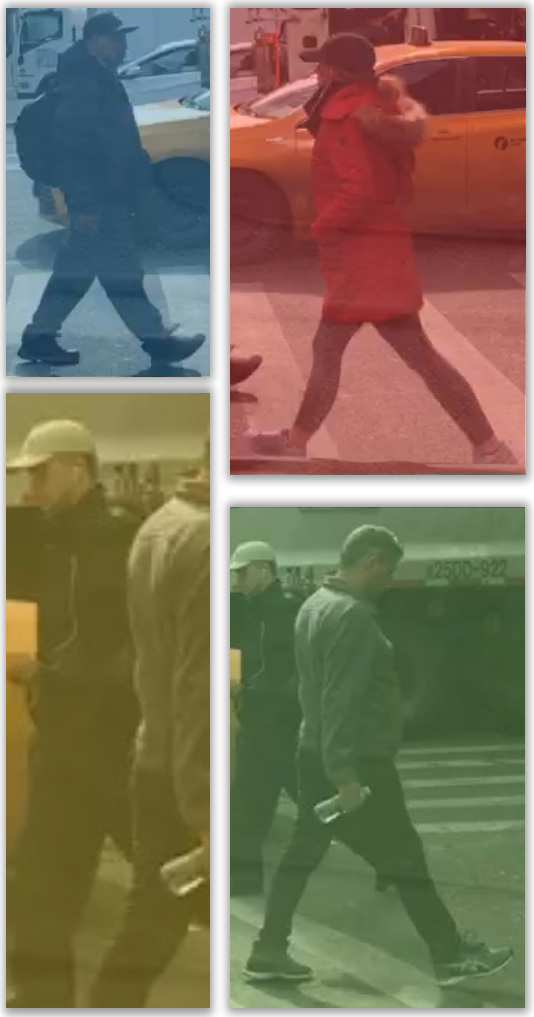
Let's play a game

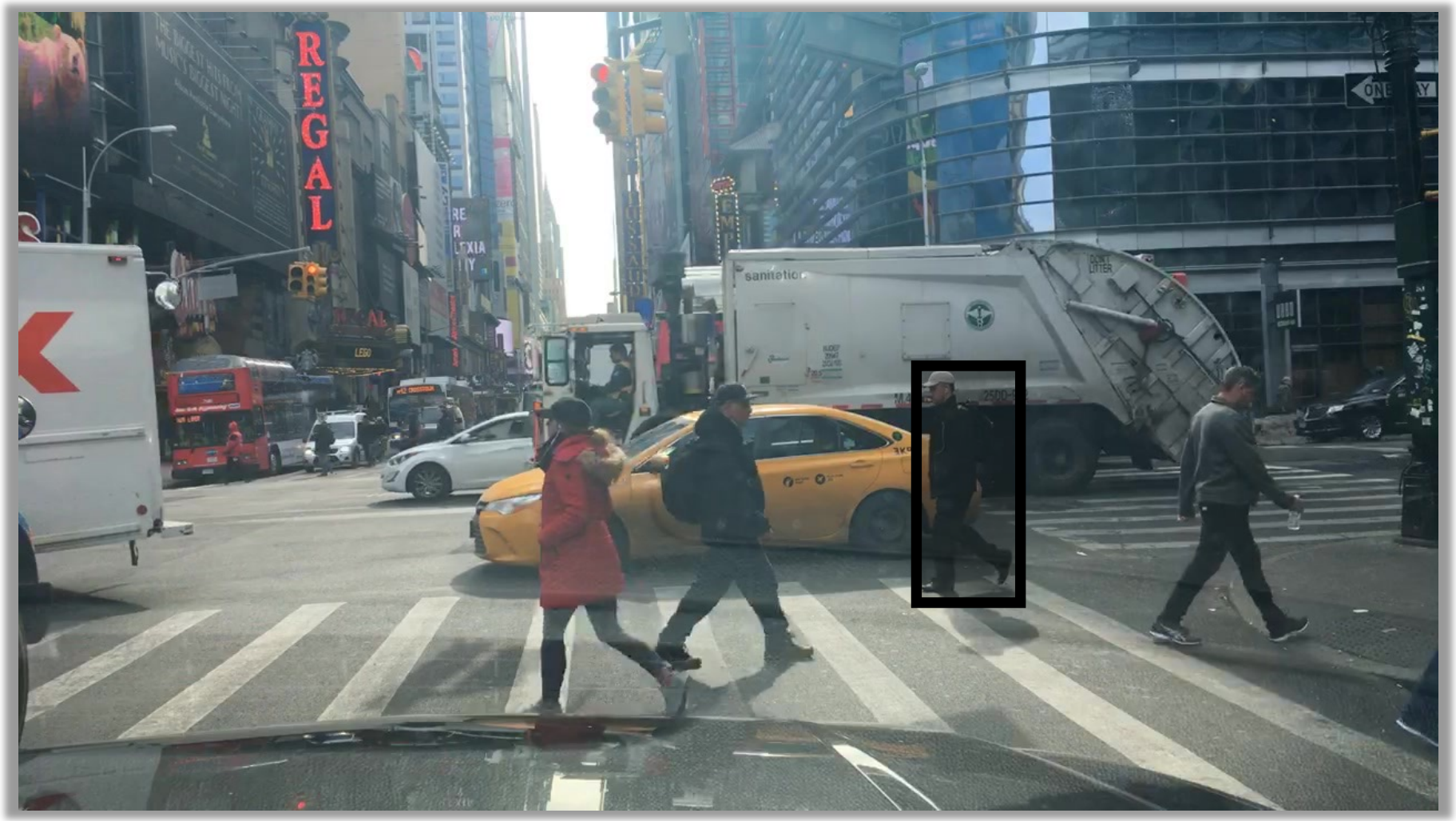




ETH zürich VIS Visual Intelligence and Systems

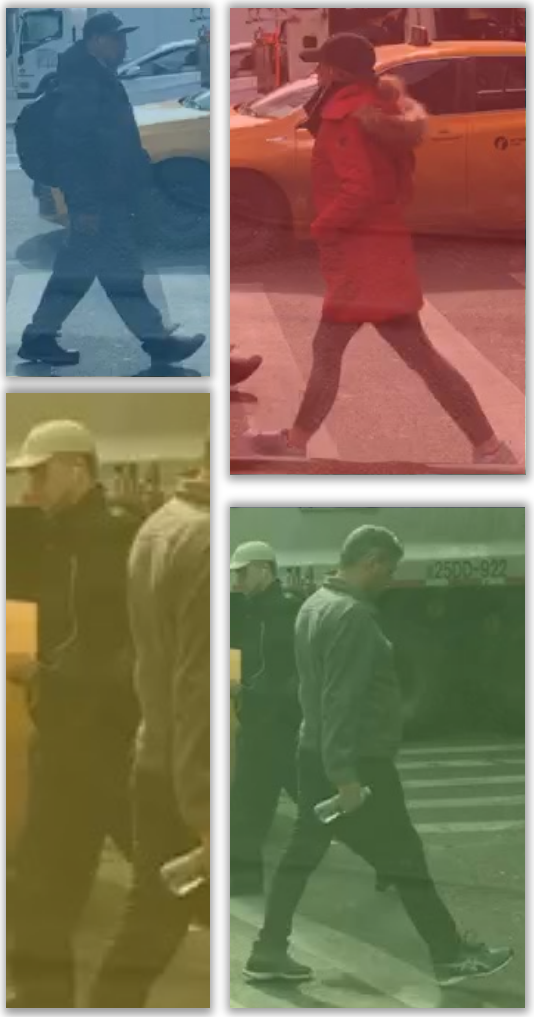
Let's play a game

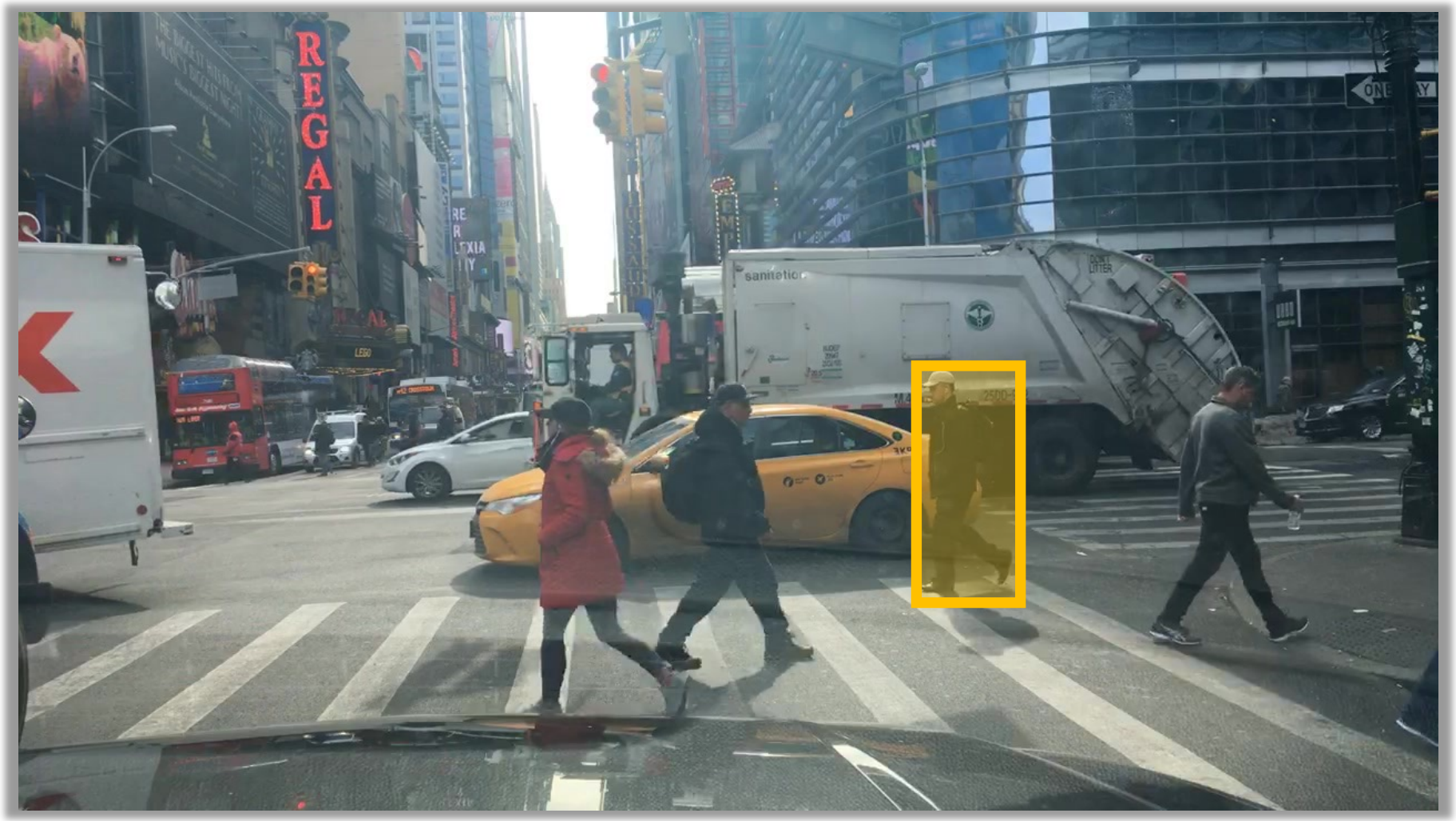




ETH zürich VIS Visual Intelligence and Systems

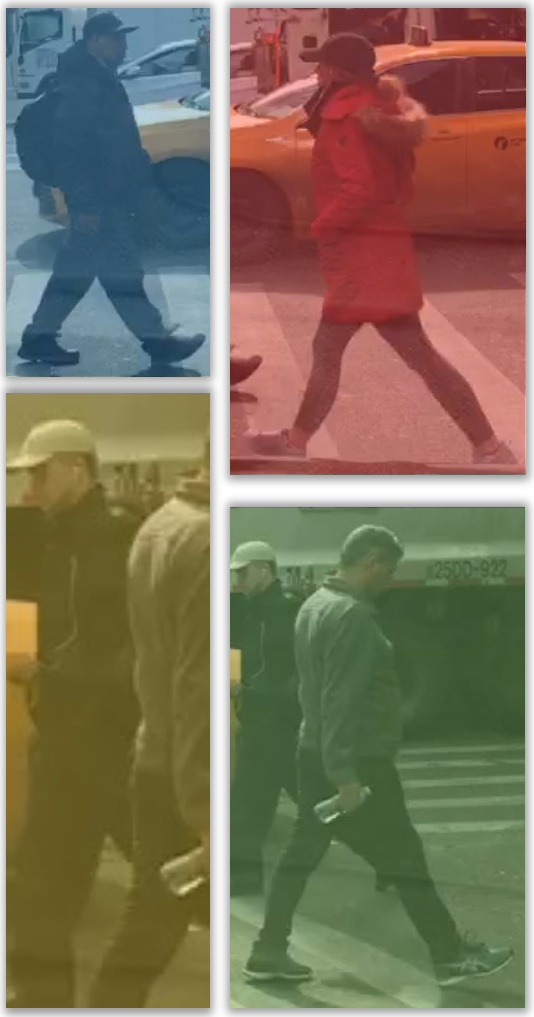
Let's play a game

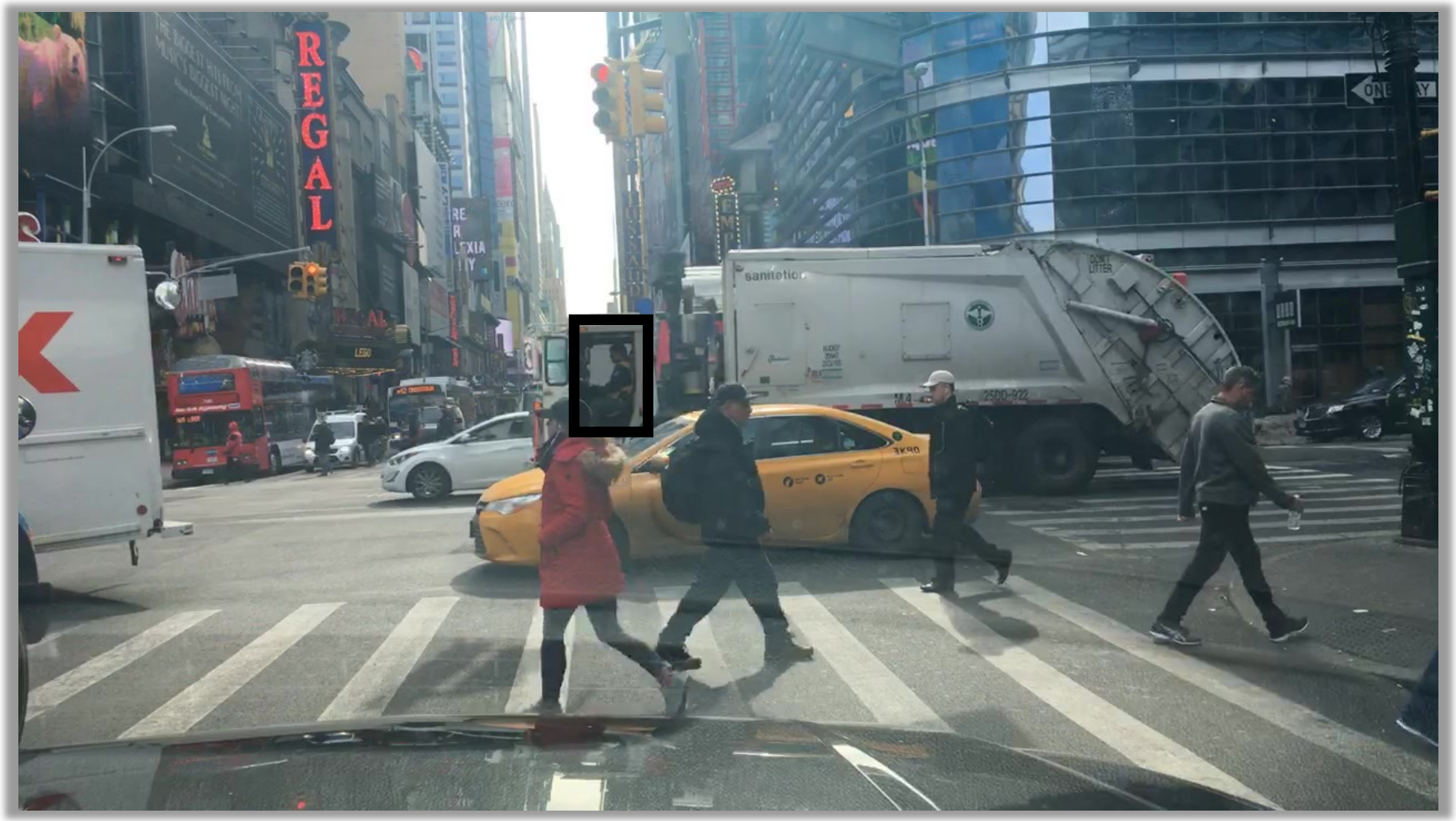




ETH zürich VIS Visual Intelligence and Systems

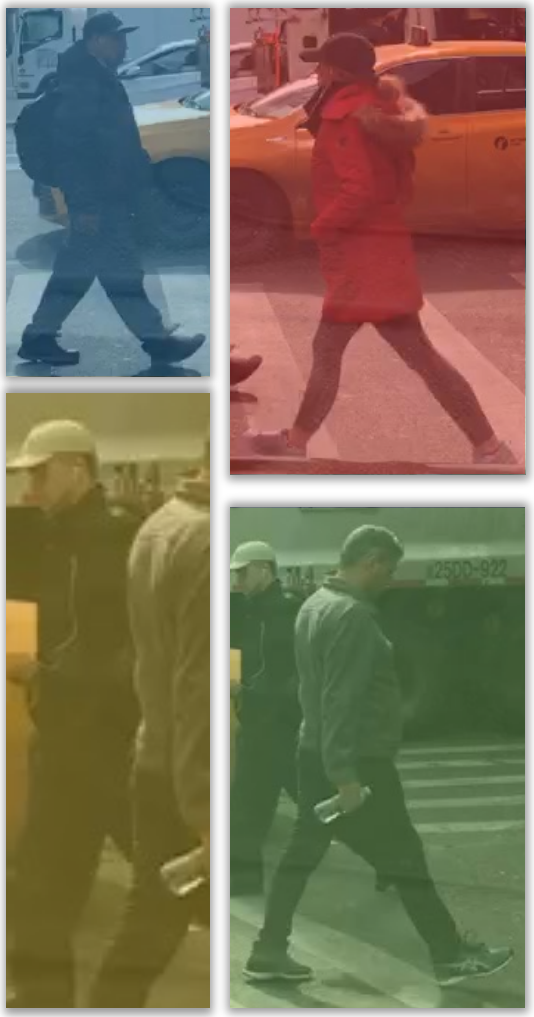
Let's play a game



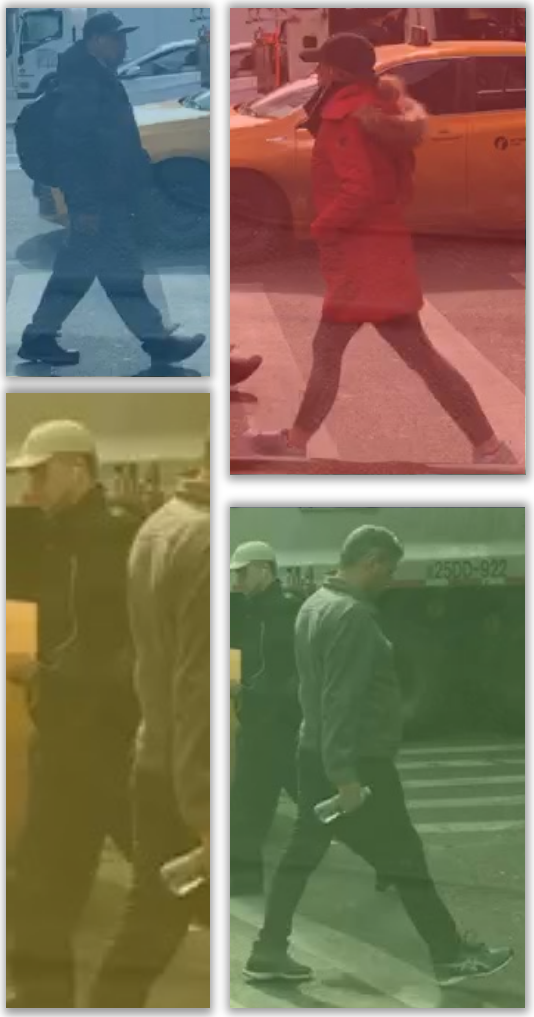


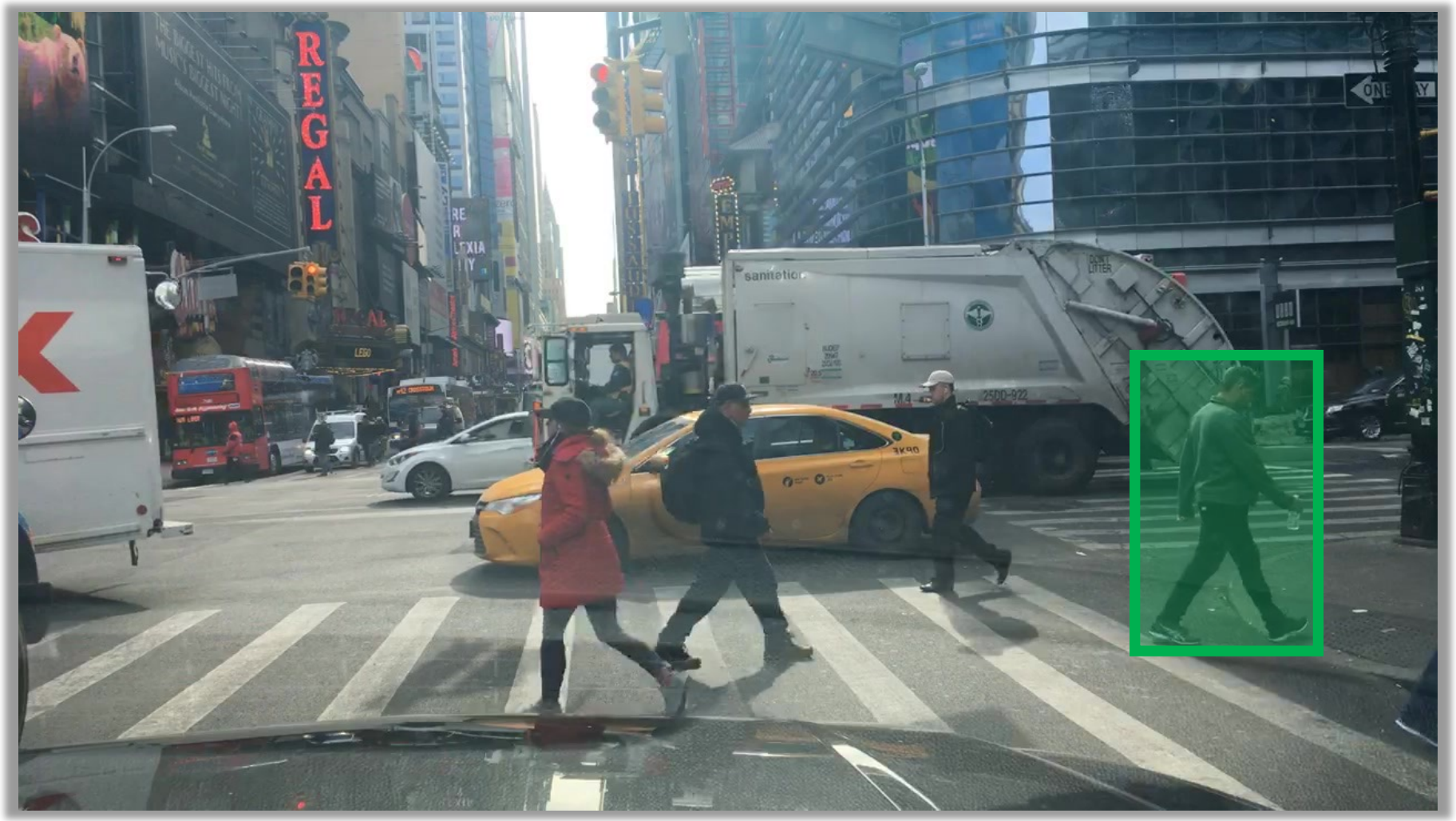
ETH zürich VIS Visual Intelligence and Systems

Let's play a game



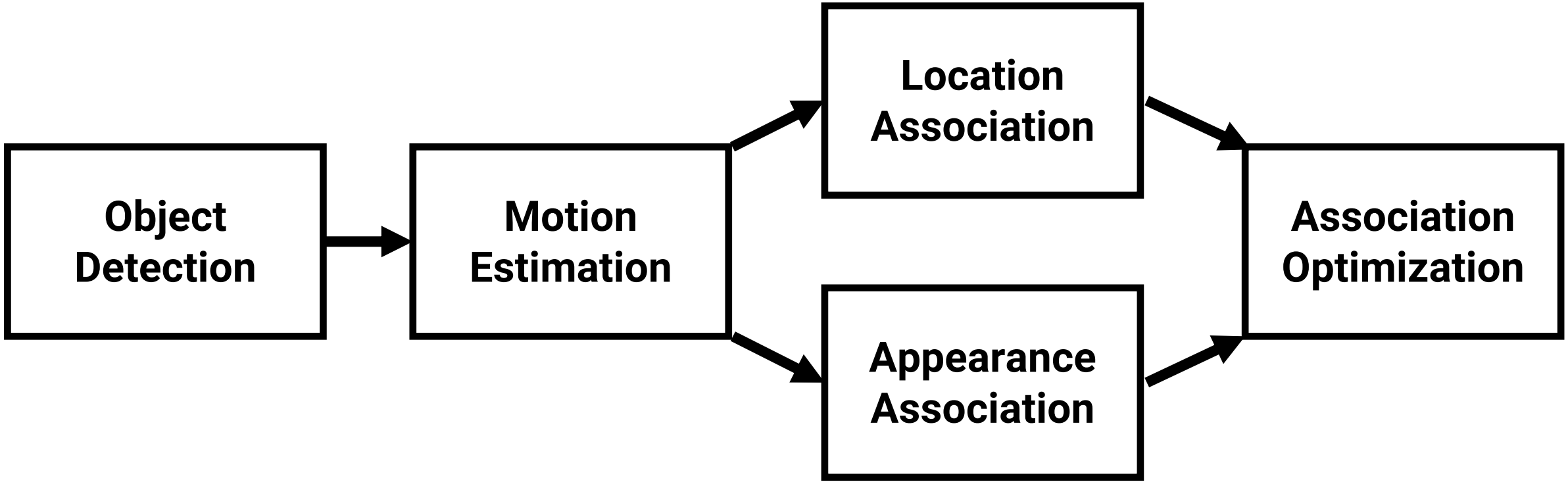
Let's play a game





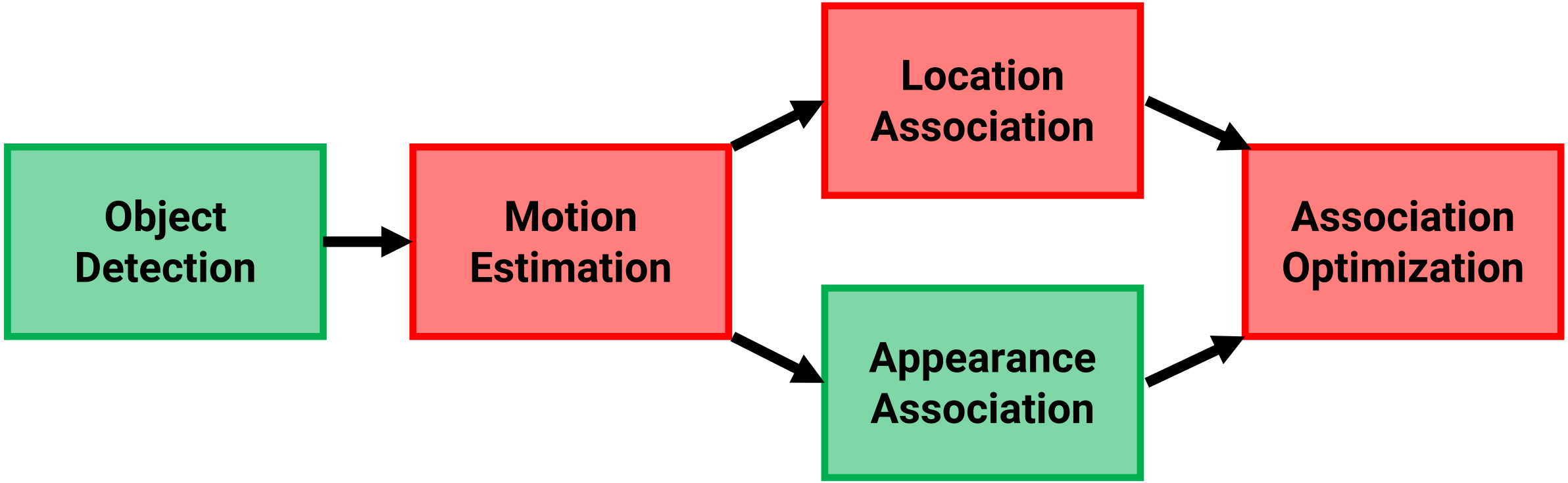
Game Conclusion

- Did we do motion estimation?
- Did we do association optimization?



Game Conclusion

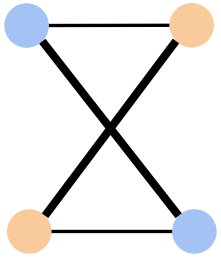
- Did we do motion estimation?
- Did we do association optimization?



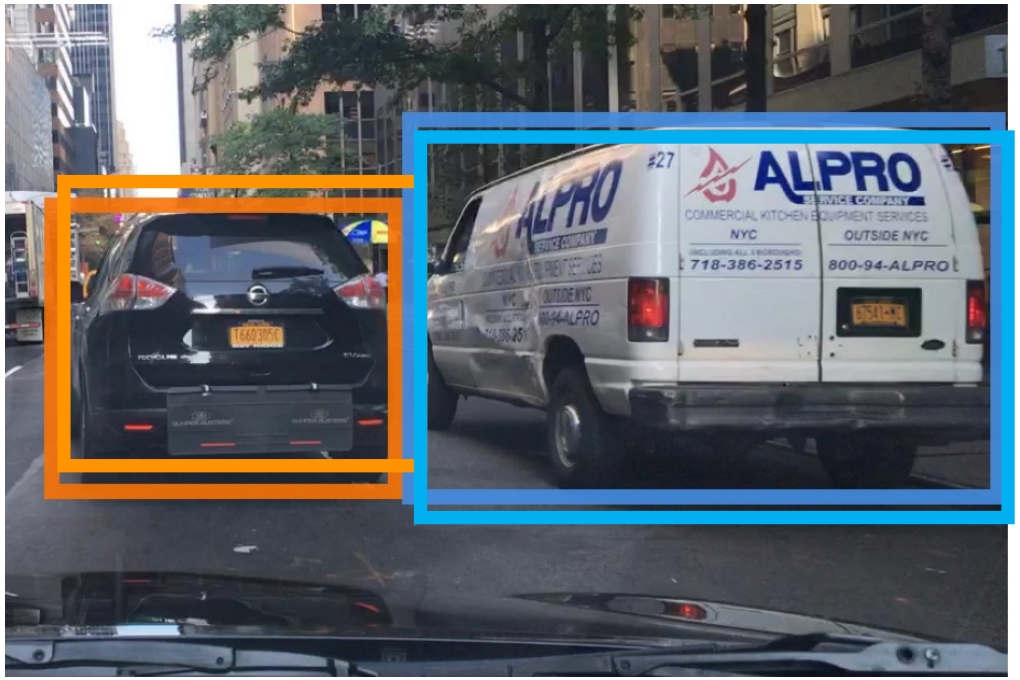
Real Question

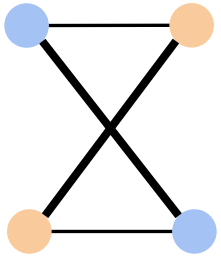
Why doesn’t appearance provide enough information in current models?

Existing Similarity Learning



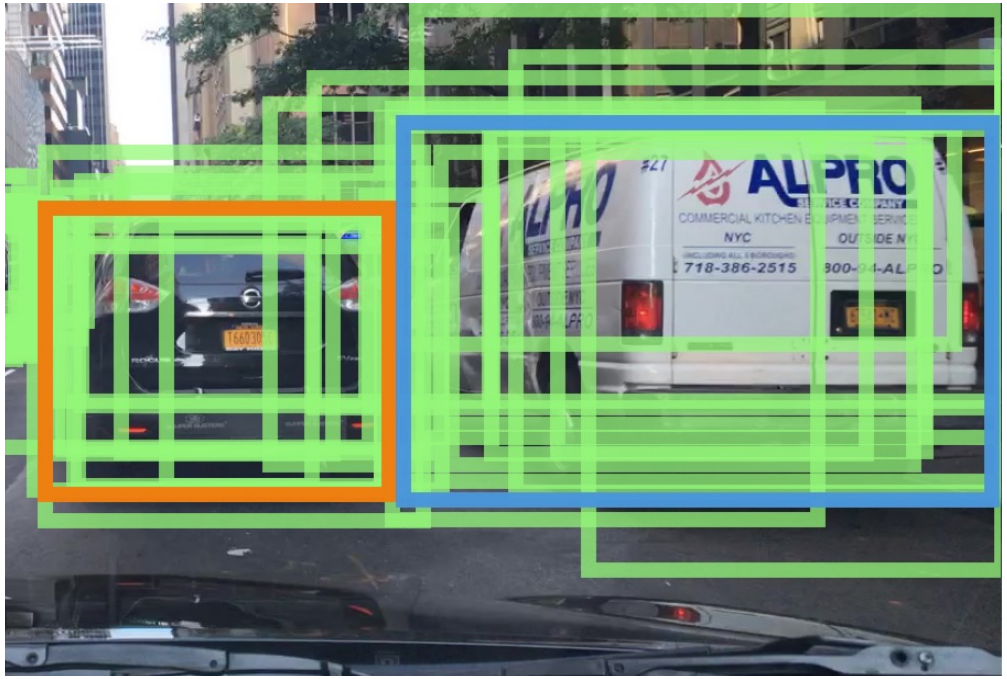
What's Missing

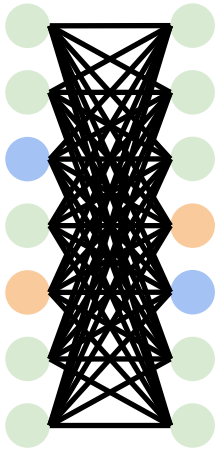


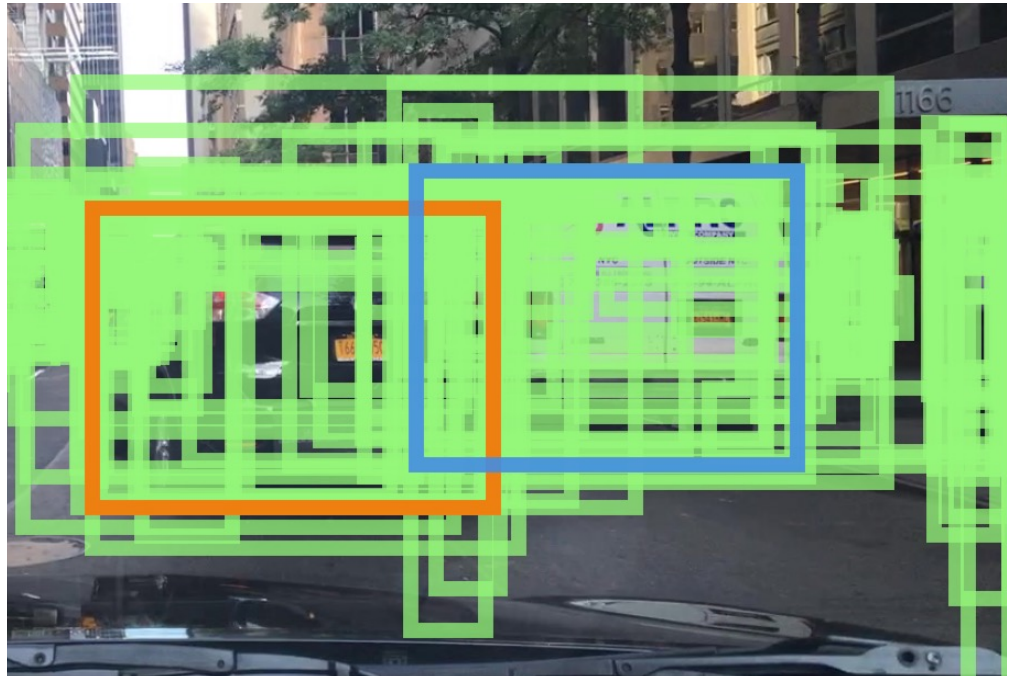


- Similar bounding boxes
- Misleading regions in the background

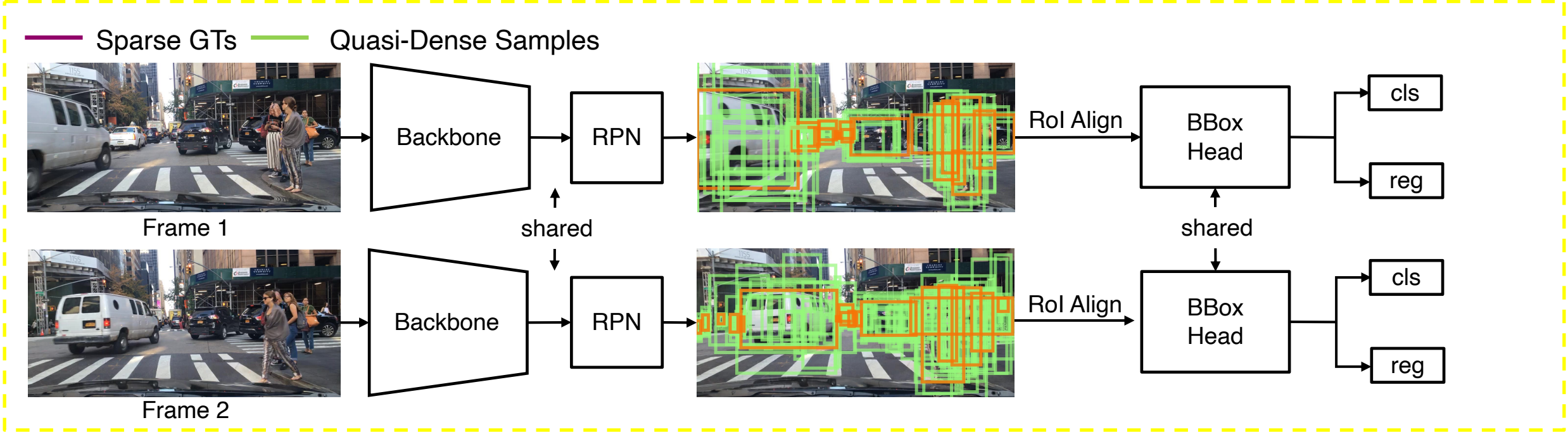
Quasi-Dense Matching





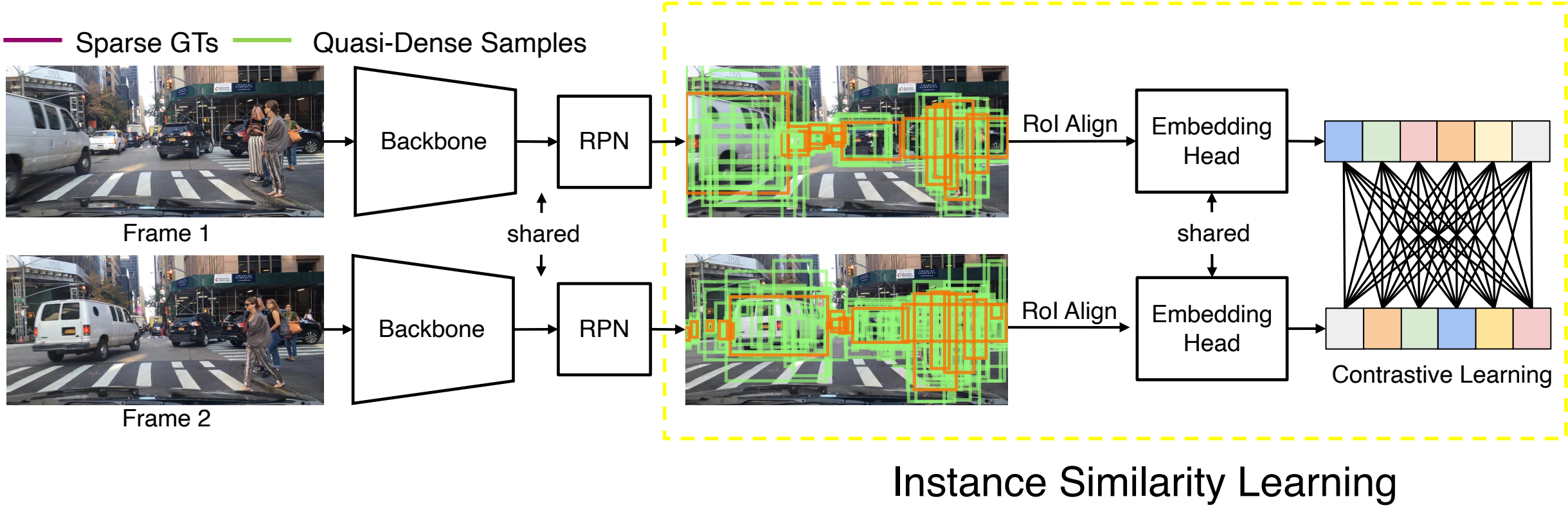


Training Pipeline

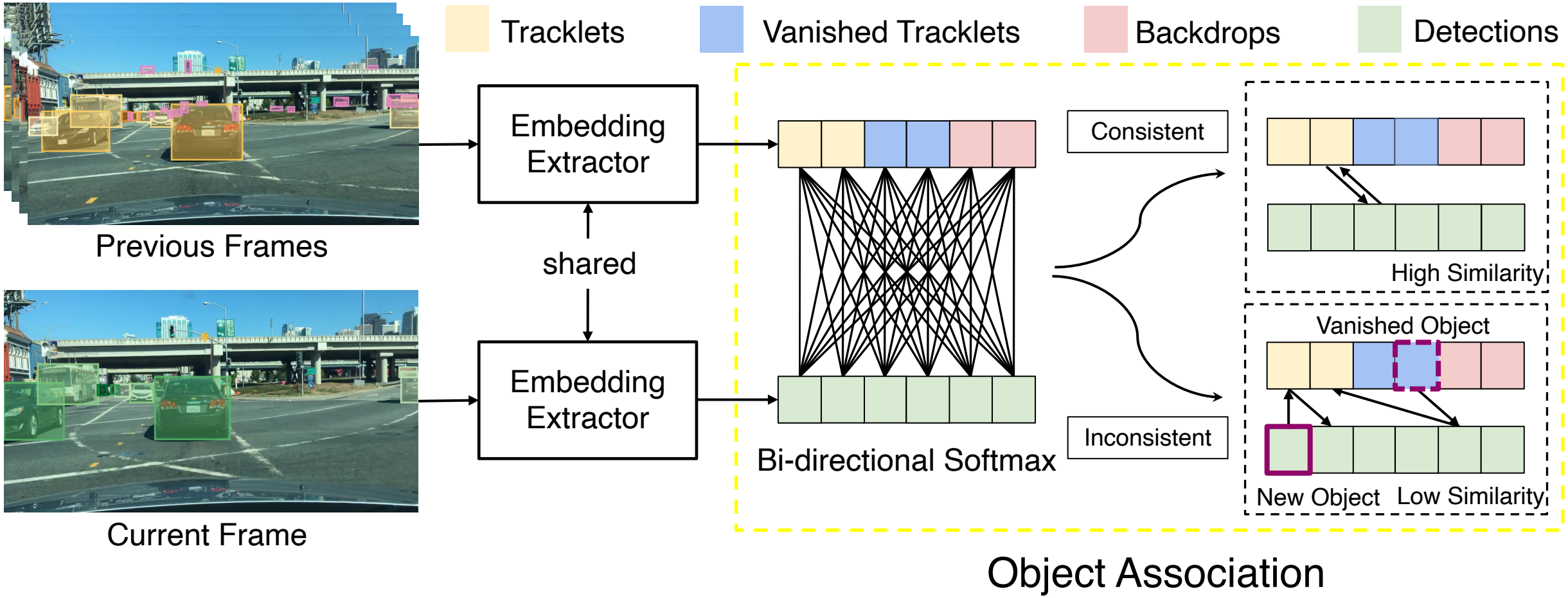


Object Detection

Training Pipeline

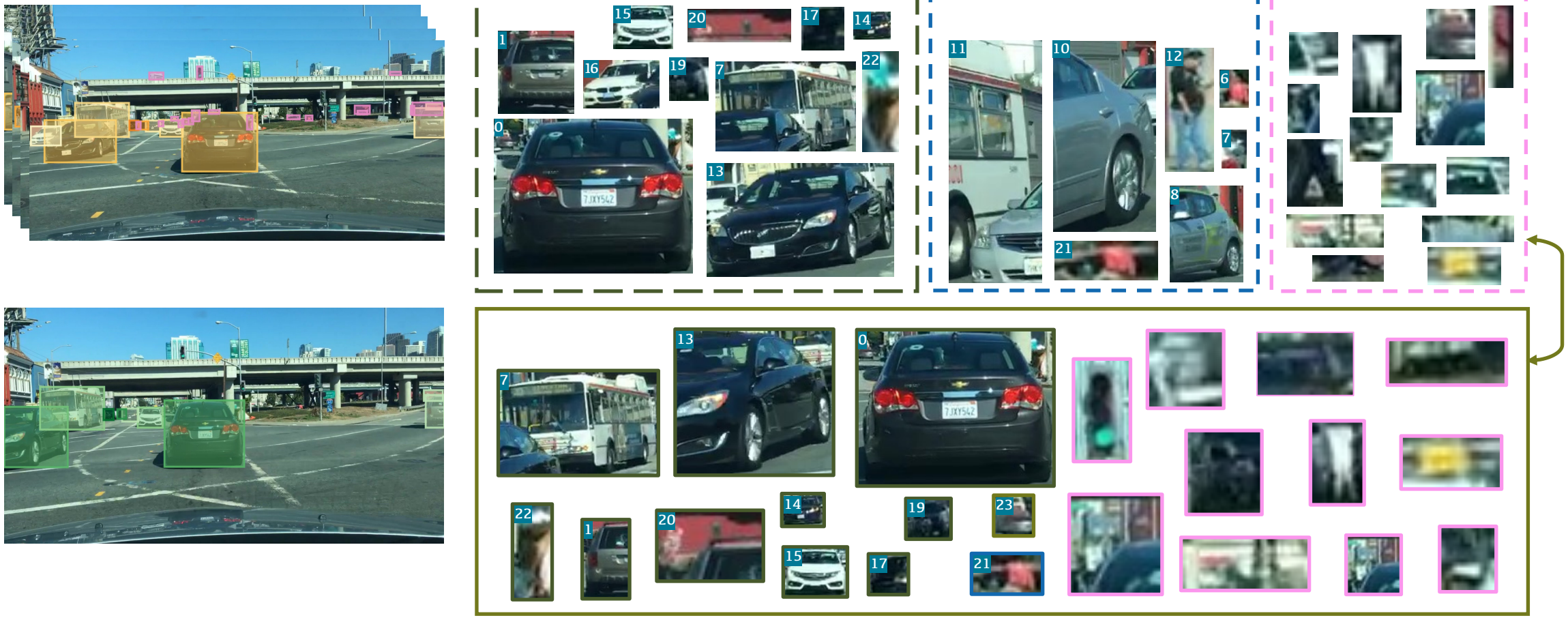


Testing Pipeline

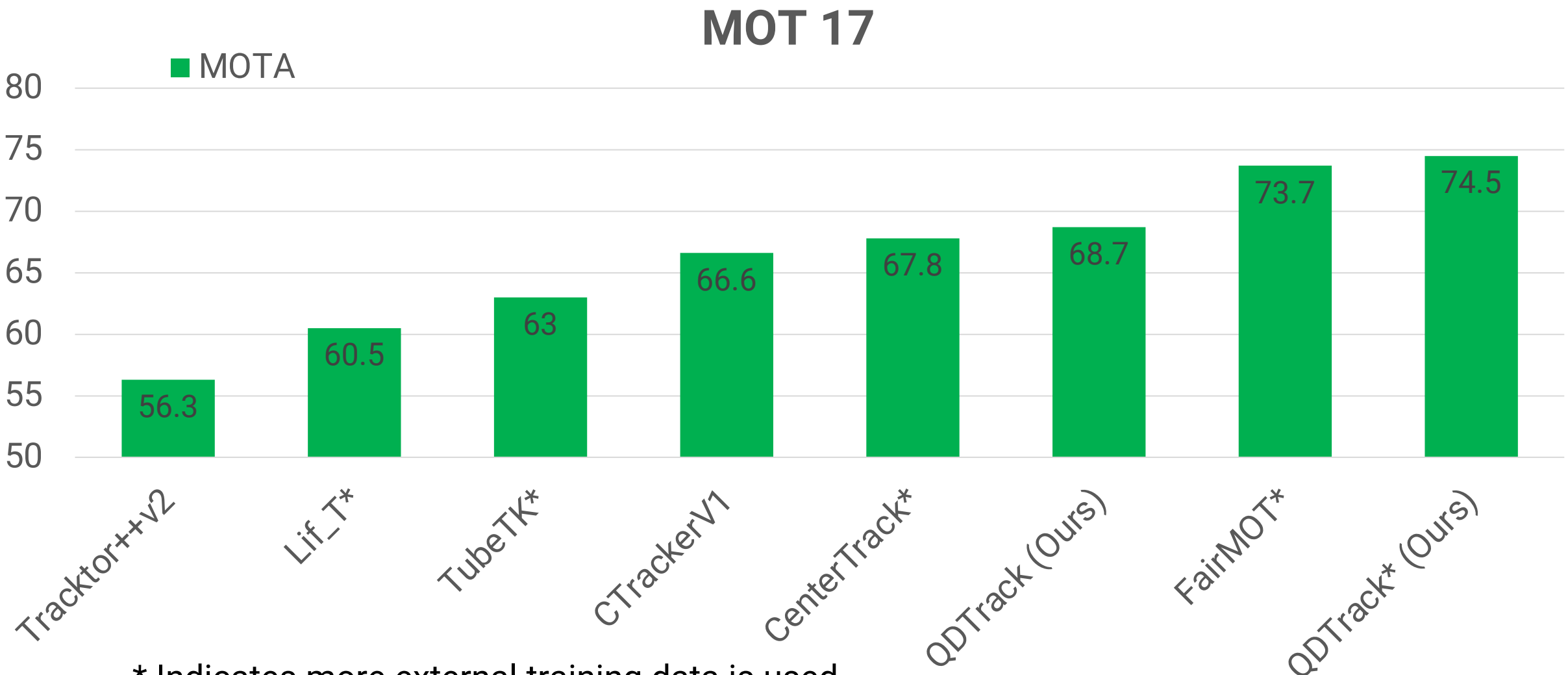


ETH zürich VIS Visual Intelligence and Systems

Object Association

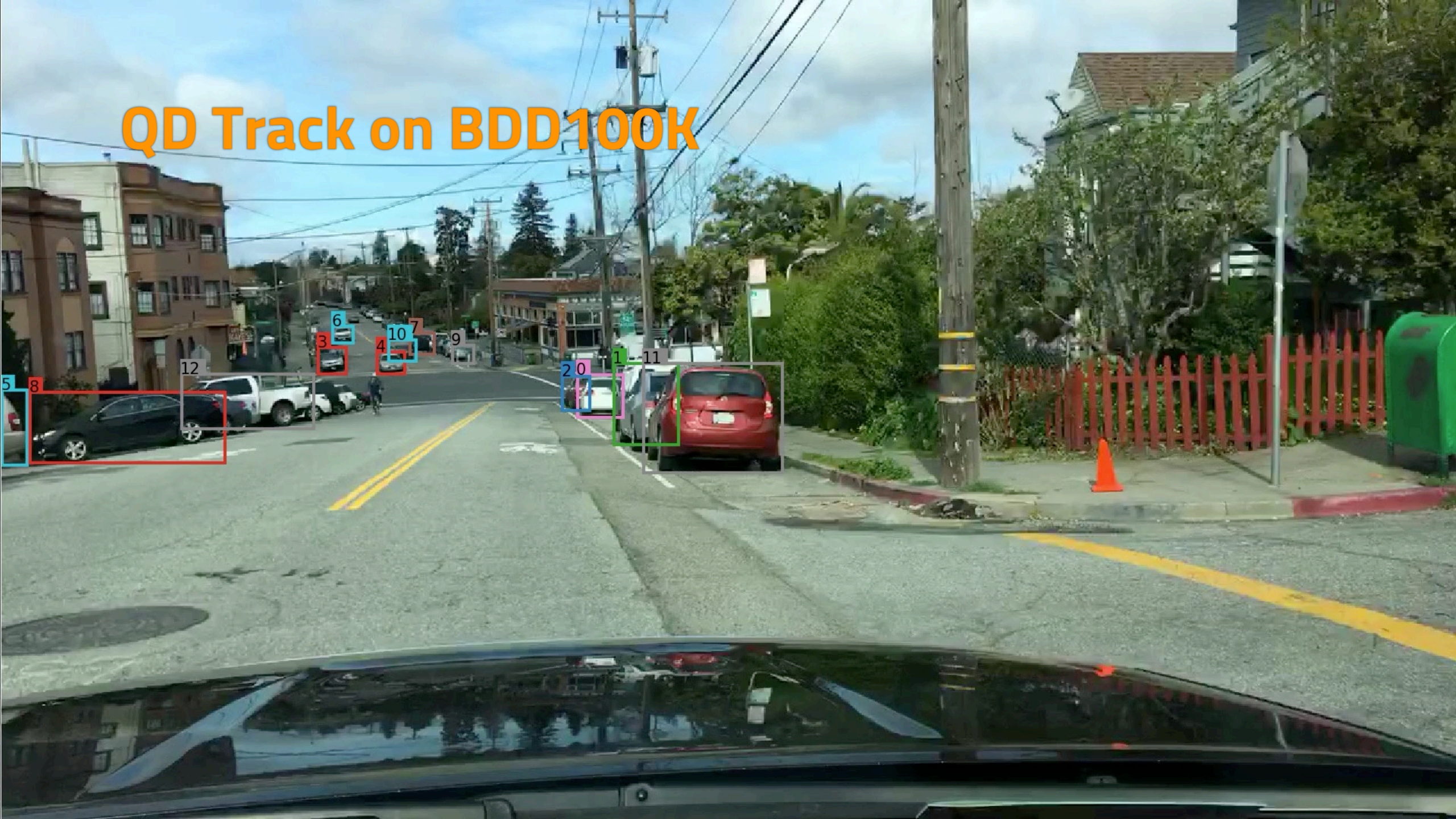


Evaluation on MOT17 Challenge



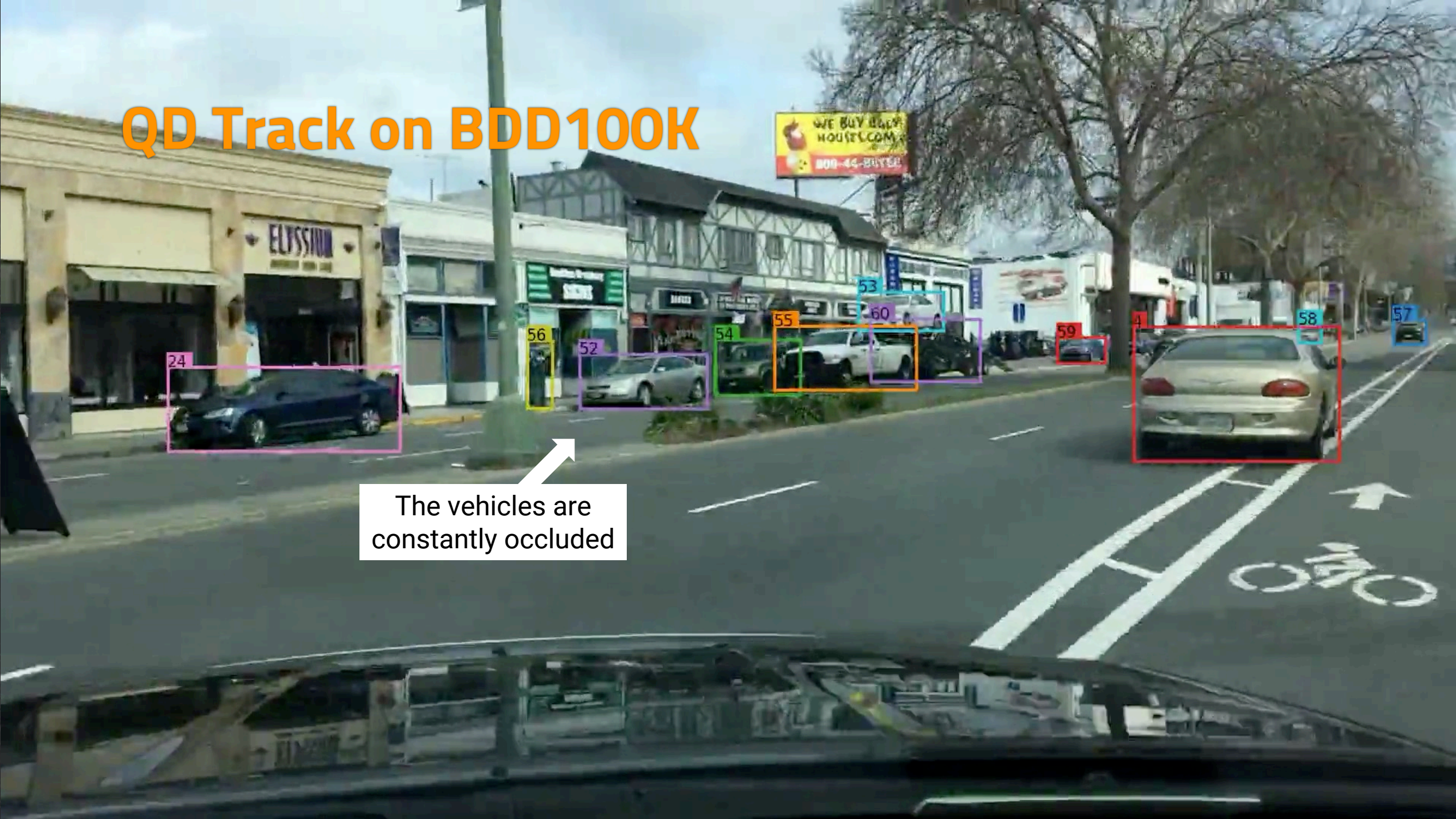
* Indicates more external training data is used

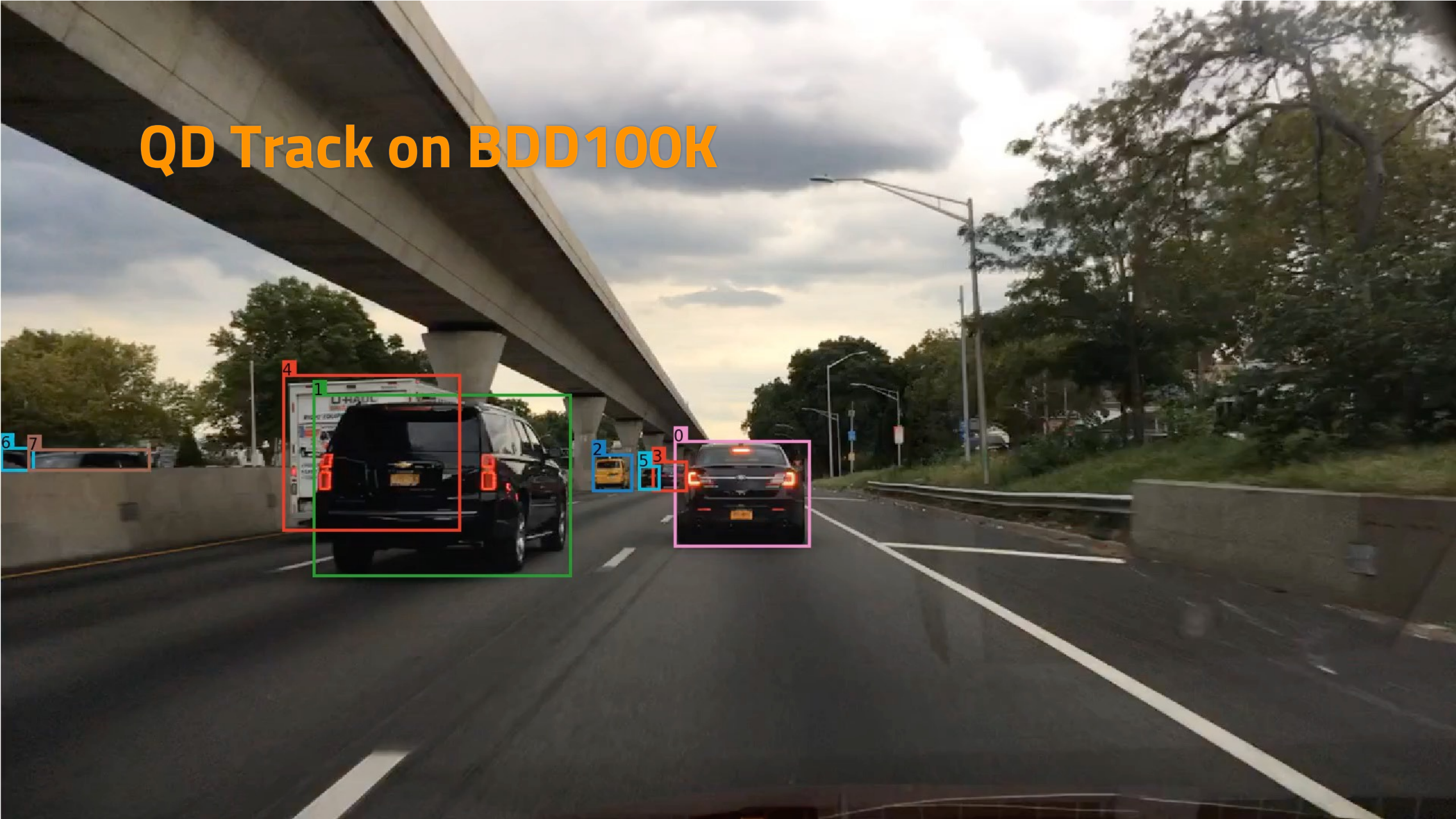

pedestrian1.00
pedestrian0.99

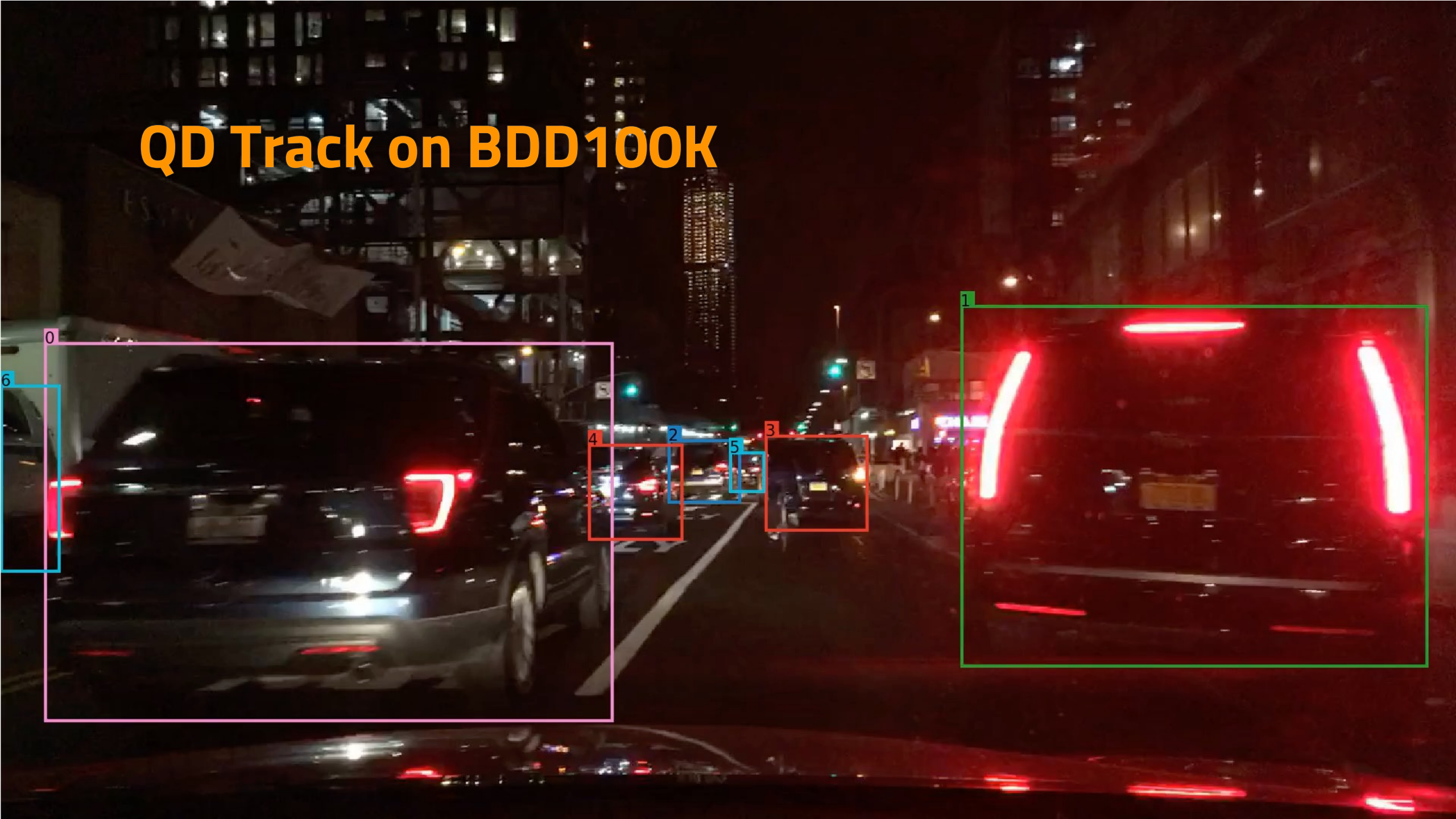

QD Track on BDD100K
5
8
12
6
3
10
7
4
9
2
0
1
11


QD Track on BDD100K
This person is about to disappear
39
41
42
6
10
7
40
9
4
37
31
18
2
28
24
32
25
17

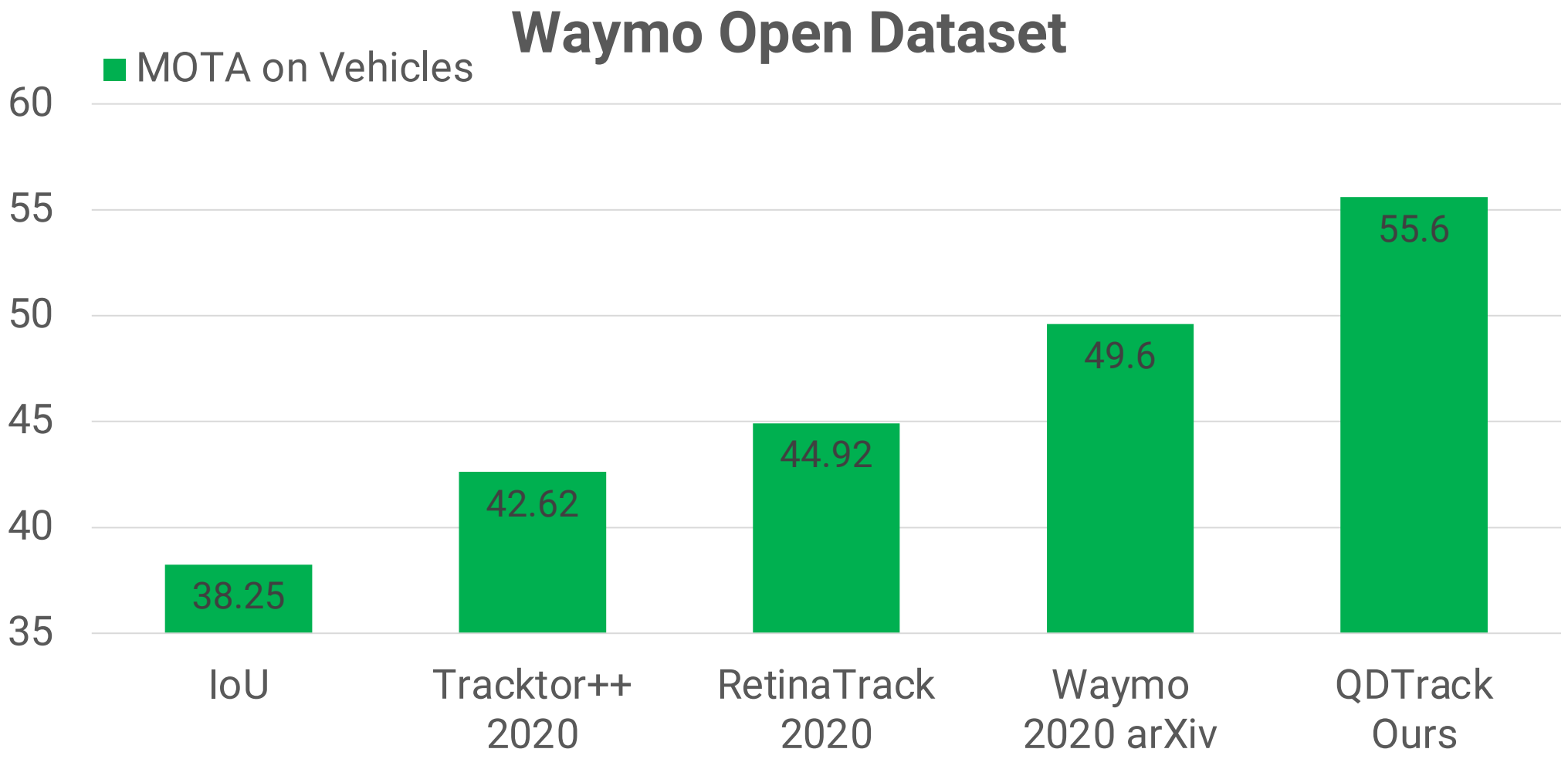

QD Track on BDD100K
This person is tracked


QD Track on BDD100K
24
56
52
54
55
53
60
59
58
57
The vehicles are constantly occluded


QD Track on BDD100K
4
1
6
7
2
5
3
0

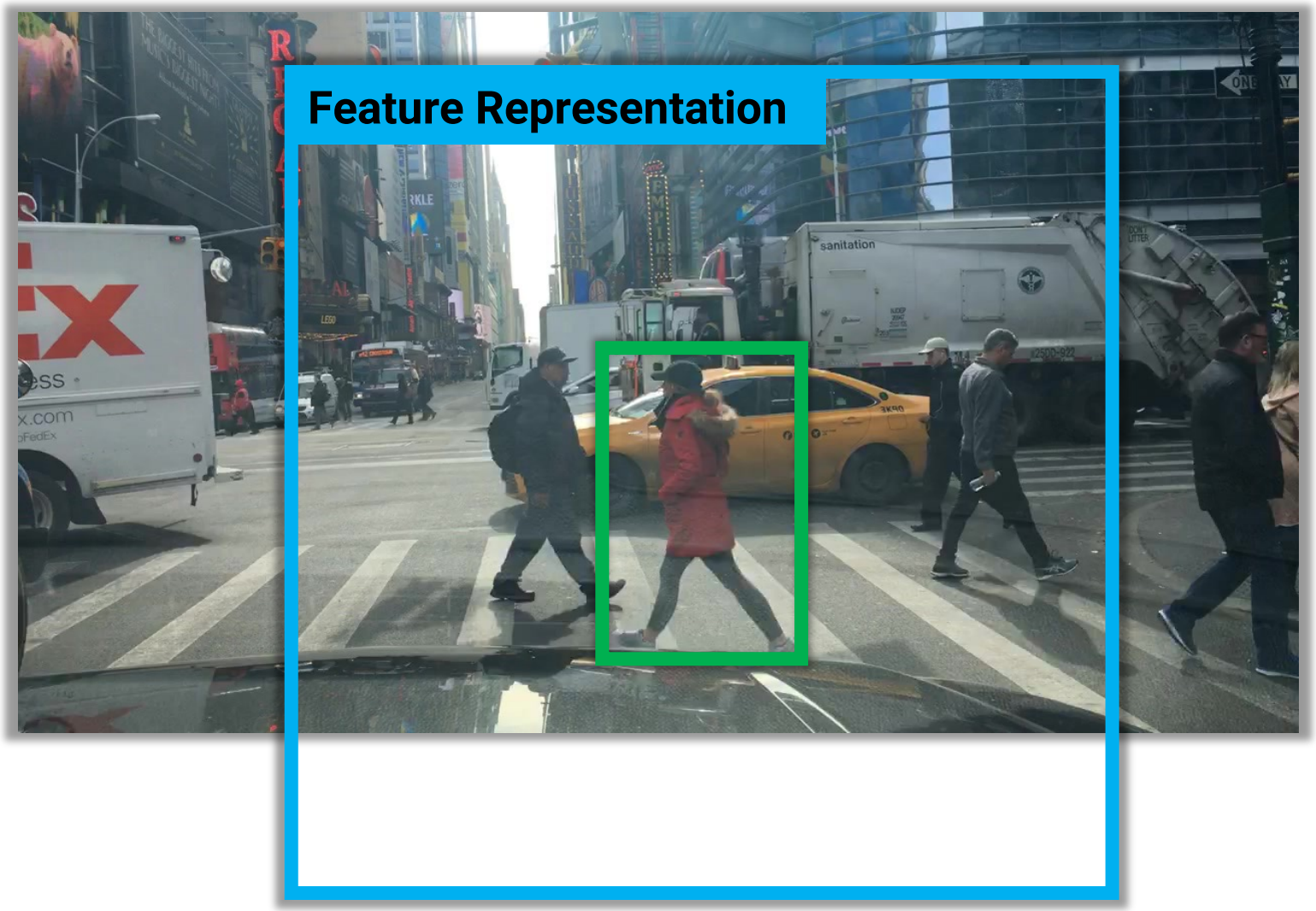

QD Track on BDD100K
0
1
2
3
4
5
6

Results on Waymo Open Dataset



ETH zürich VIS Visual Intelligence and Systems

Is appearance similarity really enough?

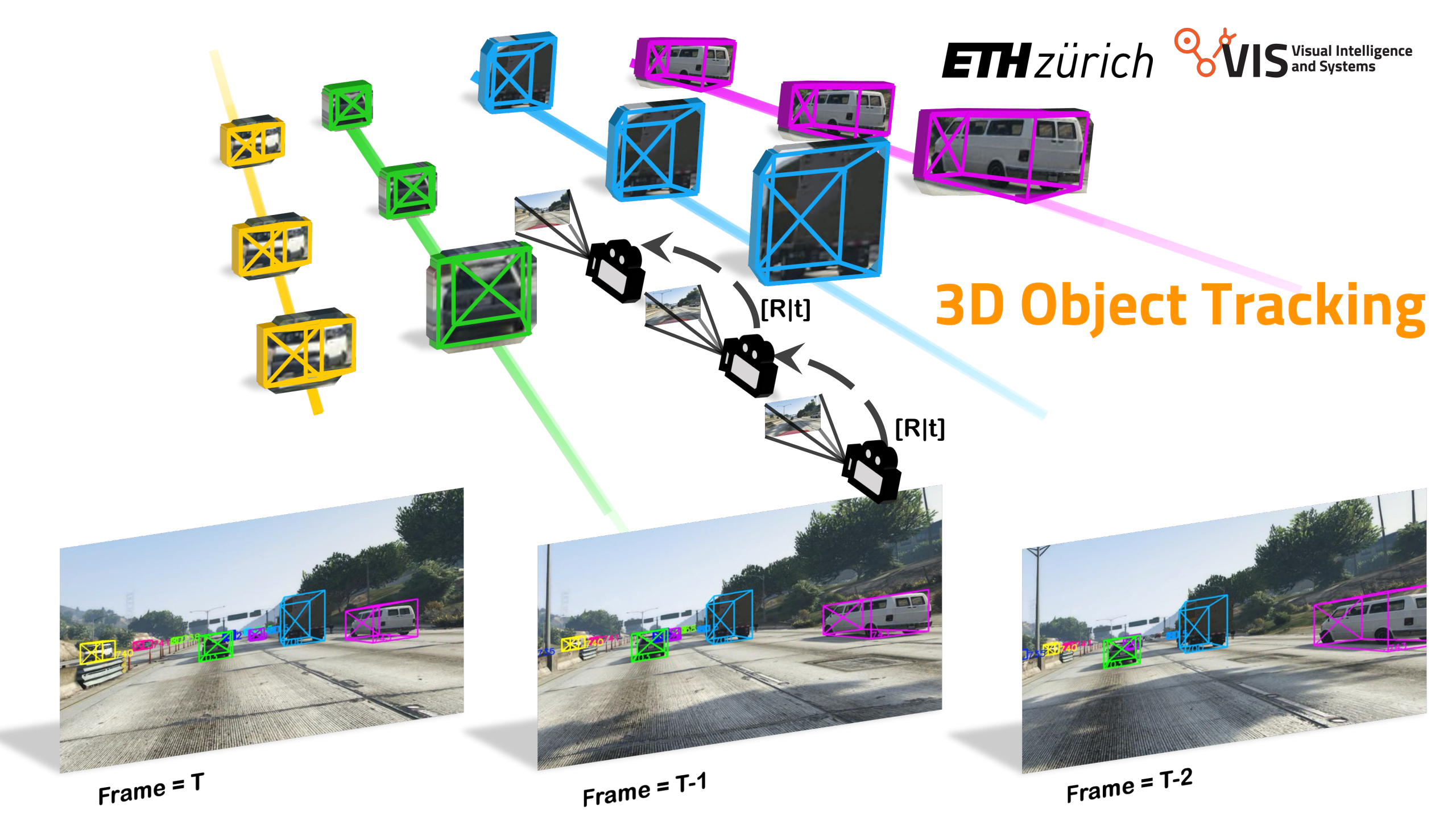


Tracking

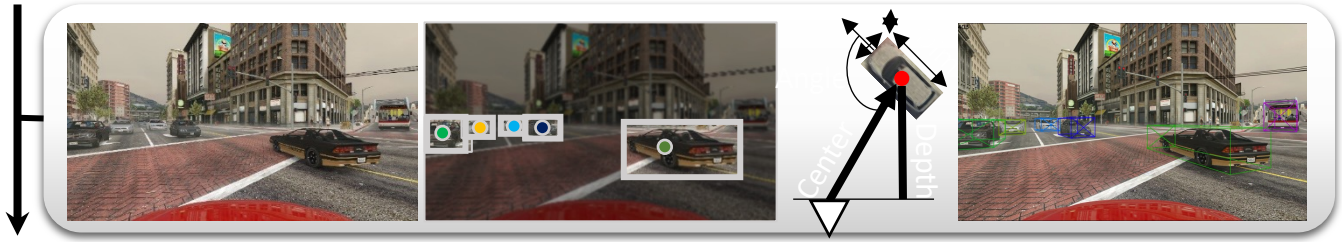
- Track a point
- Track a bigger box
- Track by detection
- Online learning
- Motion
- Multiple object tracking
- **3D object tracking**

ETH zürich VIS Visual Intelligence and Systems

3D Object Tracking

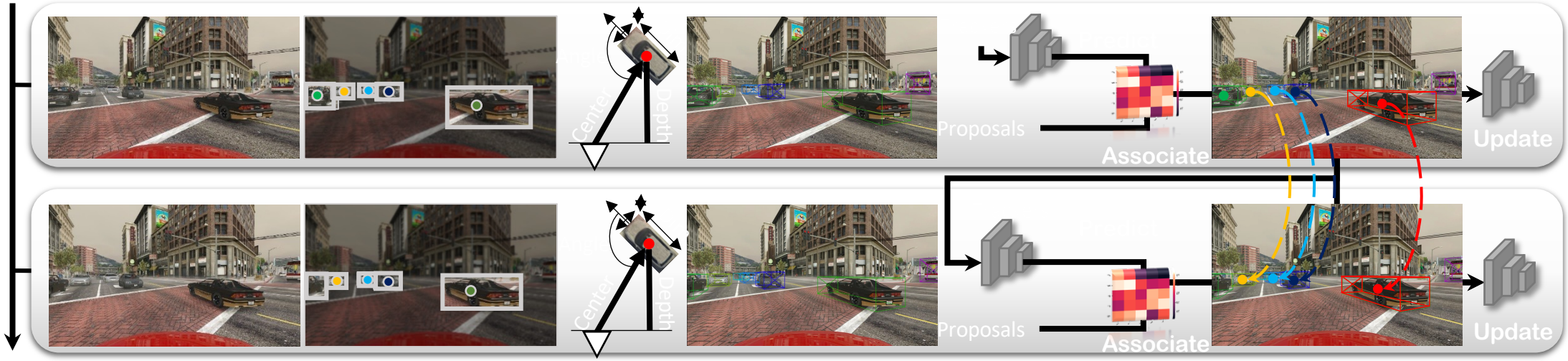


3D Tracking – Multi-frame 3D model



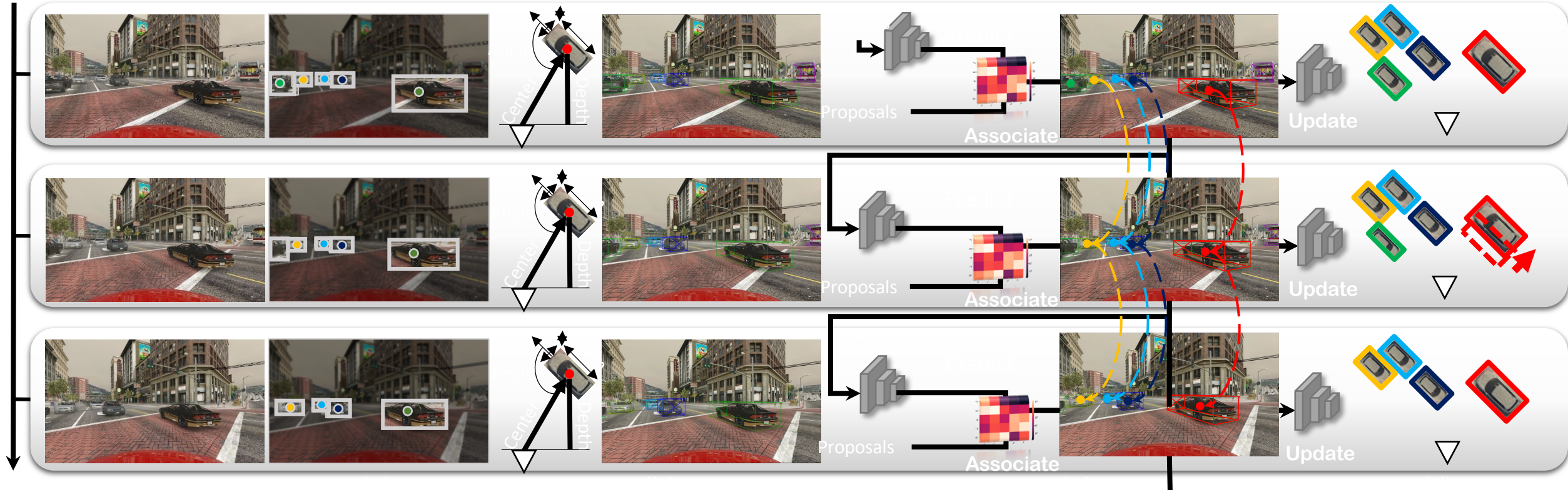
Joint Monocular 3D Vehicle Detection and Tracking, ICCV 2019

3D Tracking – Multi-frame 3D model



Joint Monocular 3D Vehicle Detection and Tracking, ICCV 2019

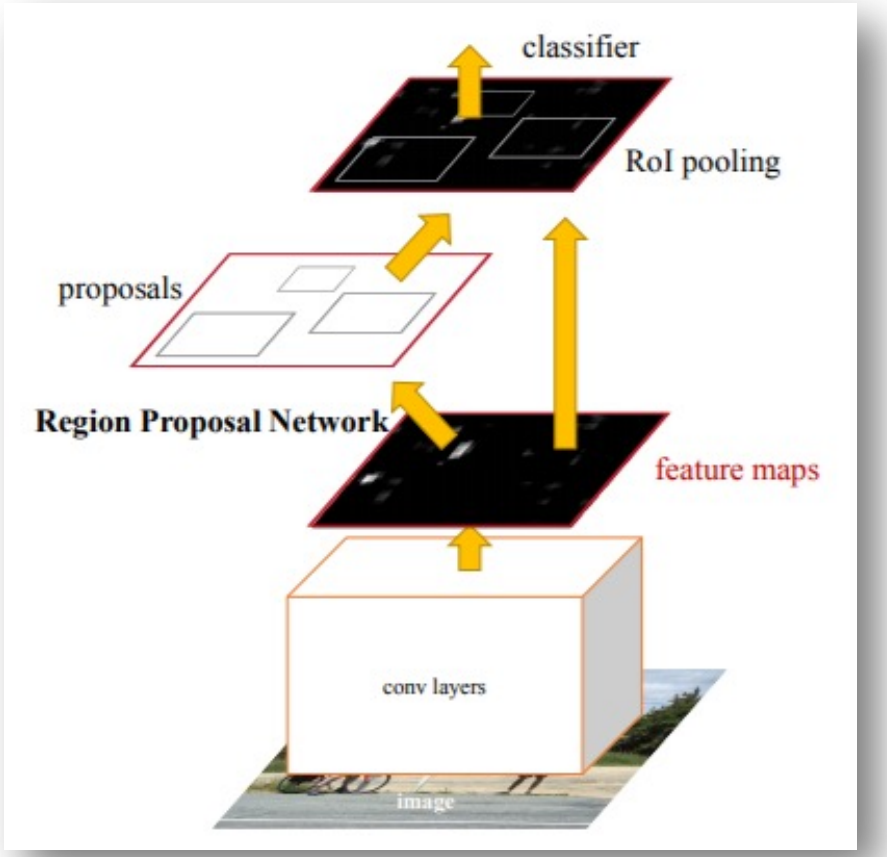
3D Tracking – Multi-frame 3D model



1. Object Detection
2. Object Dist., Orientation, Size
3. Association
4. Motion prediction

Joint Monocular 3D Vehicle Detection and Tracking, ICCV 2019

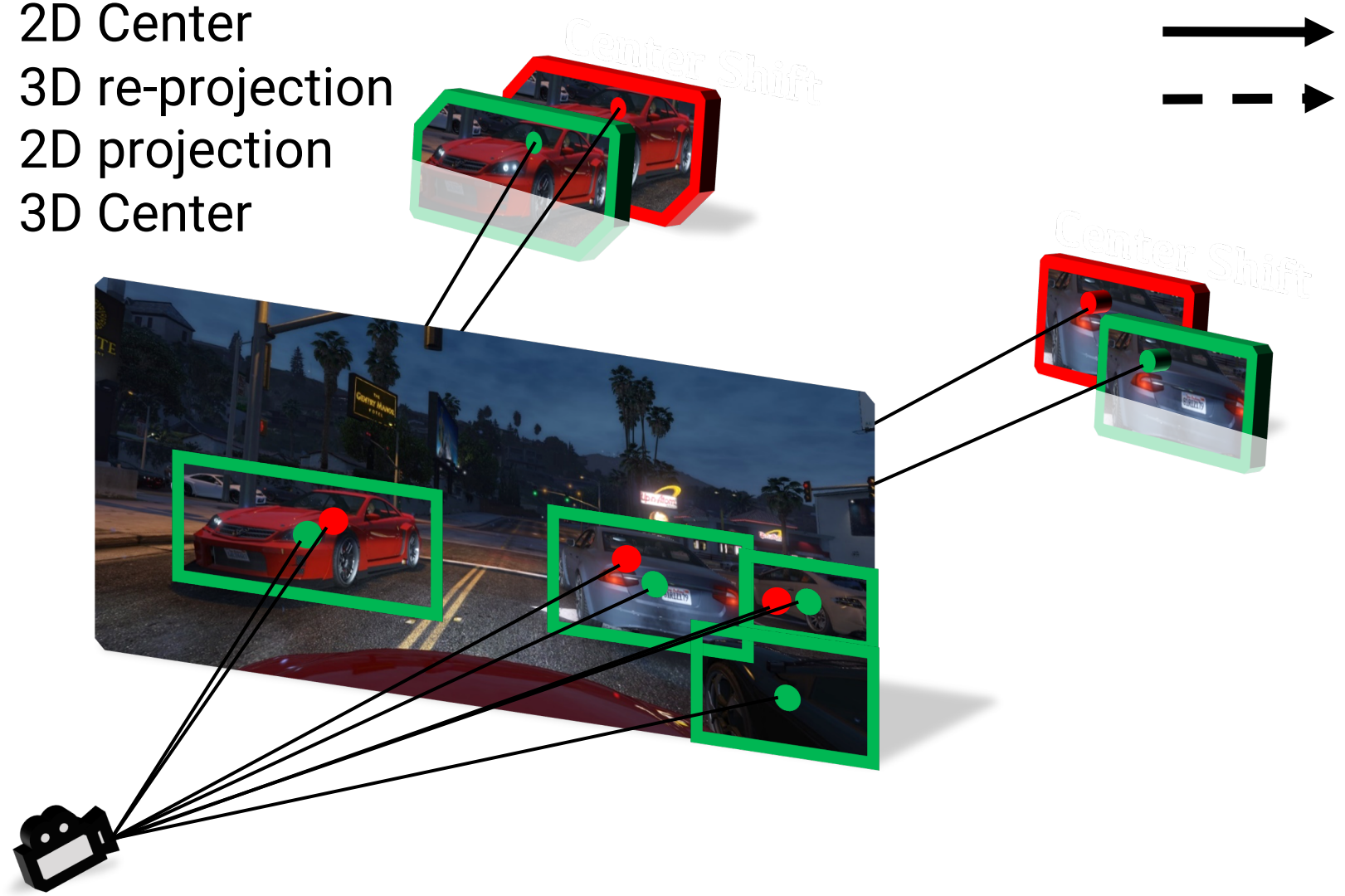
3D Tracking – 2D Detection & Localization



Using Faster RCNN to estimate 3D center re-projection with up to 300 proposals

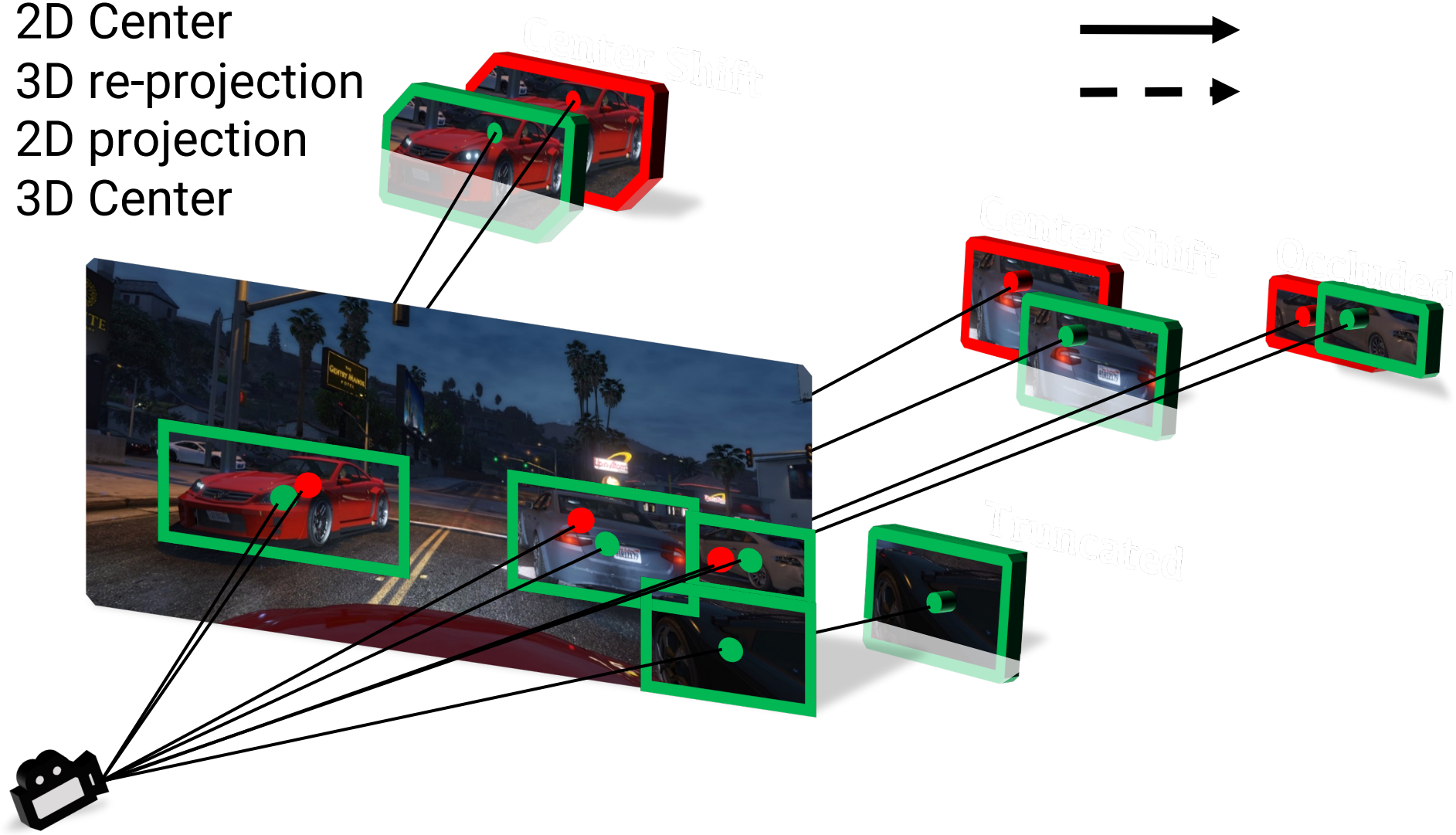
3D Tracking – 3D Estimation

- 2D Center
- 3D re-projection
- 2D projection
- 3D Center



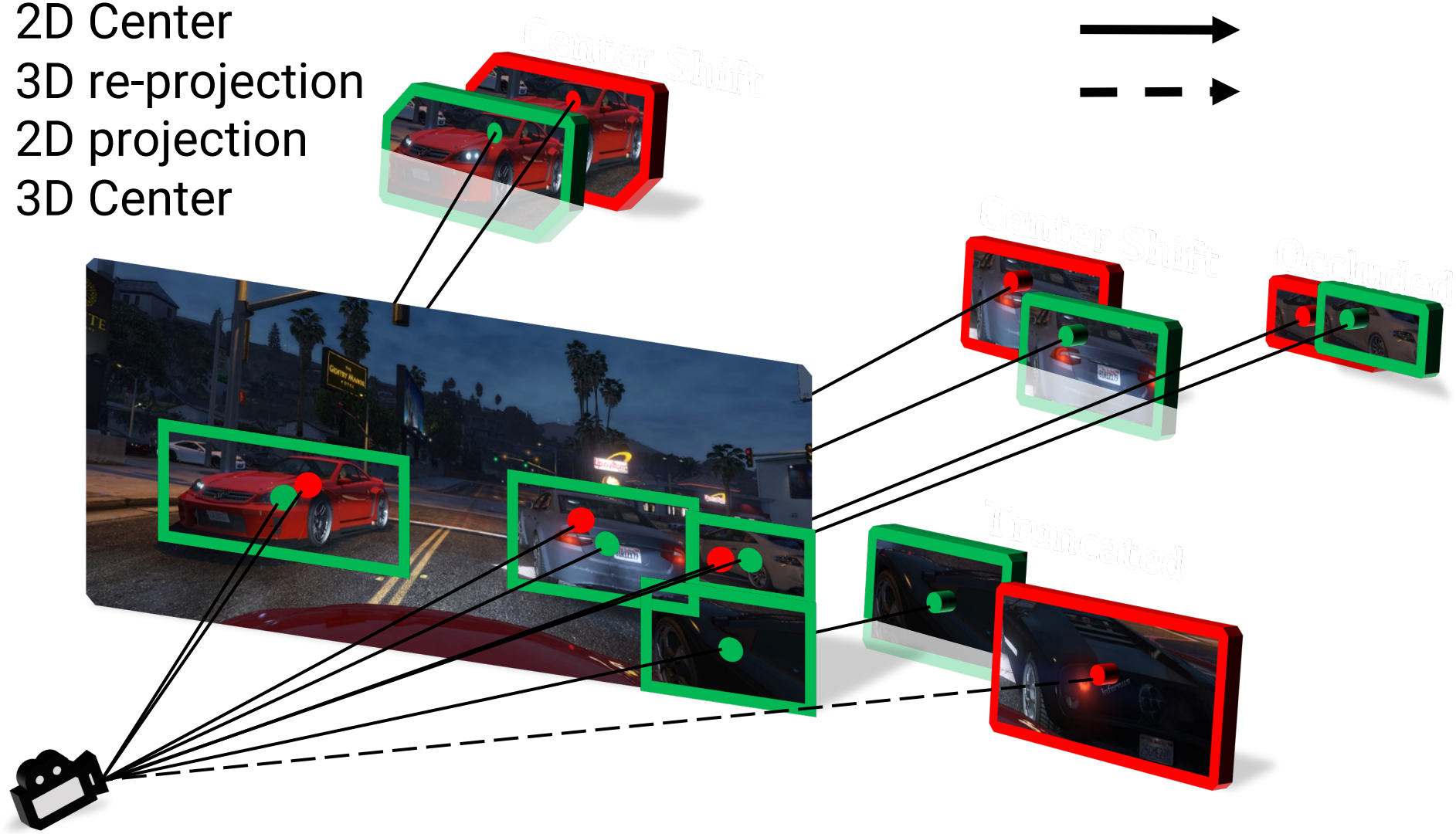
3D Tracking – 3D Estimation

- 2D Center
- 3D re-projection
- 2D projection
- 3D Center

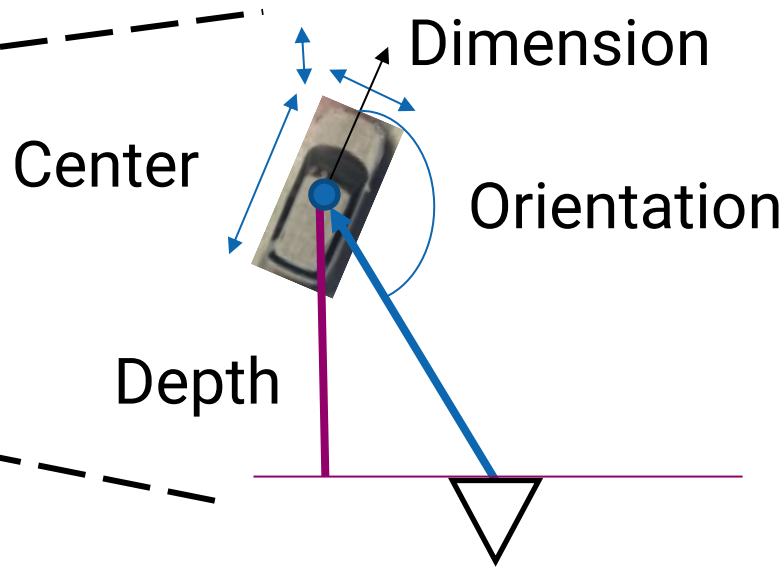


3D Tracking – 3D Estimation

- 2D Center
- 3D re-projection
- 2D projection
- 3D Center



3D Tracking – 3D Estimation



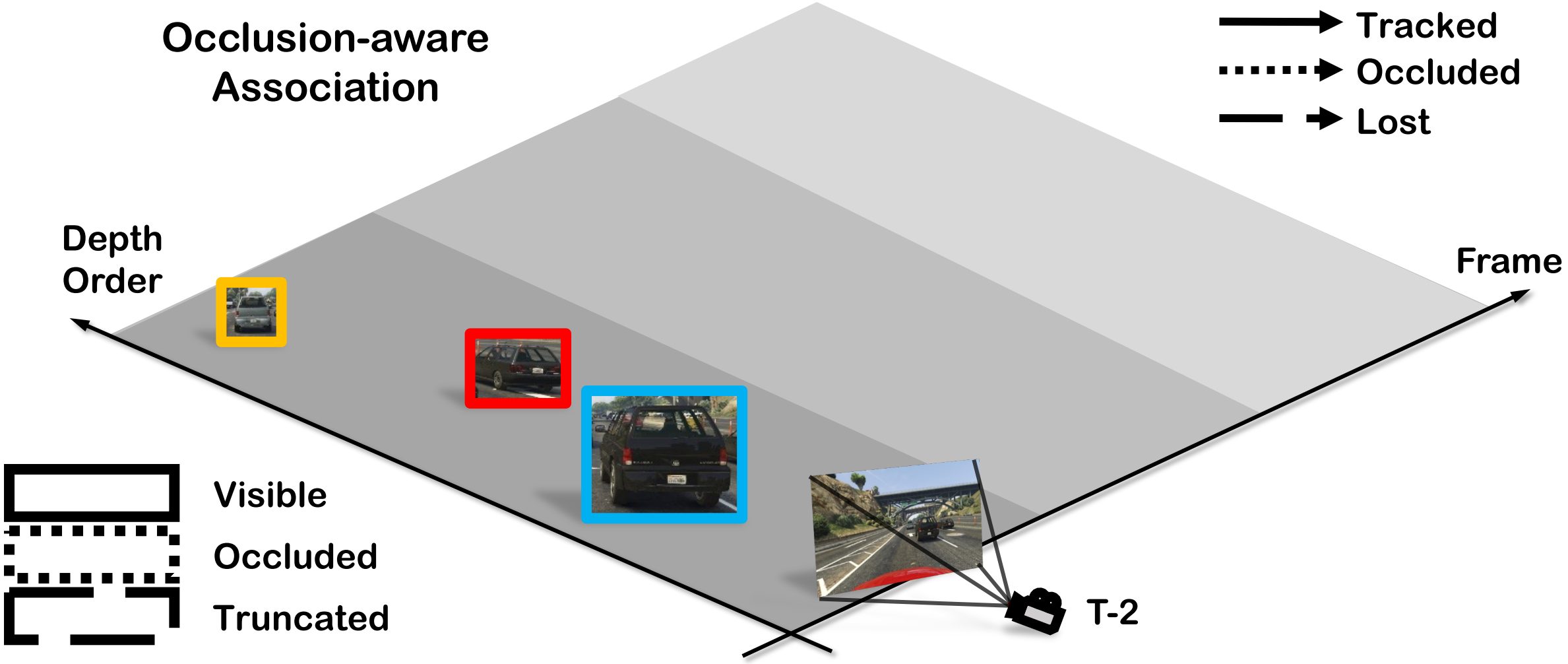
Length, Width, Height

Two Bins, each have a probability pair and an angle pair

1 / Depth (Disparity)

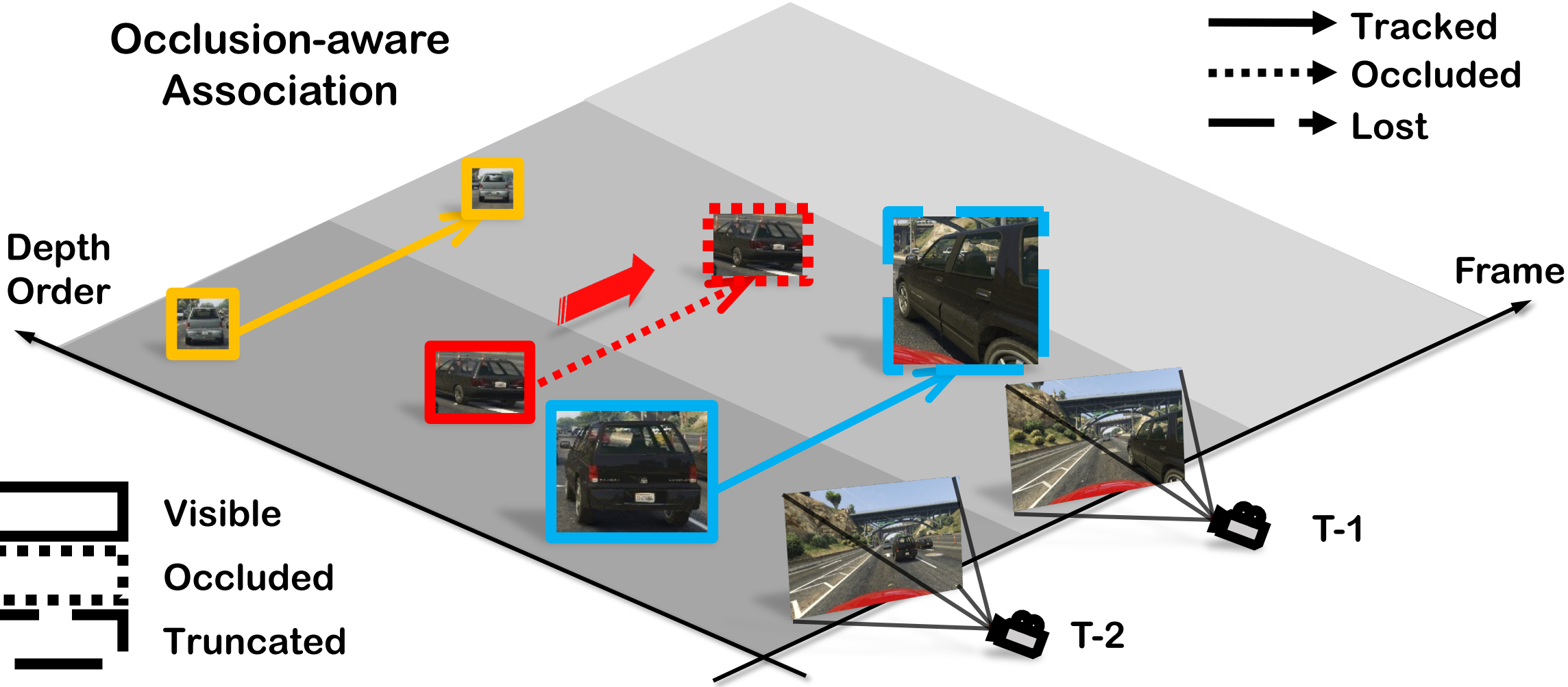
Using DLA-up 34 as base to compute convolutional representation
Pyramid shape convolution module as the multi-head layer

3D Tracking – Occlusion-Aware Association



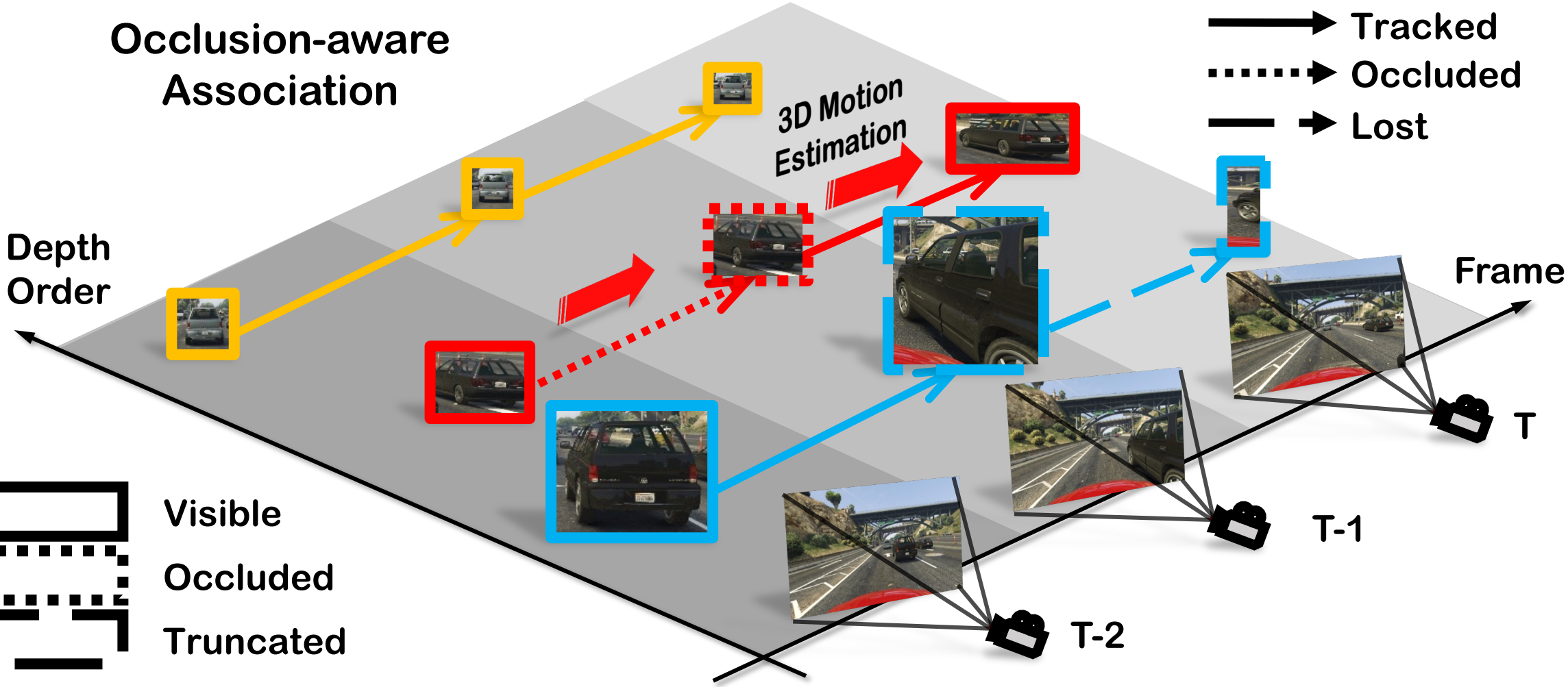
ETH zürich VIS Visual Intelligence and Systems

3D Tracking – Occlusion-Aware Association

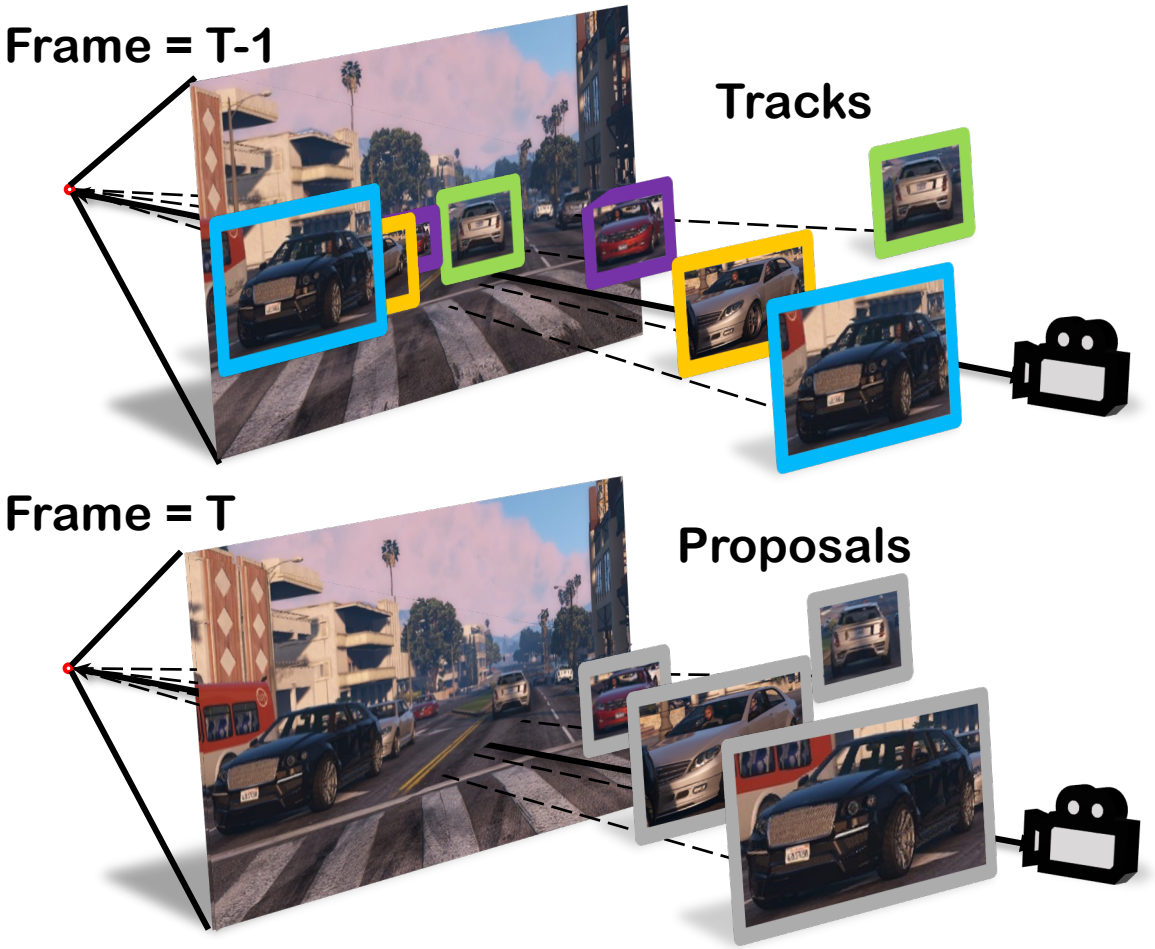


ETH zürich VIS Visual Intelligence and Systems

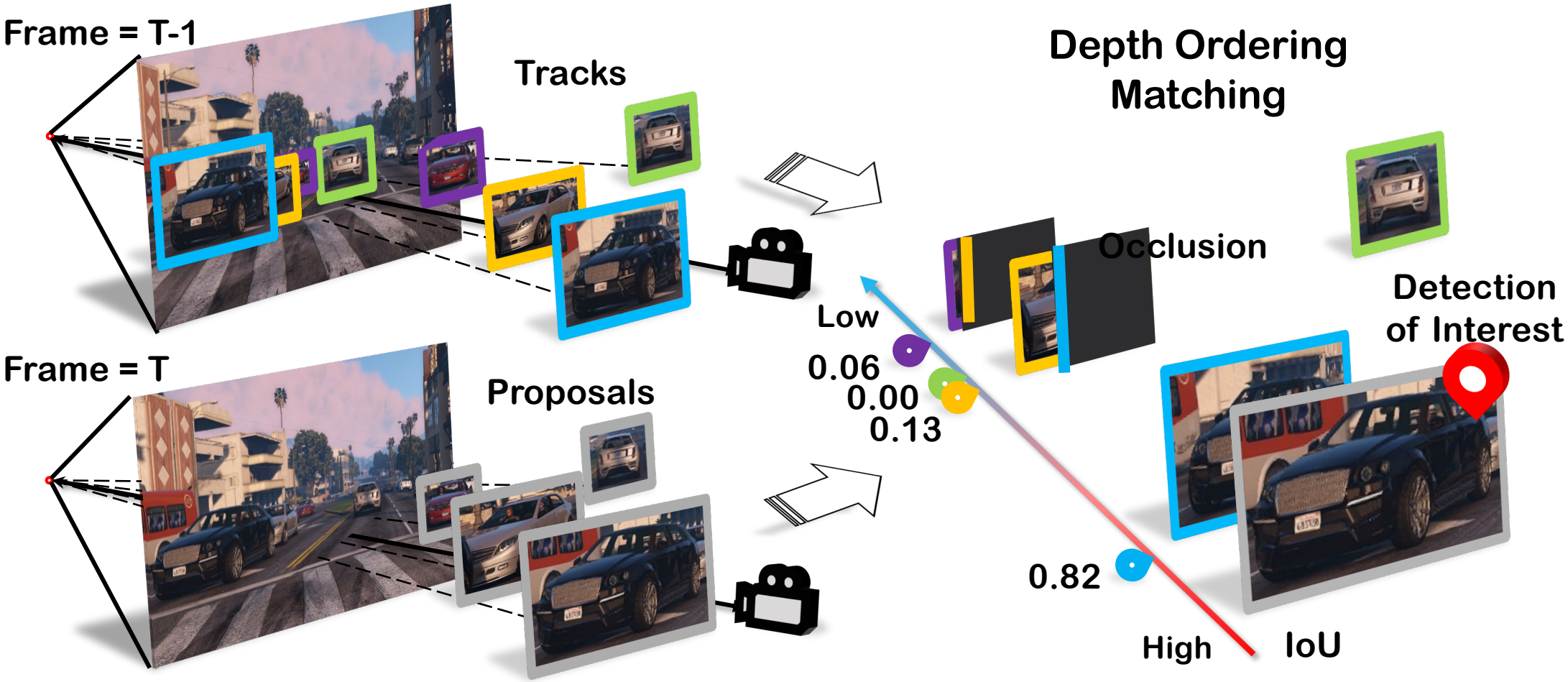
3D Tracking – Occlusion-Aware Association



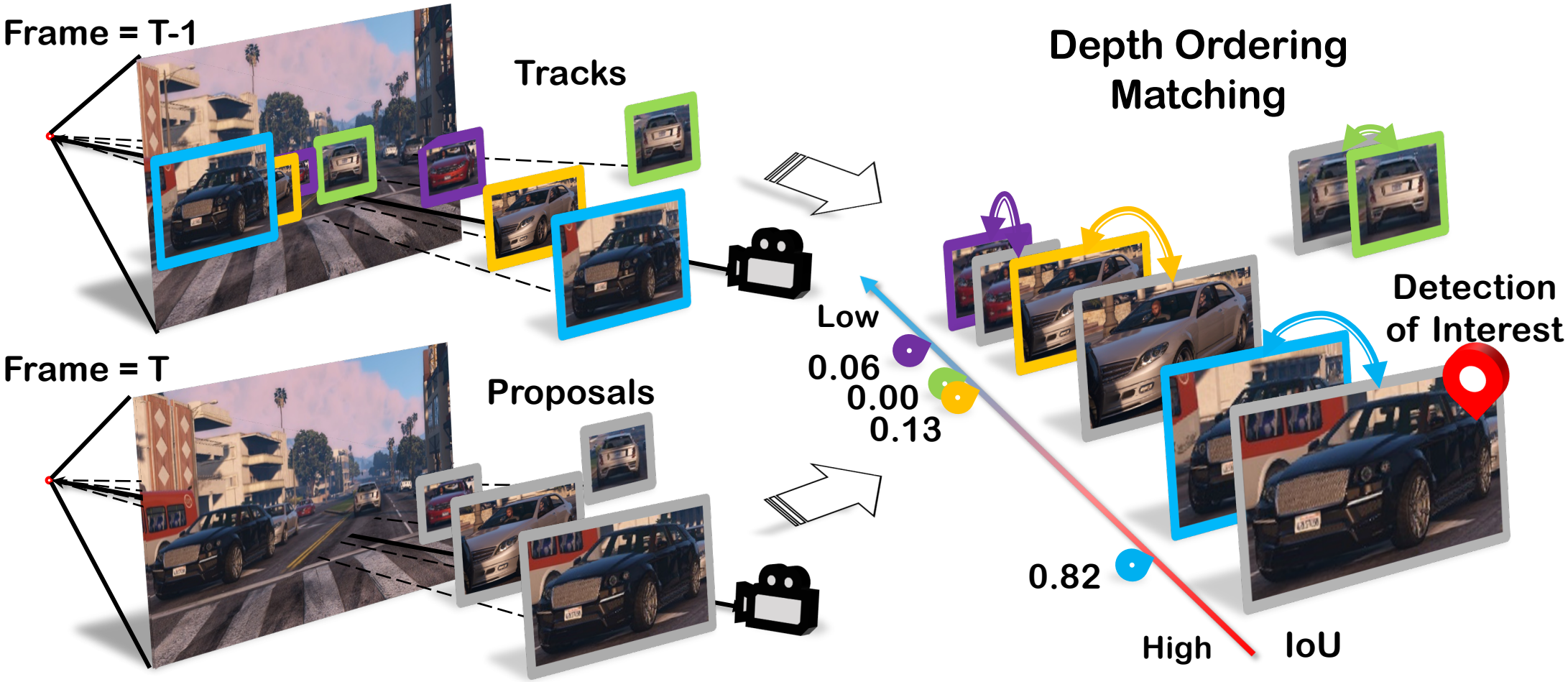
3D Tracking – Order-Aware Matching



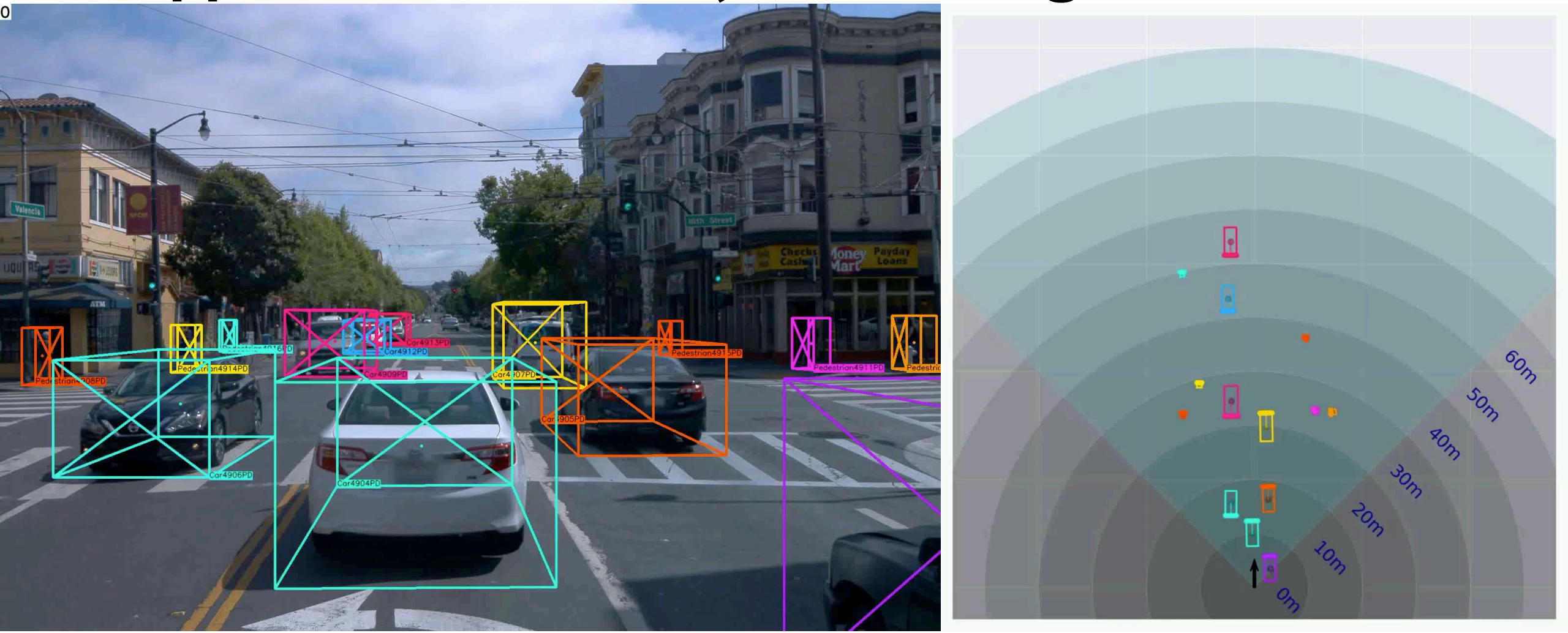
3D Tracking – Order-Aware Matching



3D Tracking – Order-Aware Matching



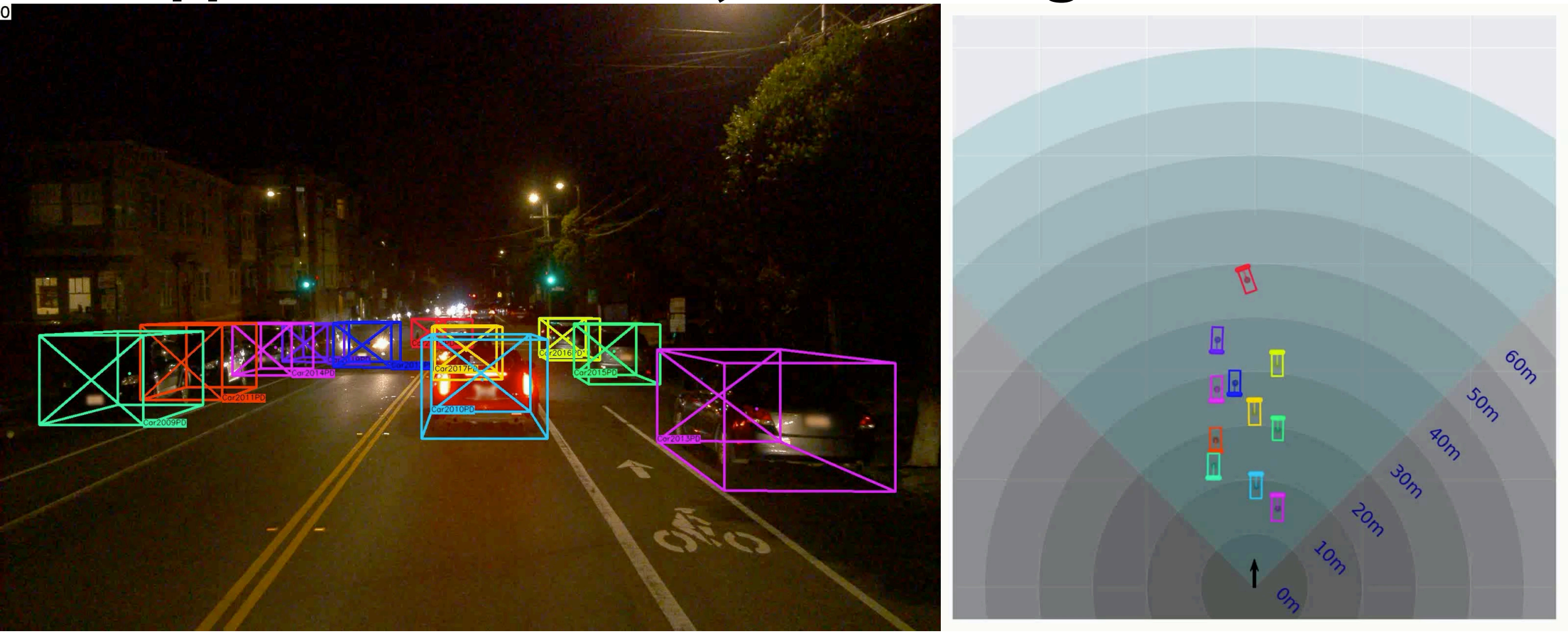
Application on 3D Object Tracking



Results on Waymo Dataset

Hu, Cai, Wang, Lin, Sun, Krähenbühl, Darrell, Yu, *Joint Monocular 3D Vehicle Detection and Tracking*, ICCV 2019

Application on 3D Object Tracking



Results on Waymo Dataset

Hu, Cai, Wang, Lin, Sun, Krähenbühl, Darrell, Yu, *Joint Monocular 3D Vehicle Detection and Tracking*, ICCV 2019

Tracking

- Track a point
- Track a bigger box
- Track by detection
- Online learning
- Motion
- Multiple object tracking
- 3D object tracking