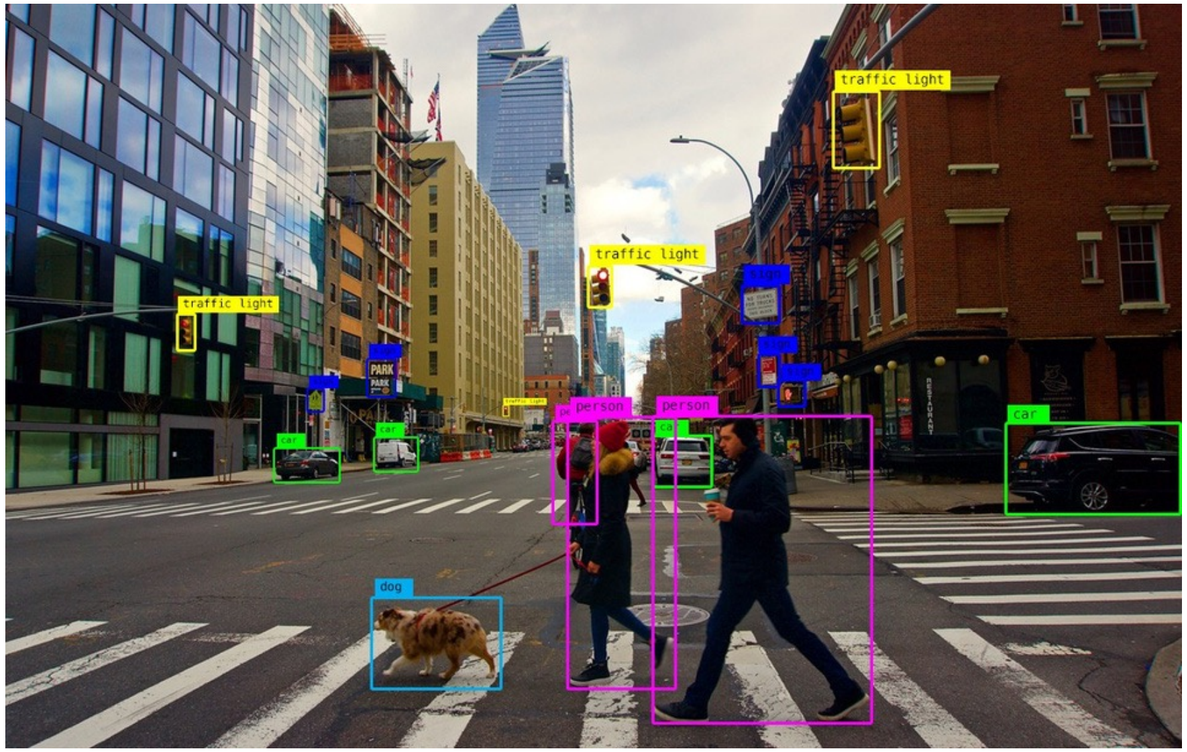
Computer Vision & Cognitive System Project

UNIMORE

Multi-Vision System for Autonomous Driving



STUDENTS

Daniel Rossi

Riccardo Salami

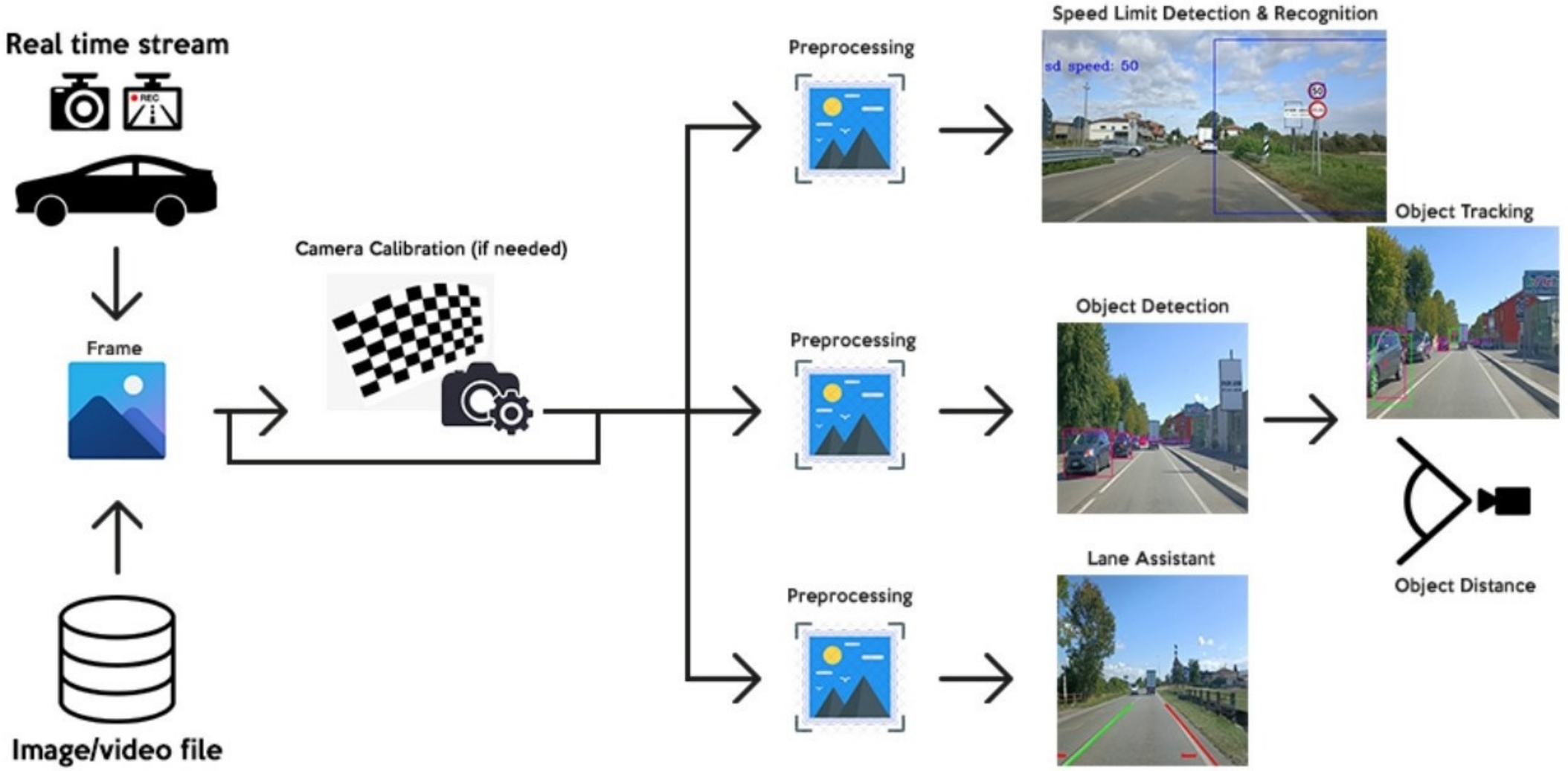
Filippo Ferrari

TUTORS

Roberto Amoroso

Silvia Cascianelli

Pipeline

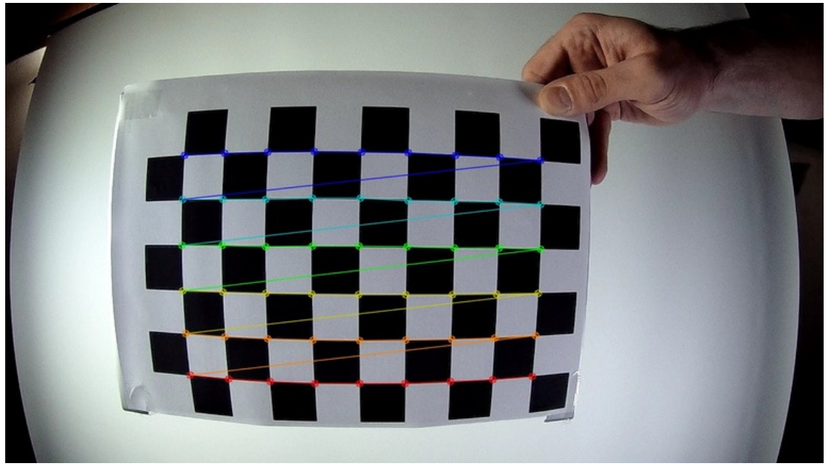

Real time stream
Frame
Image/video file
Camera Calibration (if needed)
Preprocessing
Preprocessing
Preprocessing
Speed Limit Detection & Recognition
sd speed: 50
Object Detection
Lane Assistant
Object Tracking
Object Distance

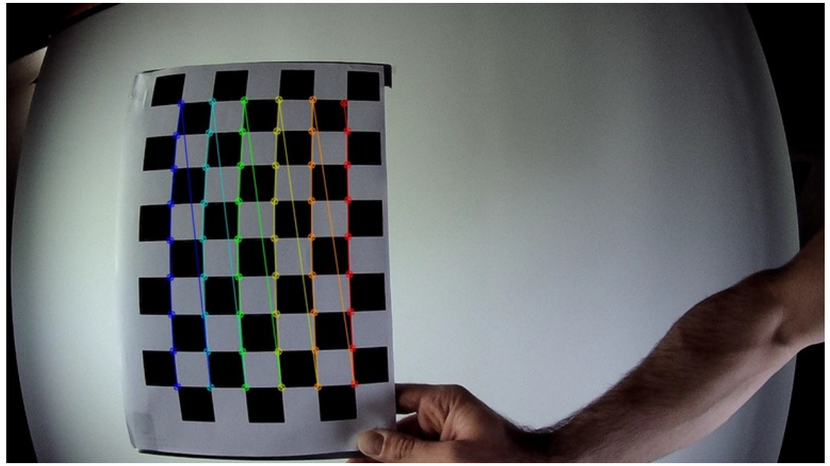
Classical Techniques

Camera Calibration

The wide-angle lens of the camera we use creates a barrel distortion, so it was necessary to correct the image through camera calibration.

For the calibration task, we used Zhang's technique: we printed a chessboard and we acquired dozens of images of it, in order to estimate the camera parameters and obtain undistorted images.

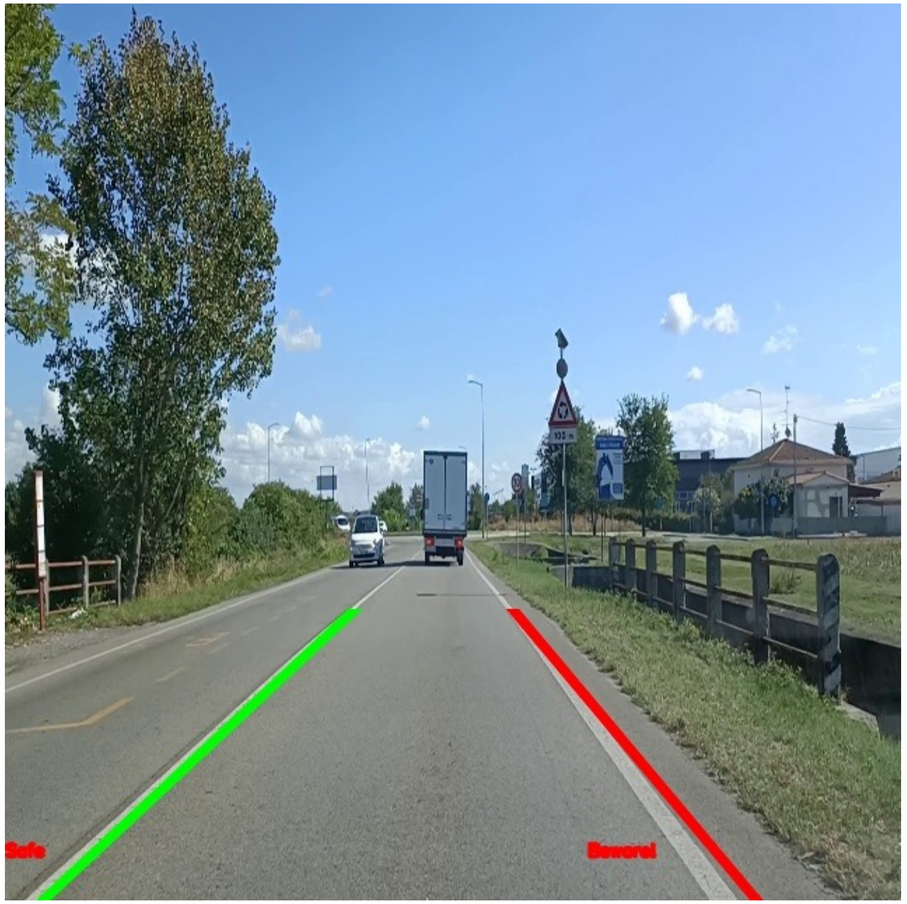




Lane assistant

The lines at the sides of the road are detected and the driver is notified if he's crossing either the left or the right one, preventing him from endangering him and the other passengers.

The trickiest problem for this part was surely the removal of noise.
First, we remove the grass from the image, this removal is done using HSV color-space.
Then, based on the mean luminance of the entire image, we decide whether to enable Equalization or not and the parameters to use for the bilateral filter.
Finally, we apply Canny and we look for possible lines through the Hough transform.

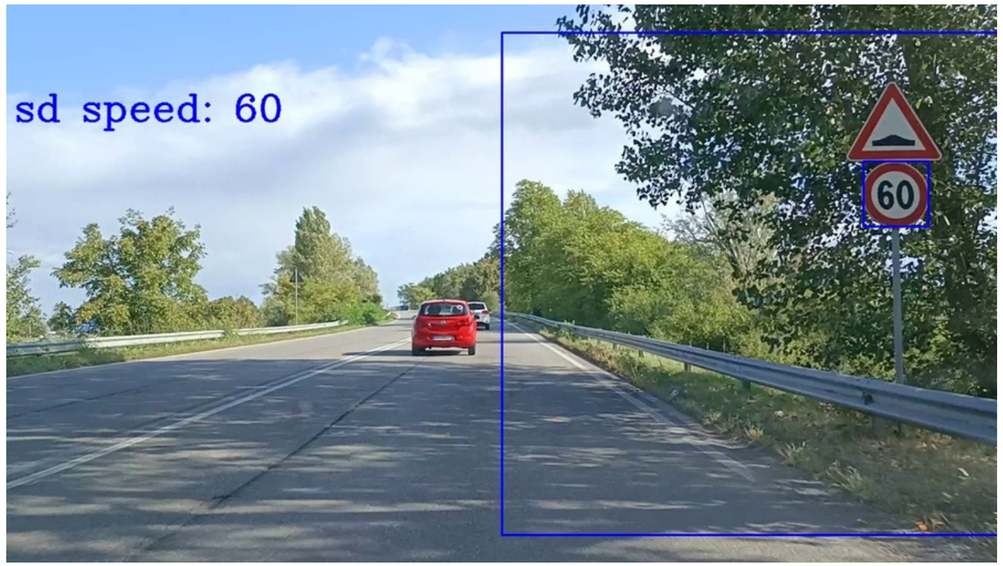


Speed Limit Detection

We are able to locate the region of the sign with high accuracy by taking advantage of the red color present in the edge of the sign and using Hough.

We have collected a series of reference street sign images, specifically one for each speed limit that may be encountered on the street. We then pass these images to SIFT, along with the newly acquired and preprocessed street image, to extract keypoints.

Once the keypoints are extracted, we use the KNN algorithm to classify the main image.



ItalianSigns

ItalianSigns is a dataset containing images of the road context captured by phone camera. It contains 362 labelled images: each image contains a road sign (speed limit only), of which the number associated with the speed limit, and the bounding box corresponding to the ROI of the sign were annotated.

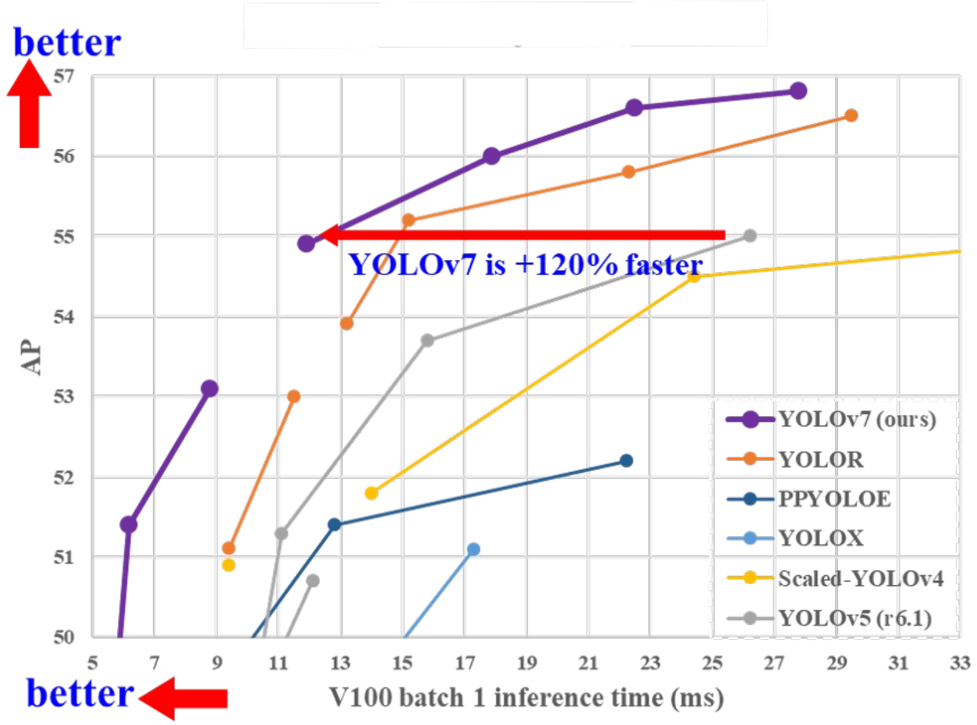
To recognize the road signs, we applied preprocessing techniques clean the image, used HoughCircles to find the spatial location of the sign and extract the bounding box, and used a KNN on feature vectors extracted via SIFT between ground truth images, and the crop of the inference image

kaggle

https://www.kaggle.com/datasets/officialprojecto/italiansigns

Deep Learning Techniques

YOLOv7

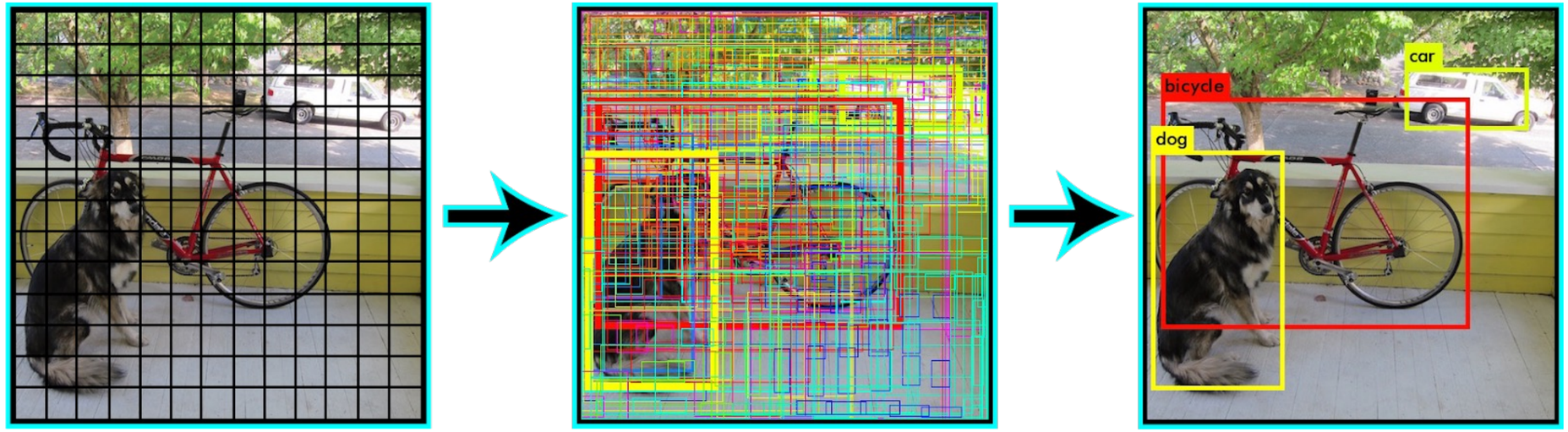


YOLO is one of the premiere object detection networks.

YOLOv7 is the latest version of the popular algorithm, and improves significantly on its predecessors.

A number of new changes were made for YOLOv7:

- the model scales the network depth and width simultaneously while concatenating layers together
- the technique "Extended efficient layer aggregation networks" or E-ELAN is used to accelerate inference time
- used gradient flow propagation paths to analyze how re-parameterized convolution should be combined with different networks
- uses the lead head prediction as guidance to generate coarse-to-fine hierarchical labels, which are used for auxiliary head and lead head learning, respectively



How does YOLO work ?

YOLO works to perform object detection in a single stage by first separating the image into N grids. For each grid, bounding box coordinates, B, for the potential object(s) are predicted with an object label and a probability score for the predicted object's presence.

To handle this redundancy and reduce the predicted objects down to those of interest, YOLO uses Non-Maximal Suppression to suppress all the bounding boxes with comparatively lower probability scores.

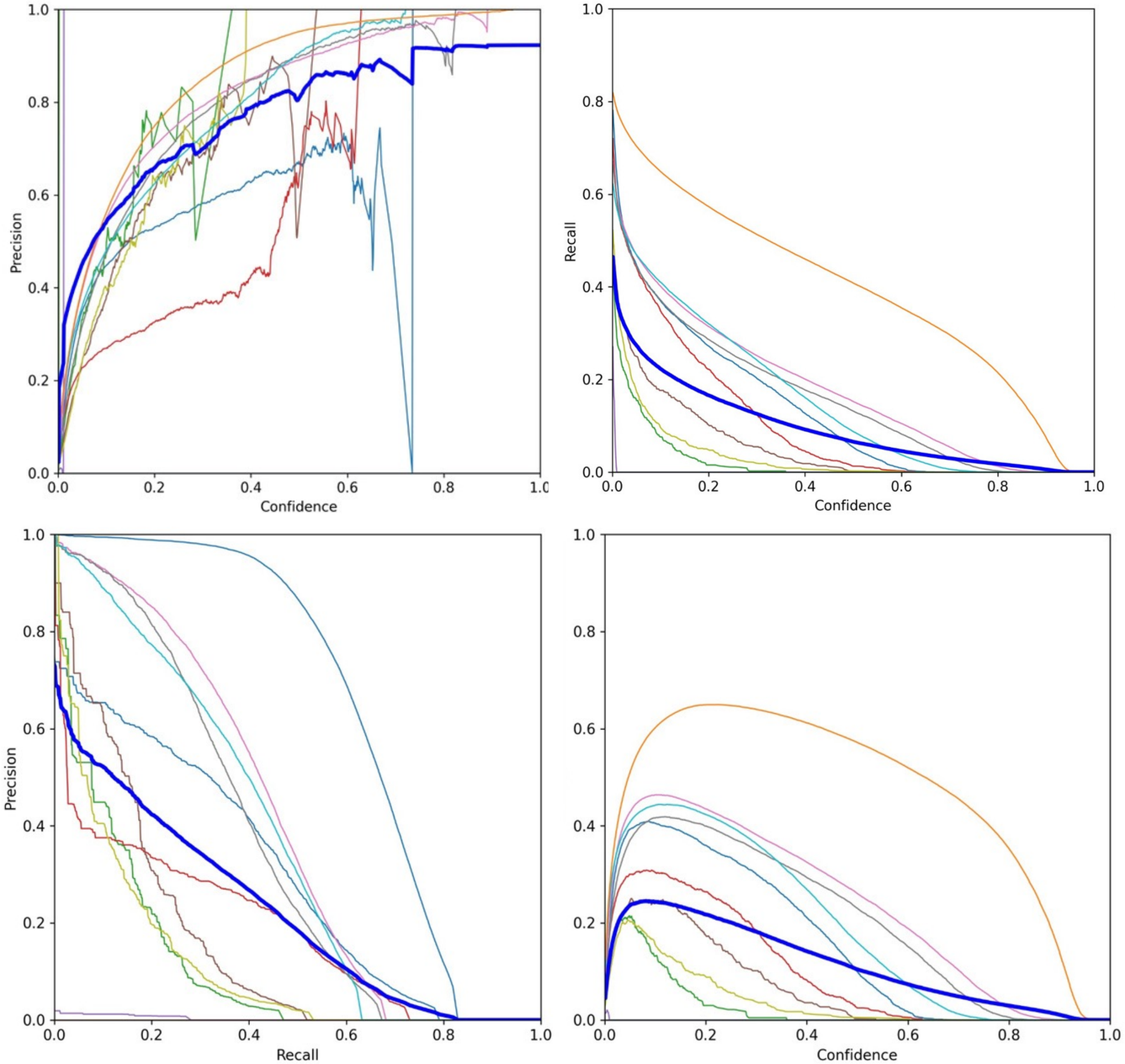
CustomYOLOv7

We use BDD100K to fine tune YOLOv7's detection layer. We chose to use this dataset because we mount the camera inside the cockpit as in the images of BDD.
BDD100K can be used for different tasks: we used it by taking advantage of the object detection labels, which include 10 classes: pedestrian, rider, car, truck, bus, train, motorcycle, bicycle, traffic light, traffic sign.

To do the training, we obtained a partnership with the company LeaderGPU. Which made available to us for 15 hours 5 GPU rtx 3090.

Results


truck
other person
motorcycle
bus
other vehicle
rider
traffic sign
pedestrian
bicycle
traffic light
train
car
trailer
all classes 0.46 at 0.000


Precision
Confidence
Recall
Confidence
Precision
Recall
Confidence

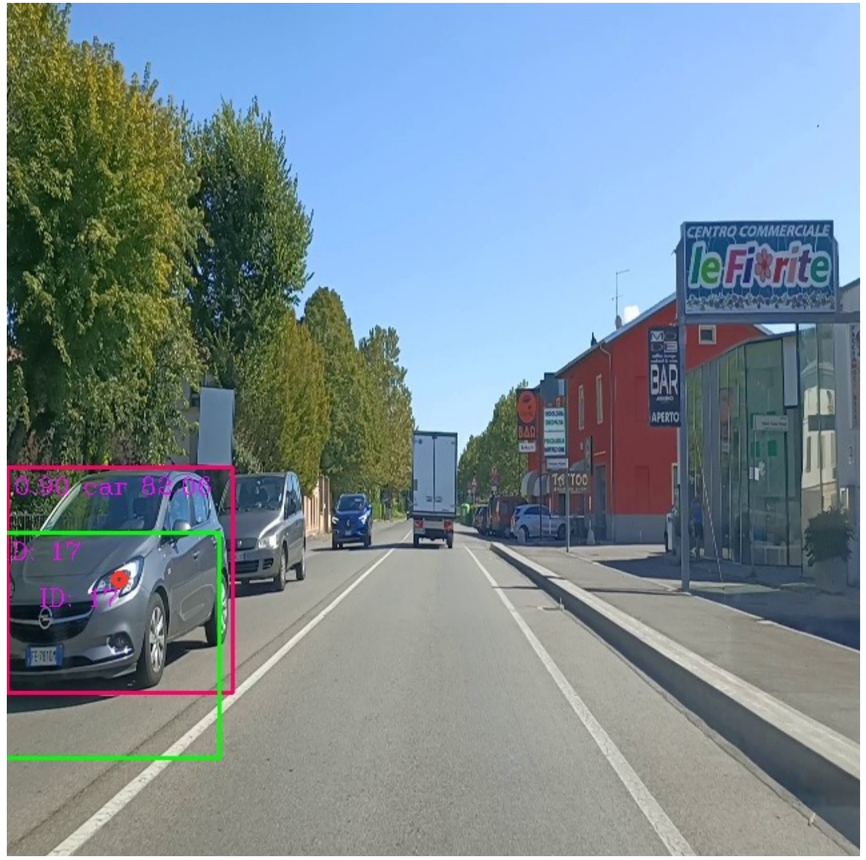
Future Developments

Tracking

Starting with the bounding box which describe where the object is located, we carry out the tracking of it using Kalman filters. These filters tend to produce increasingly accurate estimates of the object's future location as we add information about the trajectory it has traveled so far.

The Kalman filter at work then takes advantage of the current bounding box and the ROI in HSV format to update the model that it exploits to perform the estimation and to subsequently provide us with a prediction of future position.

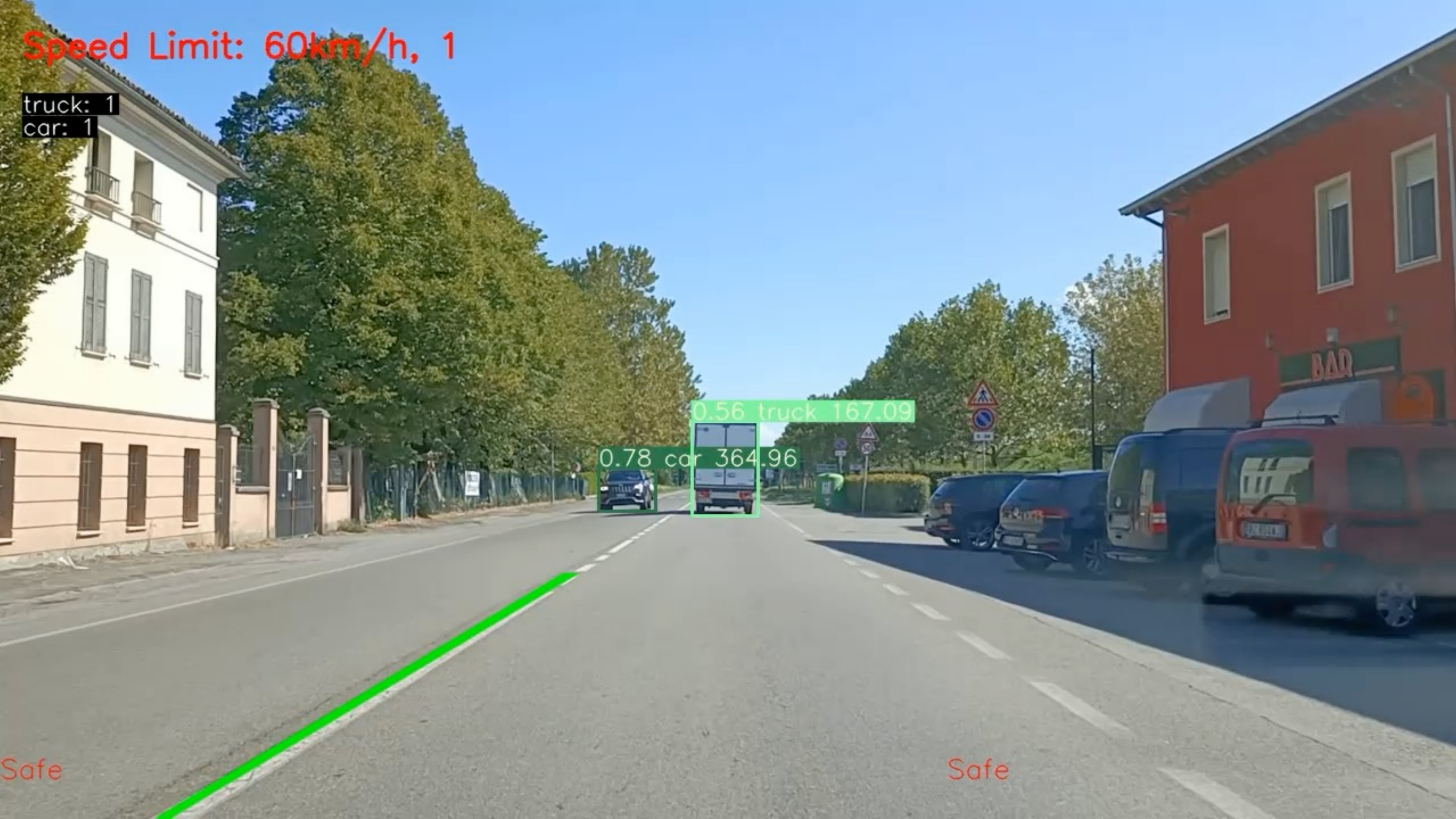
During testing, we noticed that the Kalman Filters produce an accurate prediction of the future position of the object only when the bounding box is large enough.



Object distance

To calculate the distance we use a very simple procedure: we have a reference frame in which we recognize an object, whose size and distance from the camera we know. From this data we are able to extract the focal length, which we then use to calculate the distance to a generic object. This method is rather inaccurate and needs appropriate calibration according to the camera being used.

To improve the distance calculation technique it would be sufficient to use a stereo vision camera, which unfortunately we were unable to add to the project for mainly economic reasons.


Speed Limit: 60km/h, 1
truck: 1
car: 1
0.56 truck 167.09
0.78 car 364.96
Safe
Safe

Some tools used

Partnership with [e-con Systems]

Thank you for listening