Master in
Computer Vision
Barcelona

Project
Module 6
Coordination

Week 4: Final Report

Video Surveillance for Road Traffic Monitoring

Cristian Gutiérrez, Iñaki Lacunza,
Carlos Boned, Marco Cordón

GROUP 6

Master in
Computer Vision
Barcelona

Index

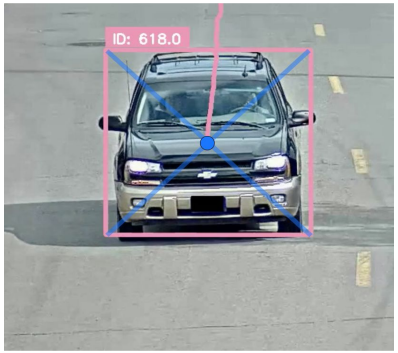
Task 1.1: Speed Estimation

CRISTIAN GUTIÉRREZ
1532397

10 FPS

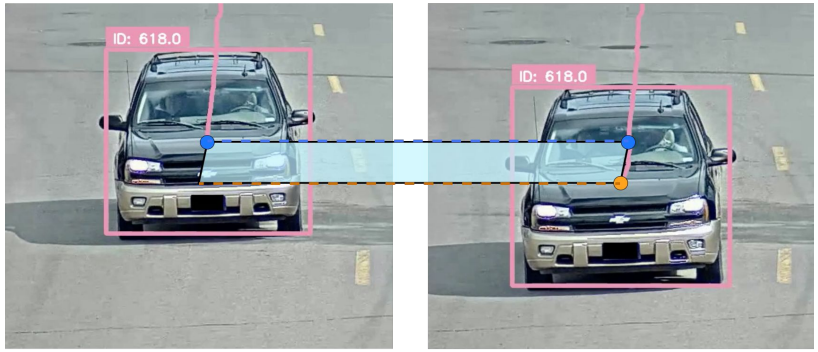
$$velocity = \underbrace{displacement * FPS}_{pixel/seconds} * \underbrace{meters/pixel * 3.6}_{km/h}$$

#1 Centroid



$$c = \left(\frac{x_{min} + x_{max}}{2}, \frac{y_{min} + y_{max}}{2} \right)$$

#2 Displacement



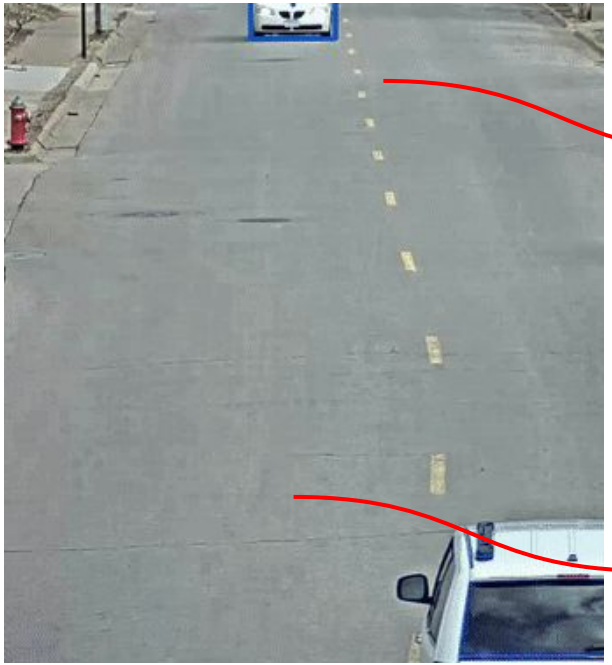
$$\sqrt{(cx_2 - cx_1)^2 + (cy_2 - cy_1)^2}$$

Task 1.1: Speed Estimation

CRISTIAN GUTIÉRREZ
1532397

Assumption #1: Fixed meters/pixel

$$velocity = displacement * FPS * \underbrace{0.05\ meters/pixel * 3.6}_{km/h}$$



PROBLEM!

For a car in a constant~ velocity, as it approaches the camera it increases the velocity.

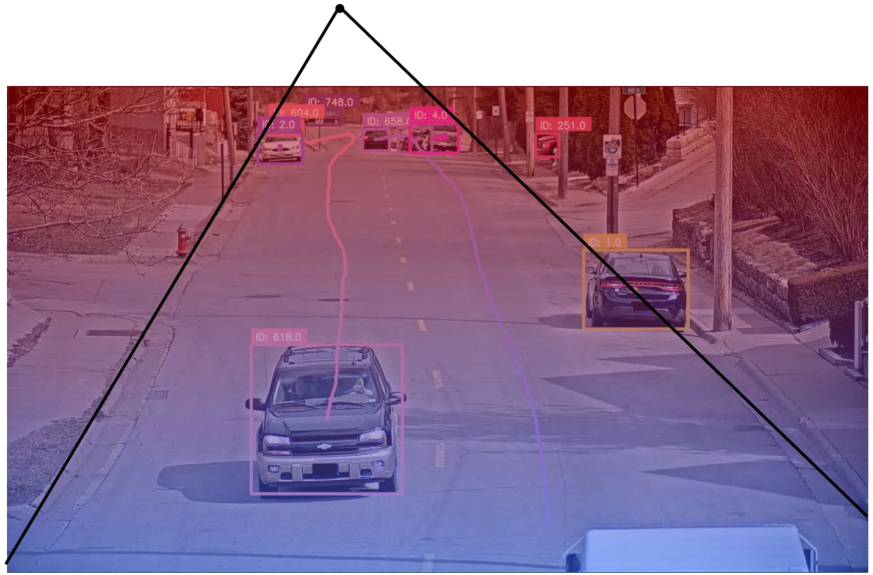
This is due to the perspective of the camera!

Task 1.1: Speed Estimation

CRISTIAN GUTIÉRREZ
1532397

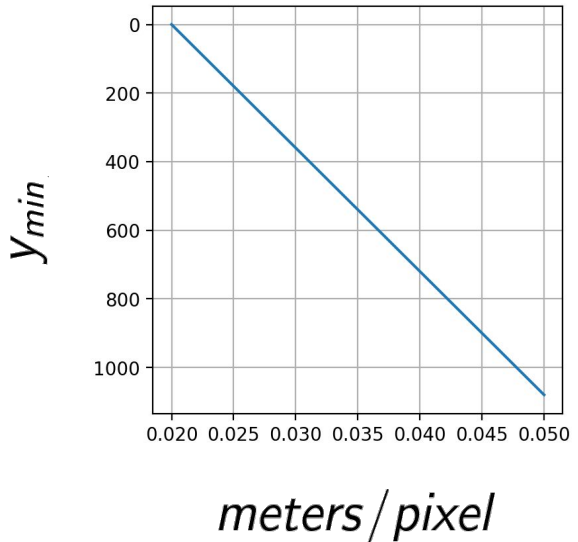
Assumption #2: Linear meters/pixel

$$velocity = displacement * FPS * f(y_{max}) meters/pixel * 3.6$$



LINEAR FUNCTION

$$a * y_{max} + b$$

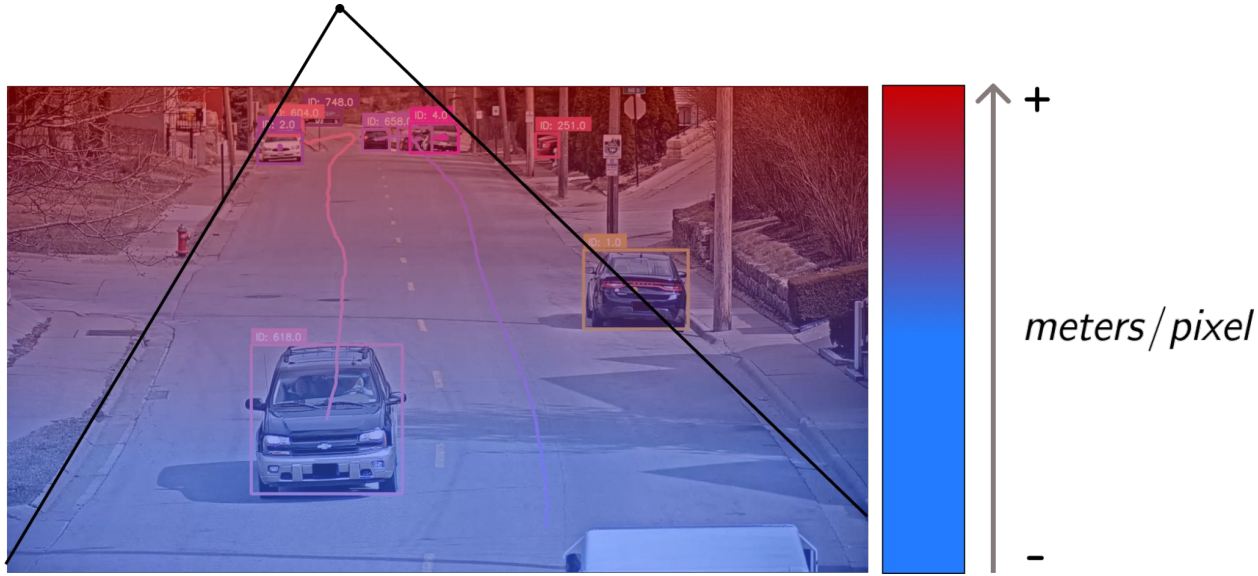


Task 1.1: Speed Estimation

CRISTIAN GUTIÉRREZ
1532397

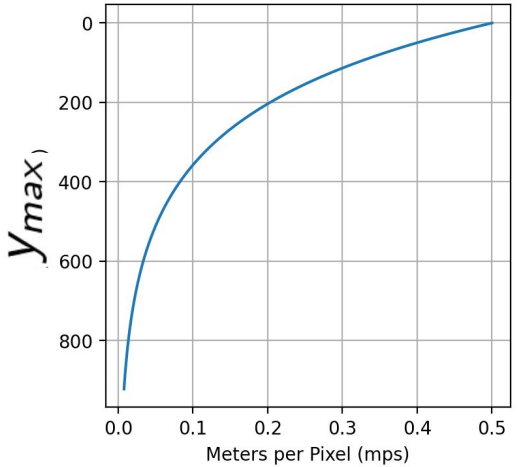
Assumption #3: Exponential meters/pixel

$$velocity = displacement * FPS * f(y_{max}) meters/pixel * 3.6$$



EXPO FUNCTION

$$e^{a*y_{max}} * b$$



$$meters/pixel$$

Task 1.1: Speed Estimation

CRISTIAN GUTIÉRREZ
1532397

Results

$$velocity = displacement * FPS * f(y_{max}) meters/pixel * 3.6$$

FIXED MPS=0.05

$a * y_{max} + b$

$e^{a*y_{max}} * b$

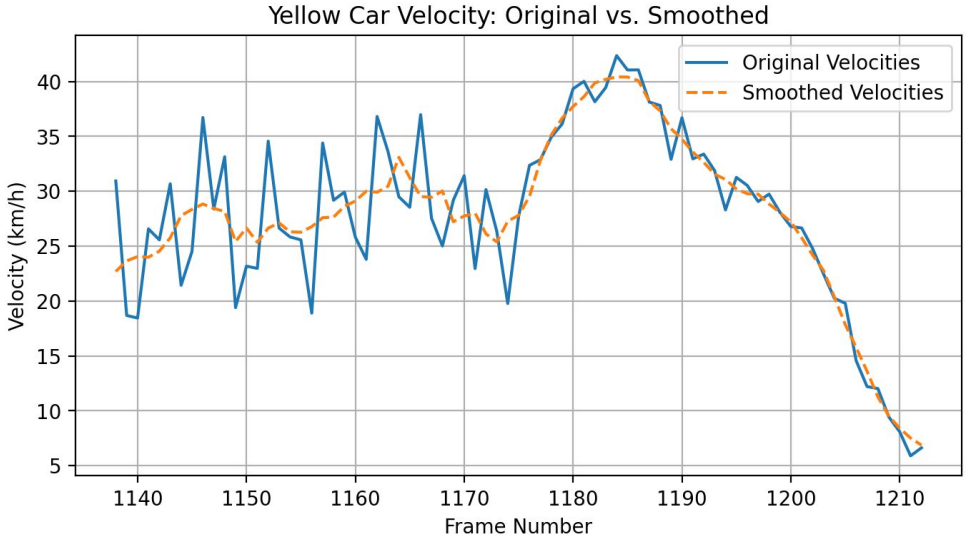
Task 1.1: Temporal Smoothing

CRISTIAN GUTIÉRREZ
1532397

$$velocity = displacement * FPS * \underbrace{f(y_{max})}_{e^{a*y_{max}} * b} meters/pixel * 3.6$$

We “self-calibrated” our sequence camera

WINDOW = 5



Task 1.1: Remove Parked Cars

CRISTIAN GUTIÉRREZ
1532397

$$velocity = displacement * FPS * \underbrace{f(y_{max})}_{e^{a*y_{max}} * b} meters/pixel * 3.6$$

x_{max}

frame count > 25

	Avg speed (km/h)
Vehicle 1	41.86
Vehicle 2	52.39
Vehicle 3	30.75
Vehicle 4	33.74
Vehicle 5	20.67
Vehicle 6	29.07
Vehicle 7	48.48

	Avg speed (km/h)
Vehicle 8	42.36
Vehicle 9	30.91
Vehicle 10	38.66
Vehicle 11	24.72
Vehicle 12	37.32
Vehicle 13	36.37
Vehicle 14	38.07

To remove parked cars

WINDOW = 10

THRESHOLD = 10

```
velocity = compute_velocity(...)
smooth_velocity = moving_avg(w=10)
if smooth_velocity < 10:
    continue
```

FURTHER WORK:
CALIBRATE ALSO FOR X-AXIS AND FOR FURTHER AWAY CARS

Carlos Boned
1533190

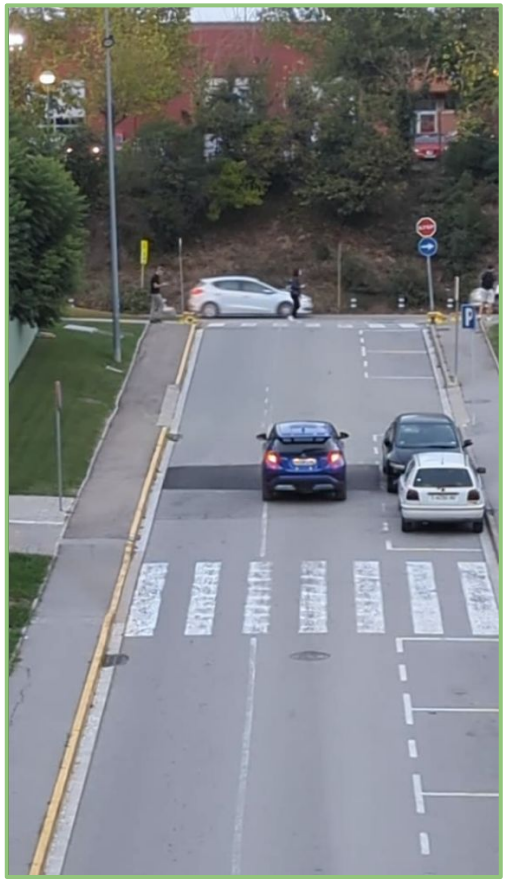
Task 1.2: Customize speed detection

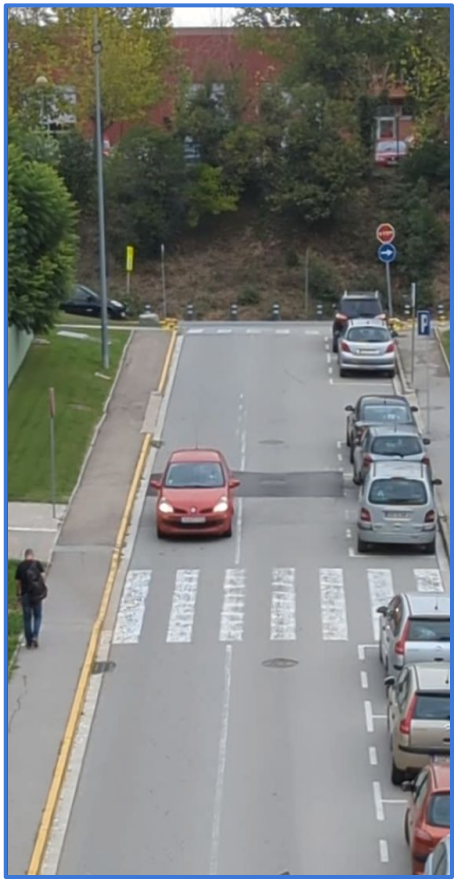
Custom Sequences:

- **CVC SMALL**: 5 min video, in the evening (less traffic)
- **CVC BIG**: 25 min video, morning (high traffic)
- **EEUU HIGHWAY**: 25 s video, road to measure use case

DIFFERENT PERSPECTIVES !!

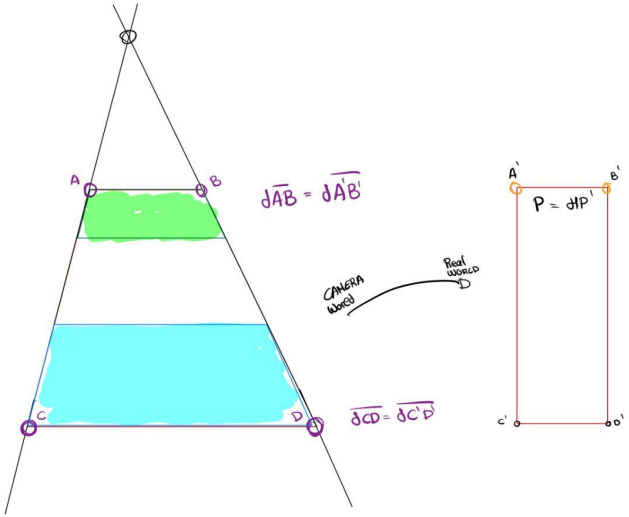
WE ASSUME CALIBRATED CAMERA

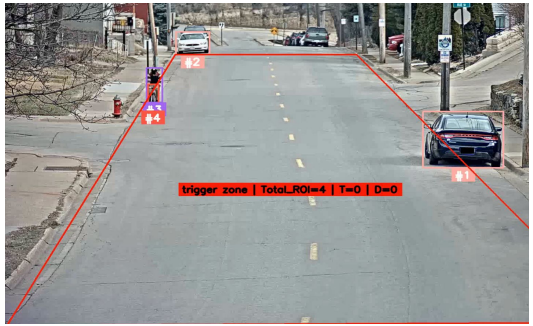


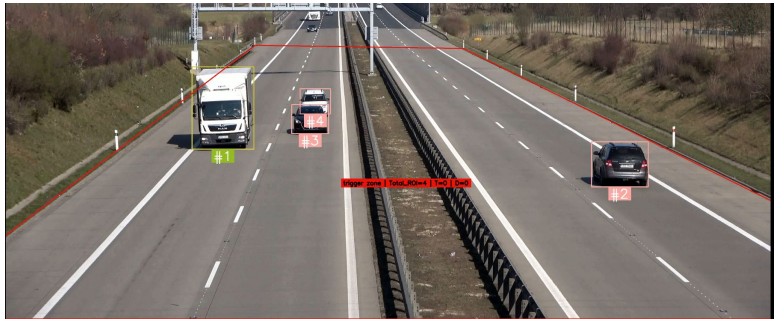


Carlos Boned
1533190

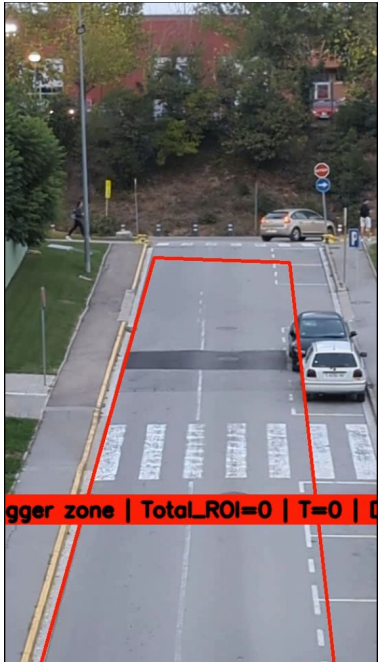
Task 1.2: Overfit Perspective Different Static Cameras







DIFFERENT PERSPECTIVES !!
DIFFERENT HOMOGRAPHY

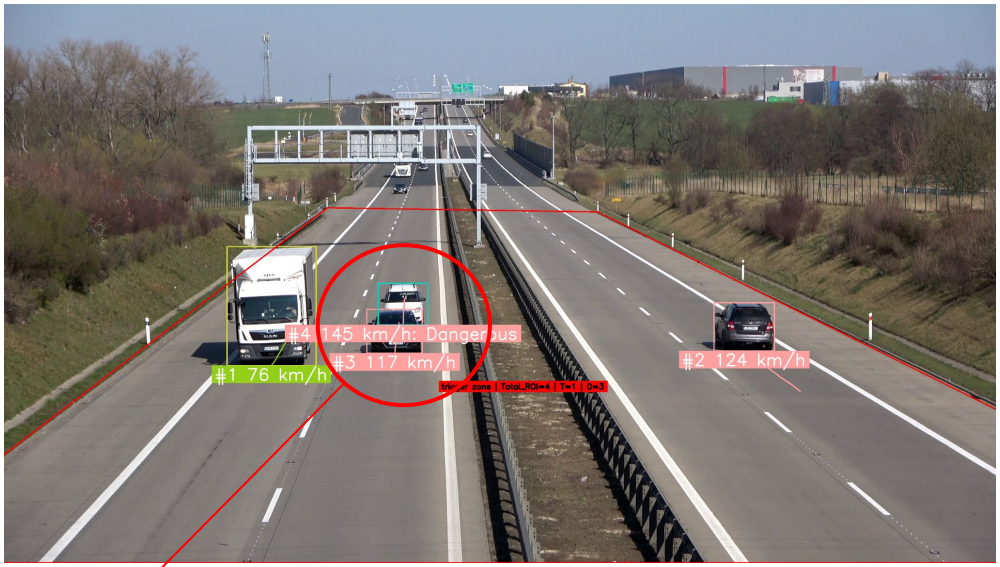


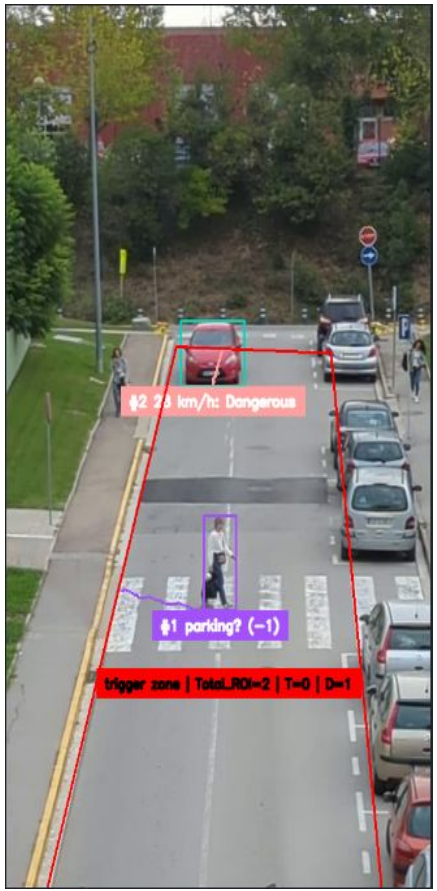
We have cameras with some perspective distortion, parallel lines are no longer parallels in the camera world, we extract the Homography to recover the parallelism

In this case we try to simulate a traffic camera, which is static and overfitted in the place it has been put. The polygon range is computed manually and it also detects de traffic jam.

Carlos Boned
1533190

Task 1.2: Detect Dangerous Situations





It also detects situation of an exceed of velocity and there are an object to track in the trigger zone, in this case a person

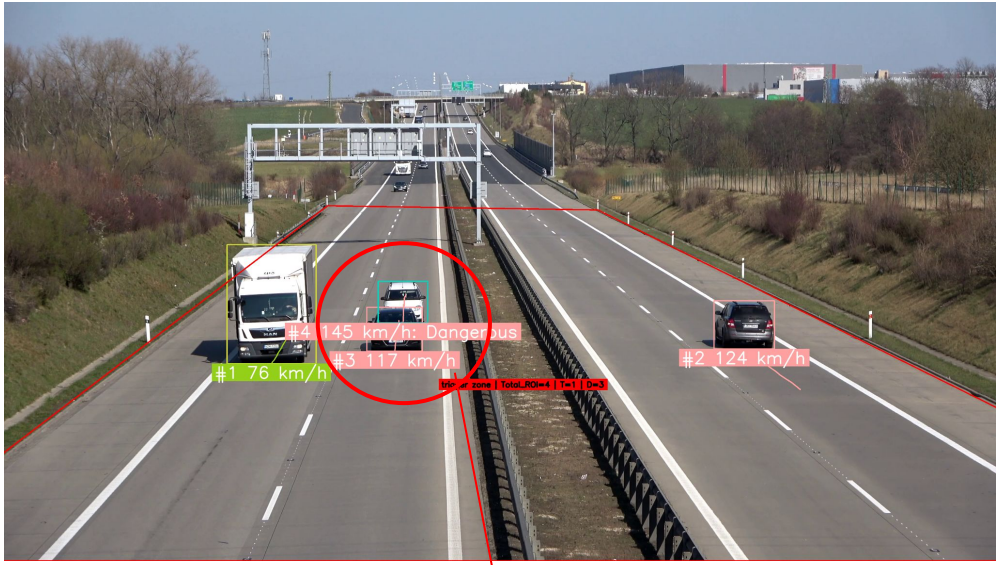
LOG

We detect dangerous situations, as for example proximity or speed.
We collect the frame instance and send it to a log file with all the information belonging to that car

As the we only care about cars, trucks etc, the label of the other objects is (-1)

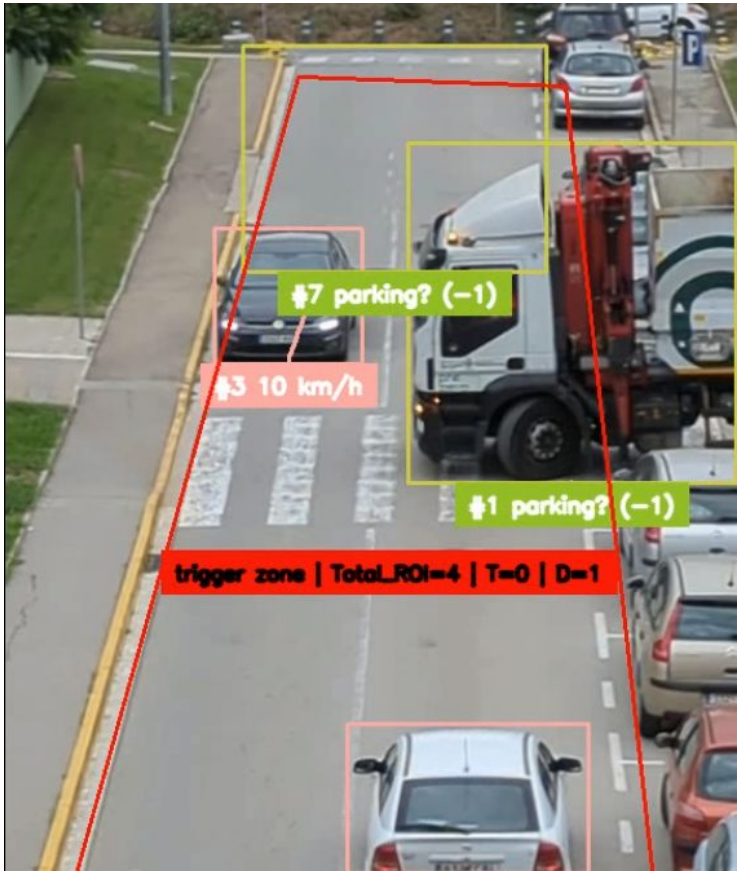
Carlos Boned
1533190

Task 1.2: Difficulties



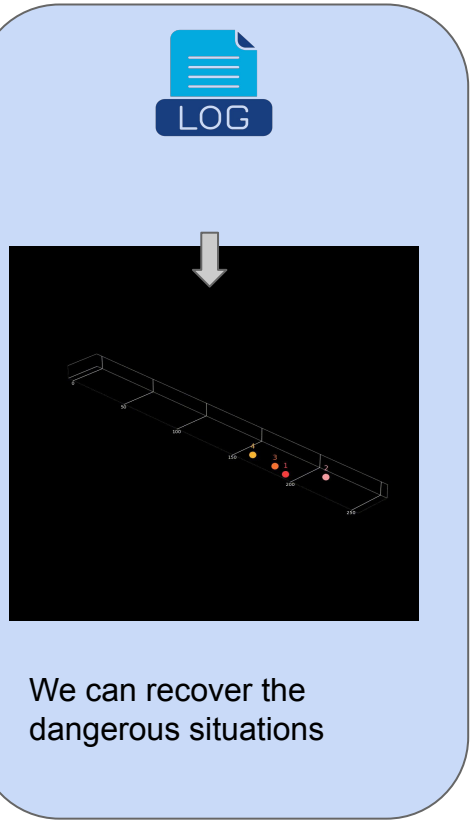
We found some difficulties when there are occlusions, the bbox had some problems in the detection, which makes the results in the computation of the speed weird.

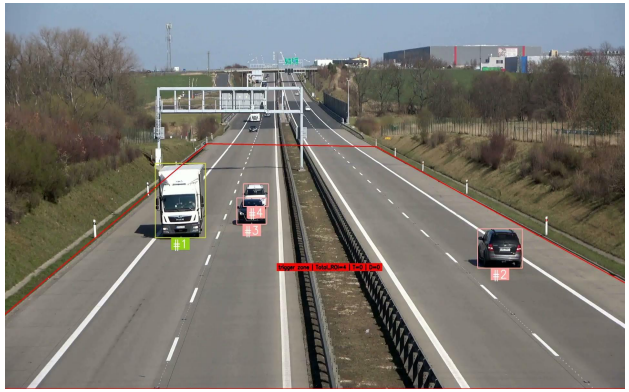
In this case exists an extra detection with a different track id, I didn't figure it out why it happens.

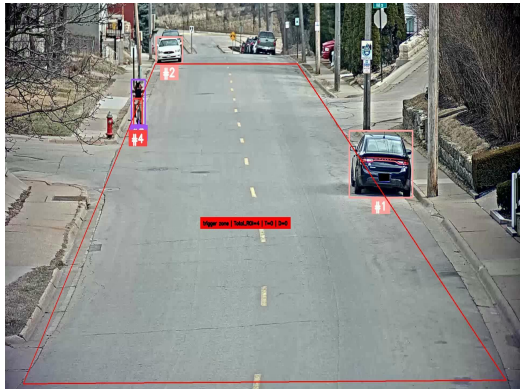


Carlos Boned
1533190

Task 1.2: Demos in different Sequences







Task 2:

Multi-camera tracking

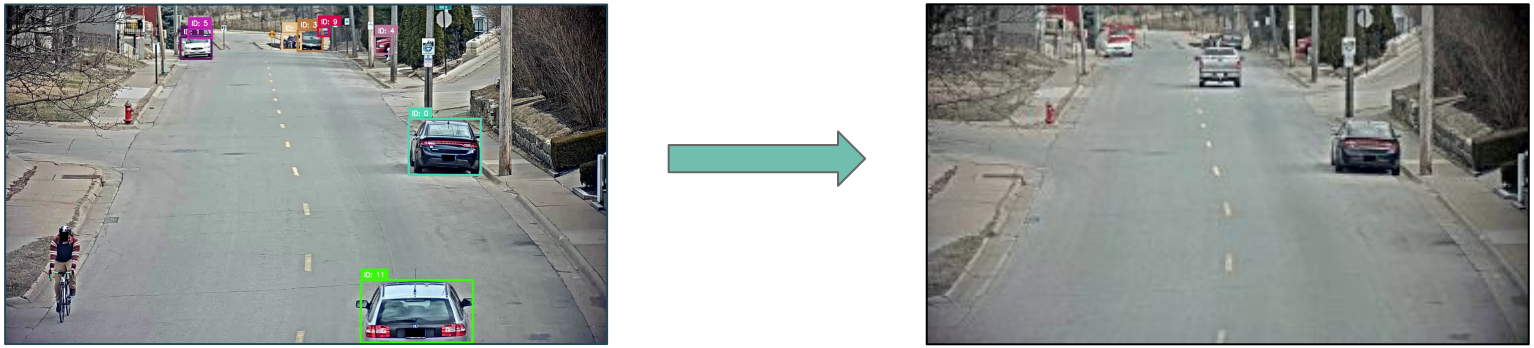
Task 2: Multi-camera tracking

MARCO CORDÓN
1691727

- RESULTS FROM LAST WEEK:

IDF1 / HOTA (SEQ03)						
c010	**c011**	**c012**	**c013**	**c014**	**c015**	**Average**
18.0/17.9	1.5/4.7	0.75/3.2	6.1/9.5	15.1/17.6	0.12/1.5	6.92/9.07

- CURRENT WEEK ADDITION: **Remove parked cars (Post-processing)**



Task 2: Multi-camera tracking

MARCO CORDÓN
1691727

- RESULTS FROM LAST WEEK:

IDF1 / HOTA (SEQ03)				
c010	**c011**	**c012**	**c013**	**c014**
18.0/17.9	1.5/4.7	0.75/3.2	6.1/9.5	15.1/17.6

- CURRENT WEEK ADDITION: **Remove parked cars**

Task 2: Multi-camera tracking

MARCO CORDÓN
1691727

- RESULTS FROM LAST WEEK:

IDF1 / HOTA (SEQ03)						
c010	**c011**	**c012**	**c013**	**c014**	**c015**	**Average**
18.0/17.9	1.5/4.7	0.75/3.2	6.1/9.5	15.1/17.6	0.12/1.5	6.92/9.07

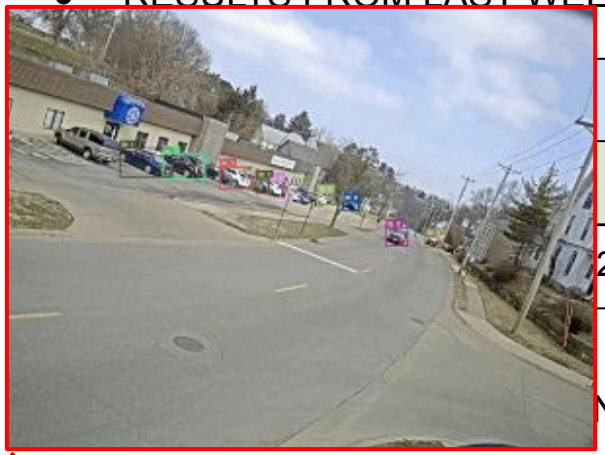
- CURRENT WEEK ADDITION: **Remove parked cars (Post-processing)**

IDF1 / HOTA (SEQ03)						
c010	**c011**	**c012**	**c013**	**c014**	**c015**	**Average**
57.7/36.6	5.33/5.82	9.69/11.5	33.4/24.0	52.18/38.6	7.42/12.04	27.6/21.3

Task 2: Multi-camera tracking

MARCO CORDÓN
1691727

- RESULTS FROM LAST WEEK:



IDF1 / H						
	c					Average
2	6.1					6.92/9.07

N: **Remov**

IDF1 / HOTA (SEQ03)						
c010	**c011**	**c012**	**c013**	**c014**	**c015**	**Average**
57.7/36.6	5.33/5.82	9.69/11.5	33.4/24.0	52.18/38.6	7.42/12.04	27.6/21.3

Task 2: Multi-camera tracking

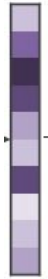
MARCO CORDÓN
1691727

Proposed idea:

Use metric learning for deciding if tracks in different cameras belong to the same car or not.

Faster-RCNN

Faster-RCNN



Distance lower than **th**

SAME CARS !!

Task 2: Multi-camera tracking

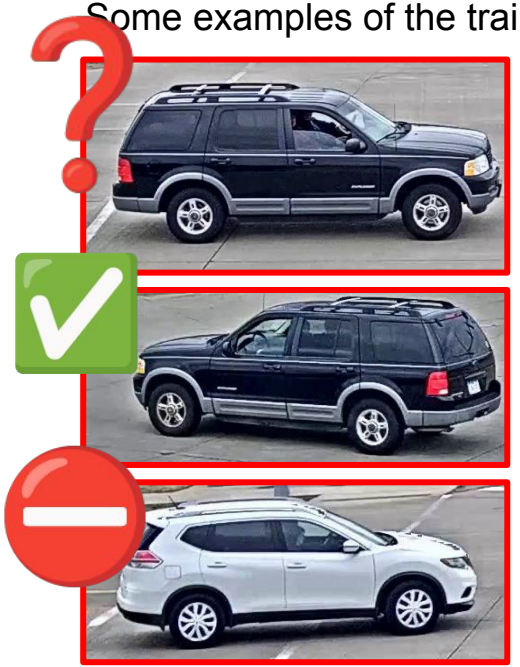
MARCO CORDÓN
1691727

Model training strategy:

Use ground truth from S01 and S04

Zheng Tang, et al., "CityFlow: A City-Scale Benchmark for Multi-Target Multi-Camera Vehicle Tracking and Re-Identification" CVPR 2019.

Some examples of the training instances:



Issue: Ground truth bounding box is different from Detectron2's bounding box detection

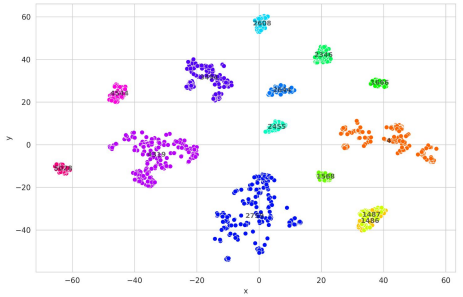
Detectron2 detection bb

Ground truth bb

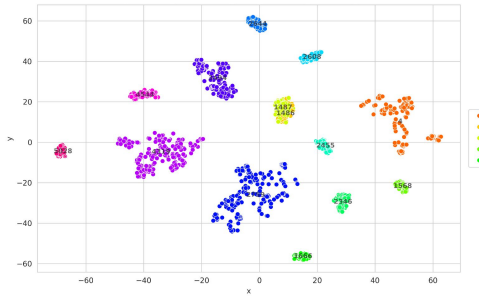
Task 2: Multi-camera tracking

IÑAKI LACUNZA
1692285

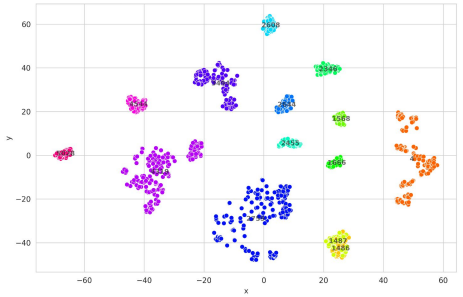
Margin 20%



Margin 10%



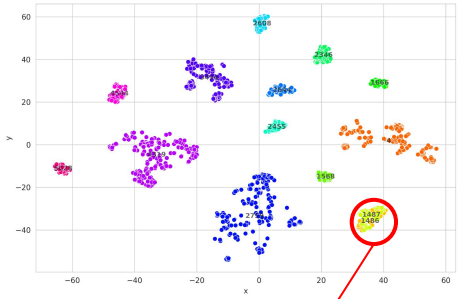
Margin 6,66%



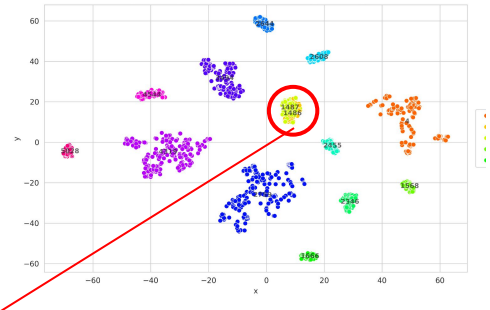
Task 2: Multi-camera tracking

IÑAKI LACUNZA
1692285

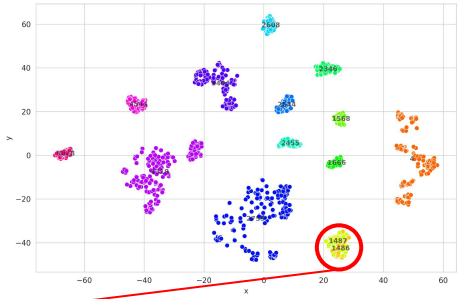
Margin 20%

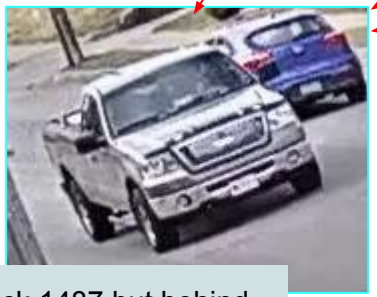


Margin 10%



Margin 6,66%





Track 1487 but behind appears the 1486.

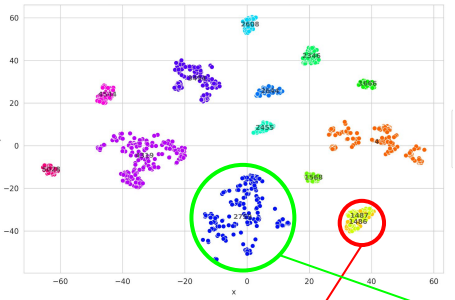
Task 2: Multi-camera tracking

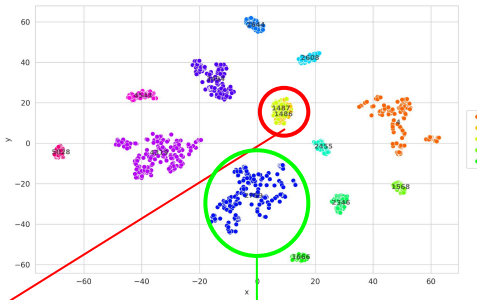
IÑAKI LACUNZA
1692285

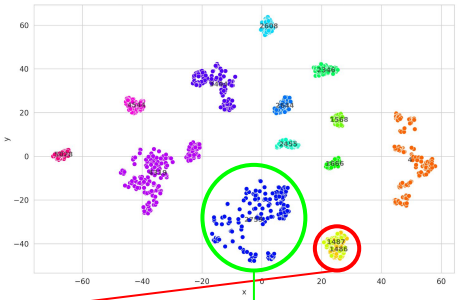
Margin 20%

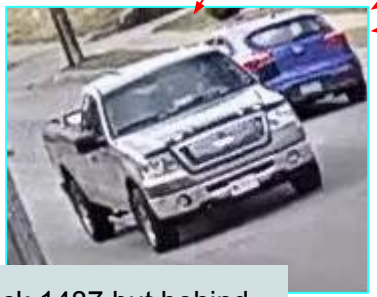
Margin 10%

Margin 6,66%









Track 1487 but behind appears the 1486.

First detections

Intermediate detections

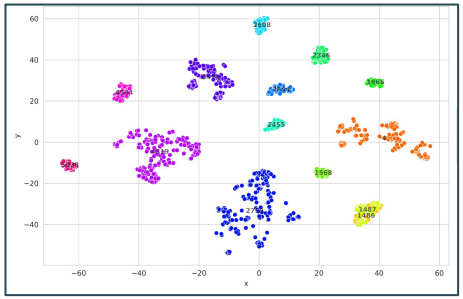
Last detections

Task 2: Multi-camera tracking

Followed pipeline:

IÑAKI LACUNZA 1692285

Get the embeddings with the first camera and apply KNN



Next camera

Single camera tracking.

Then, get embeddings for each track.

Use the KNN model to know which are the closest.

Chose the neighbor that is **closest to most detections** of the track.

Then, get the **median** of the distances to the chosen threshold.

Median distance lower than **chosen threshold?**

Same car in different cameras!!
Make their *track id*s the same

New track, not seen yet!!
Keep its original *track id*

Add the embeddings and fit again the KNN

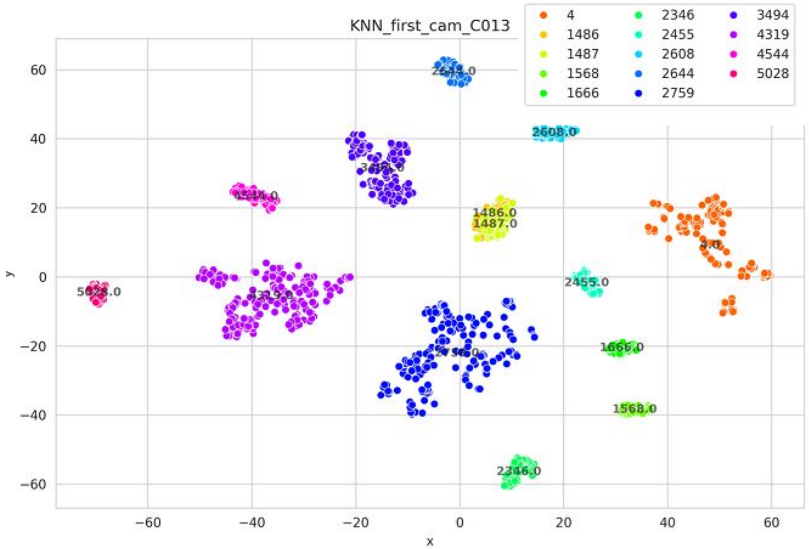
Next camera

Task 2: Multi-camera tracking

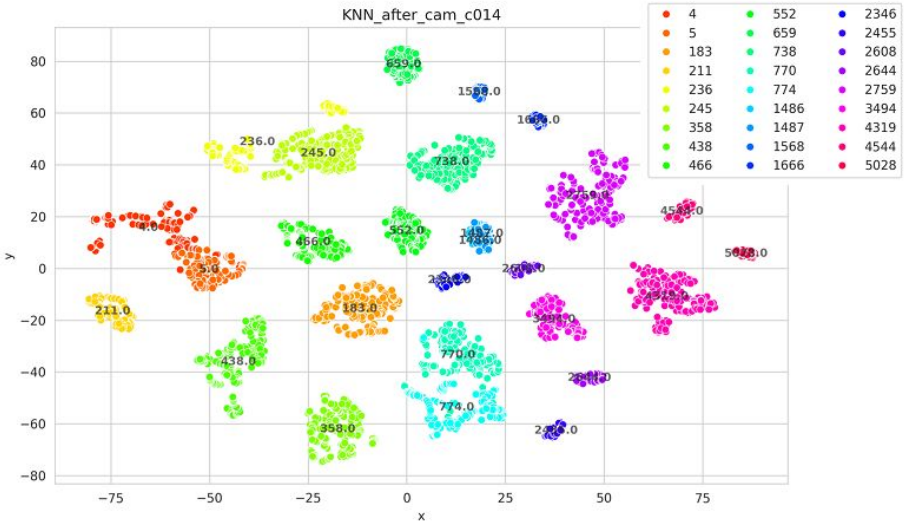
IÑAKI LACUNZA
1692285

KNN evolution: c013 ⟫⟫⟫⟫⟫⟫⟫ c014

Previous KNN



New KNN

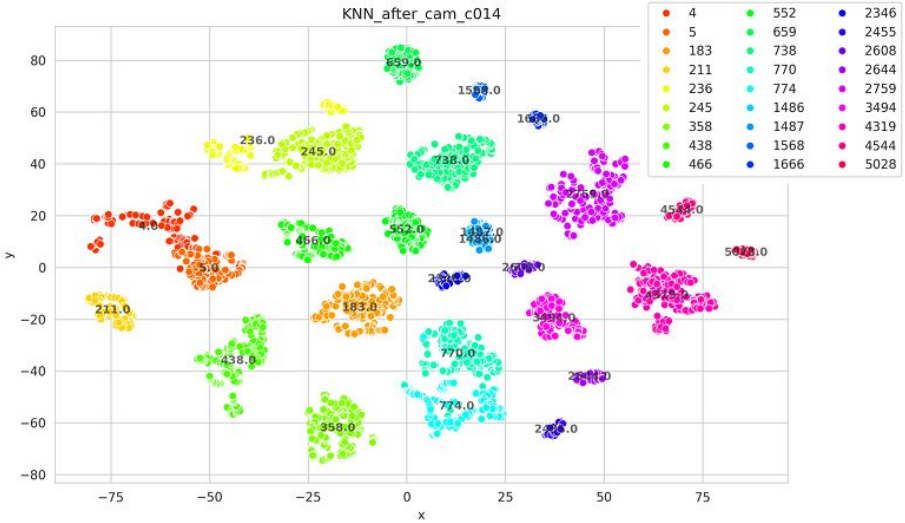


Task 2: Multi-camera tracking

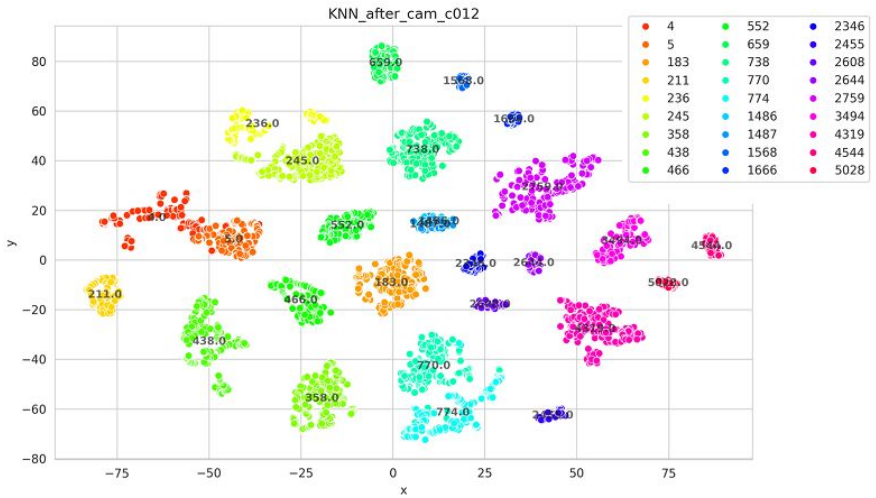
IÑAKI LACUNZA
1692285

KNN evolution: c014 ⟫⟫⟫⟫⟫⟫⟫ c012

Previous KNN



New KNN

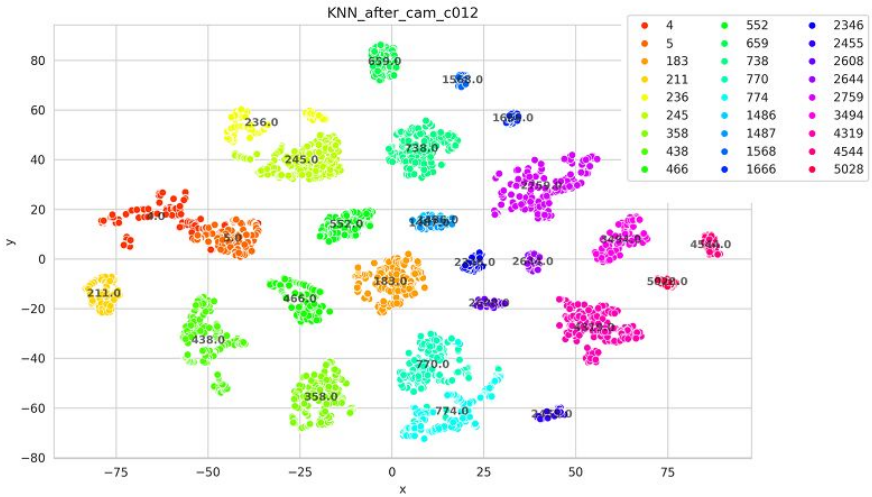


Task 2: Multi-camera tracking

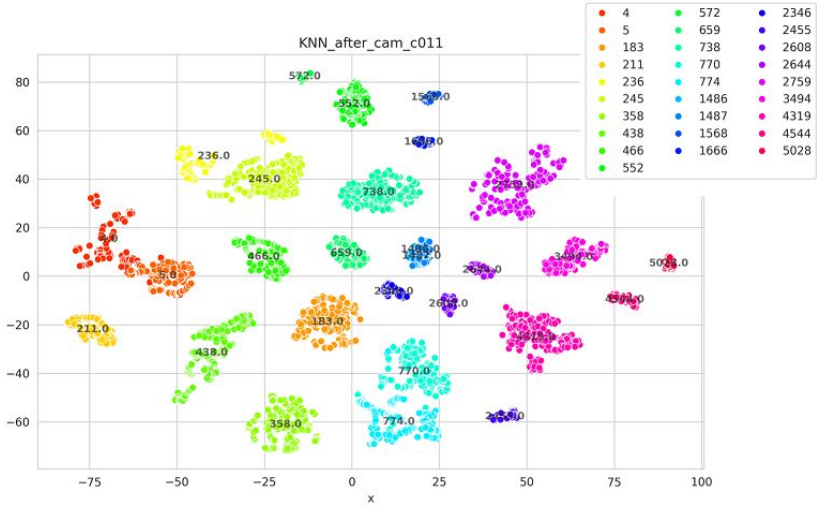
IÑAKI LACUNZA
1692285

KNN evolution: c012 ⟫⟫⟫⟫⟫⟫⟫ c011

Previous KNN



New KNN

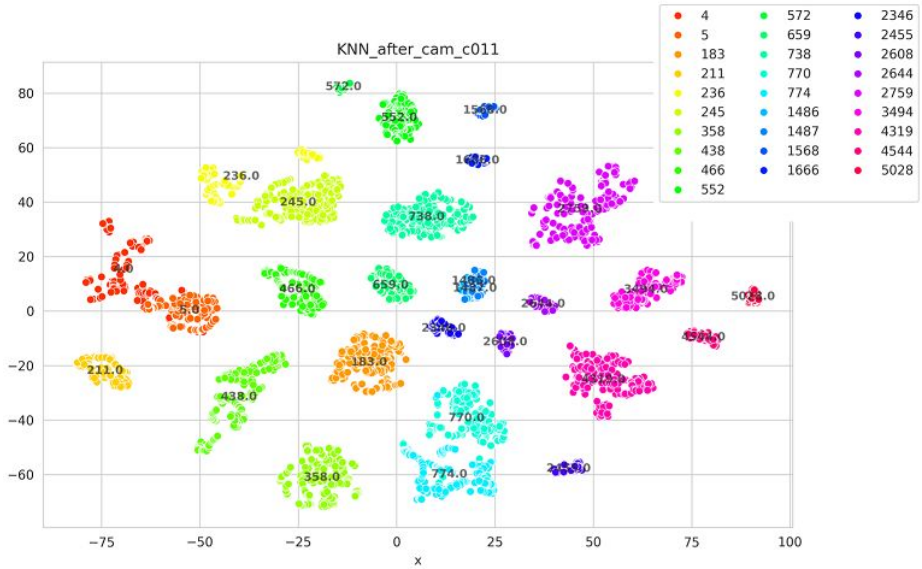


Task 2: Multi-camera tracking

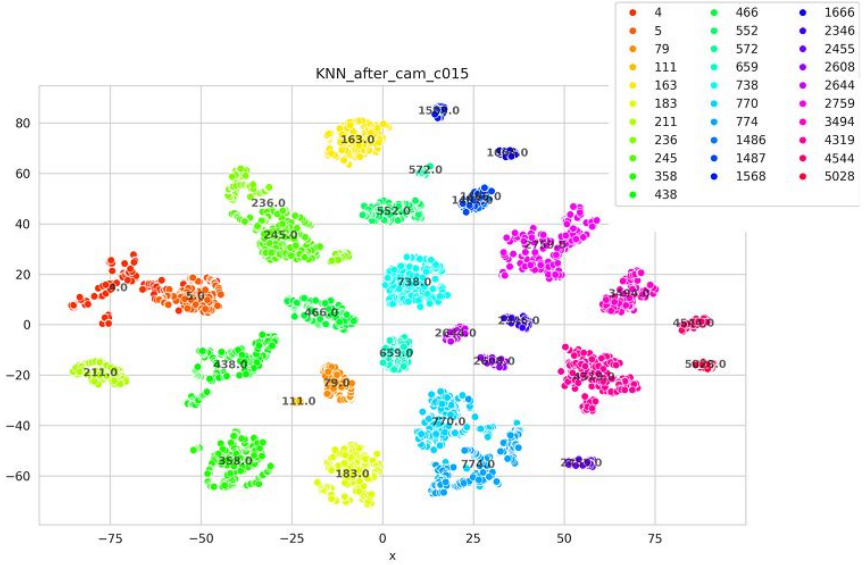
IÑAKI LACUNZA
1692285

KNN evolution: c011 ⟫⟫⟫⟫⟫⟫⟫ c015

Previous KNN



New KNN

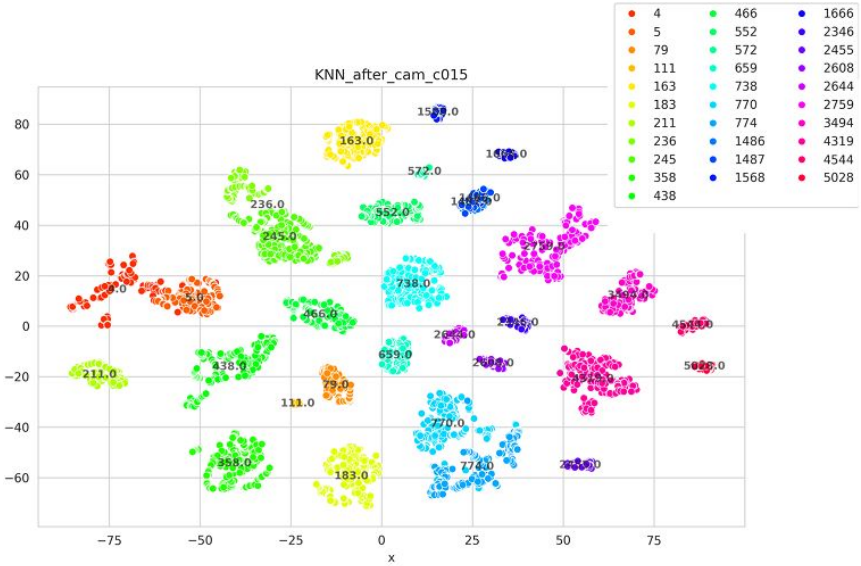


Task 2: Multi-camera tracking

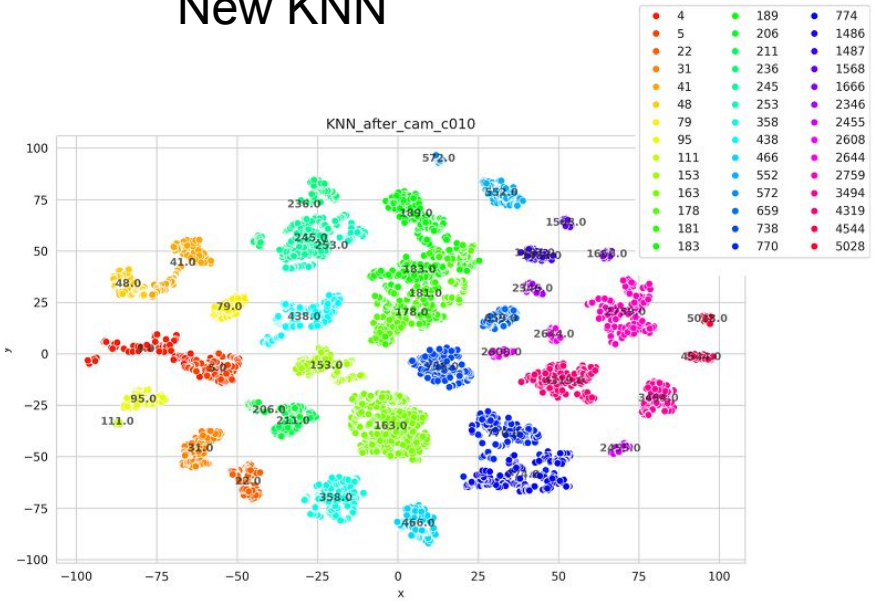
IÑAKI LACUNZA
1692285

KNN evolution: c015 ⟫⟫⟫⟫⟫⟫⟫ c010

Previous KNN



New KNN



Task 2: Multi-camera tracking

Qualitative Results:

Misclassified, might be because they are both trucks and share similarity in their shape

This would be a quite easy match to do, but the car in the left was not detected. Previous weeks' weaknesses have played an important part this week.

Misclassification, shape and color are very similar

Quantitative Results:

IDF1/HOTA
Avg S03
27.6/21.3

Only results of S03 because S01 and S04 are employed for training so they would be false results.

Final Summary

Task 1.1: Rudimentary approach

- Introduce horizontal axis into the mps estimation
- Take advantage of existing calibration methods instead.

Task 1.2: Modern approach

- We manage to apply our methods in videos with different perspectives with a decent results.
- We could extend the speed estimation with the detection of dangerous situations.

Task 2: Multi-camera tracking

- Due to the added preprocessing step, we were able to improve ~25% the results from last week.
- At first, the implemented idea looked fine, but the final results were not the expected.
- First, single-camera tracking should be improved.
- For future improvements, apart from comparing tracks using the implemented threshold, other aspects such as **time difference**, **location between cameras** and **speed estimation** could be used.