Capstone Project Three: Car Detection Model

By: Kris Simino

June 24, 2022

Introduction

Create a system that can:

- Detect the image of a car in a video frame
- Classify the type of car detected
- Display the original image with the car boxed and labeled

Data Acquisition and Wrangling

Two data sets were used. One to experiment and test the detection model and one to train the classification model

Audi Autonomous Driving Dataset (A2D2)

- Publicly available
- 12,499 dashcam images
- JSON format with image and bounding box coordinates

DVM-Car dataset

- Publicly available
- 1.5 million car images
- 899 car models

A2D2 Dataset

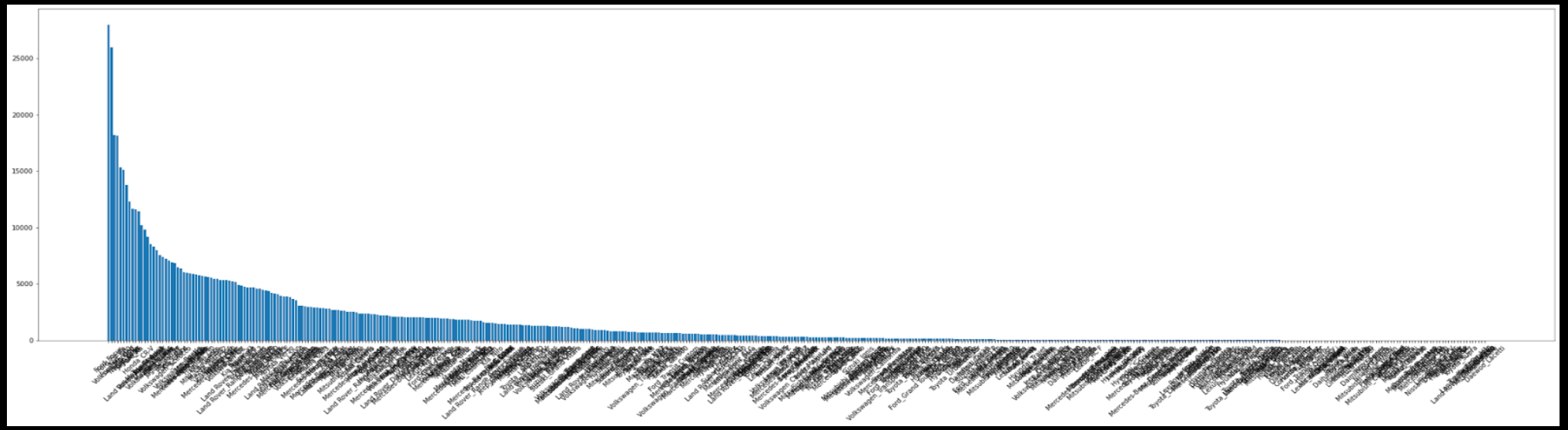
- Not used to train the detection model
- Only used to test the pretrained detection model and final detection/classification system
- Some images contain no labeled vehicles.

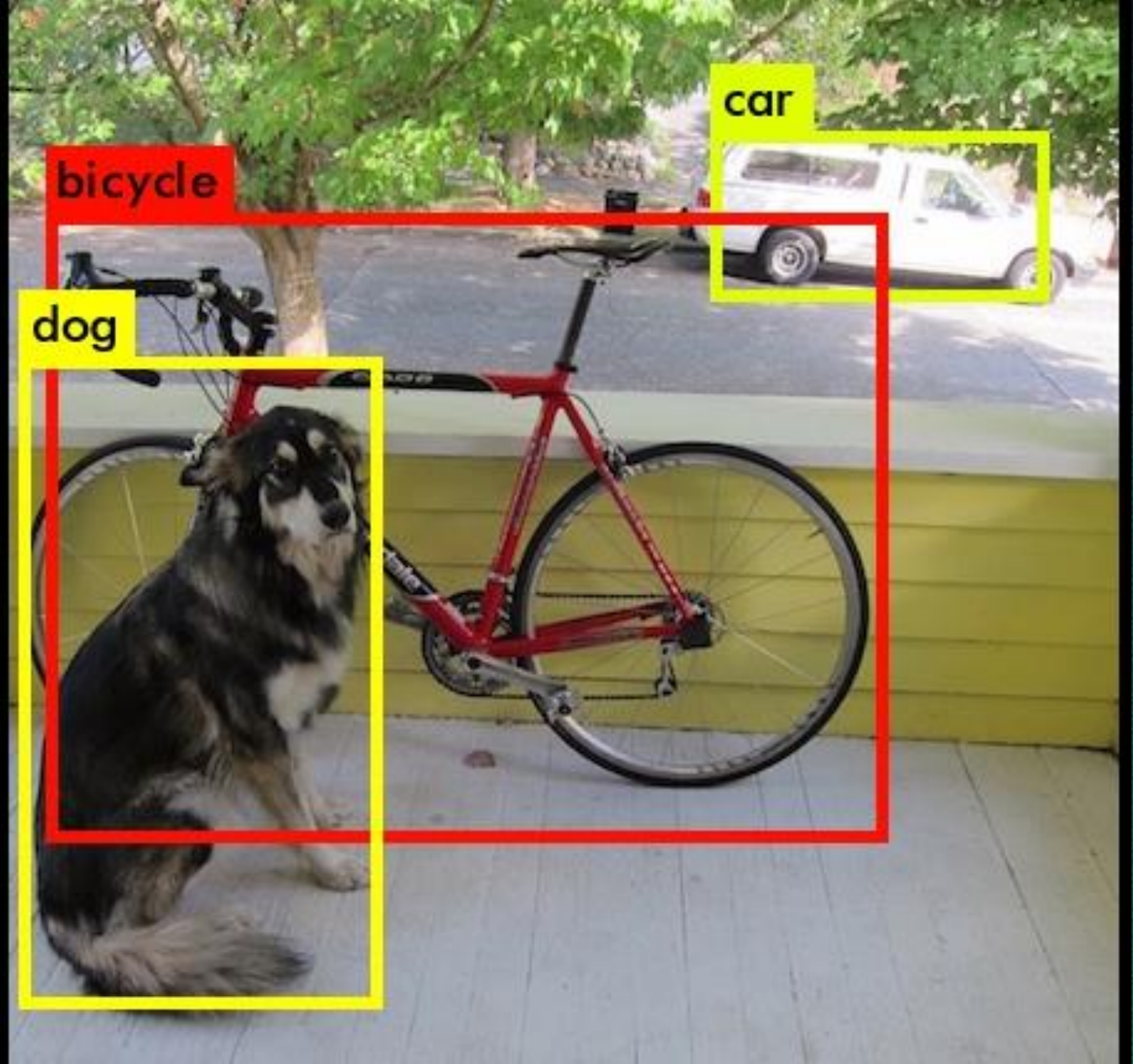
DVM-Car Dataset

- Separated by make model and year
- Images 300 x 300
- Various angles of each car cropped and set into a white background

DVM-Car Dataset Issues

- The dataset of make/model classes is highly skewed
- Largest class (Ford Focus) comprises 3.9%
- Number of classes was reduced twice.
  - Rare and foreign driven vehicles removed to make 456 classes
  - Classes changed to make only with an “unknown” classes added. Reduced to 26.

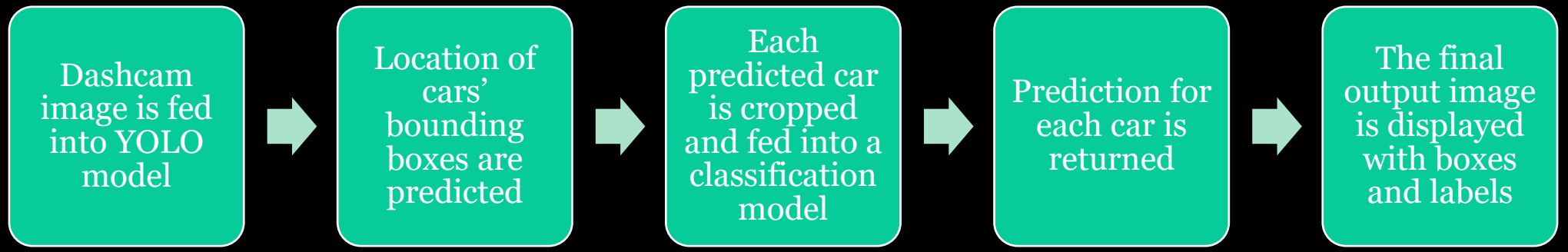


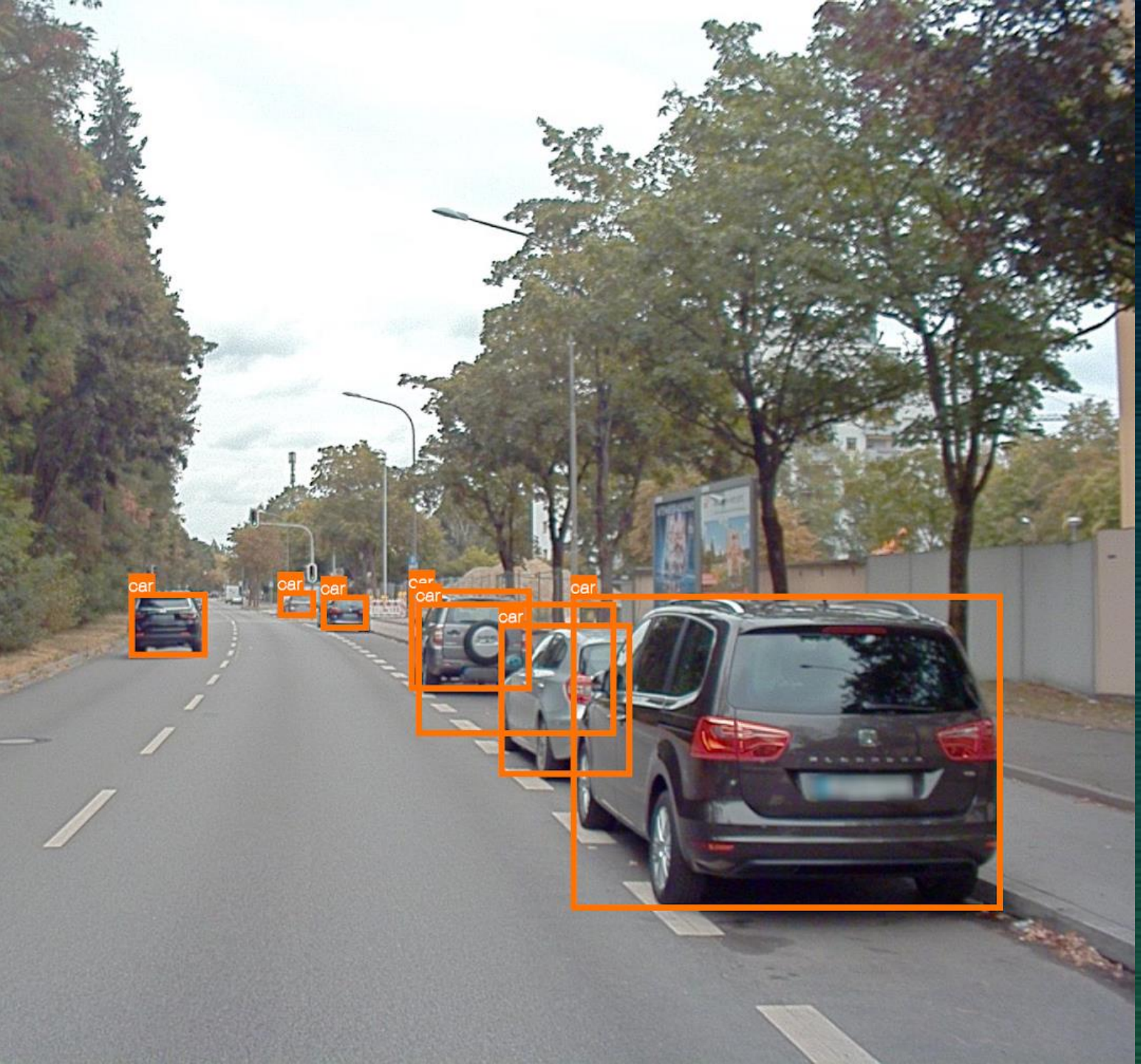


YOLO v 5

- Pretrained weights used
- Detects 80 classes that include, for example, “elephant”, “fire hydrant”, “pizza”, etc.
- Only the classes “car” and “suv” used

Modeling Strategy





Car Detection

- YOLOv5 adjusted to return only “car” and “suv” labeled as “car”
- Detects cars in the background and foreground
- Detects partially obscured cars

Car Detection Performance on A2D2 dataset

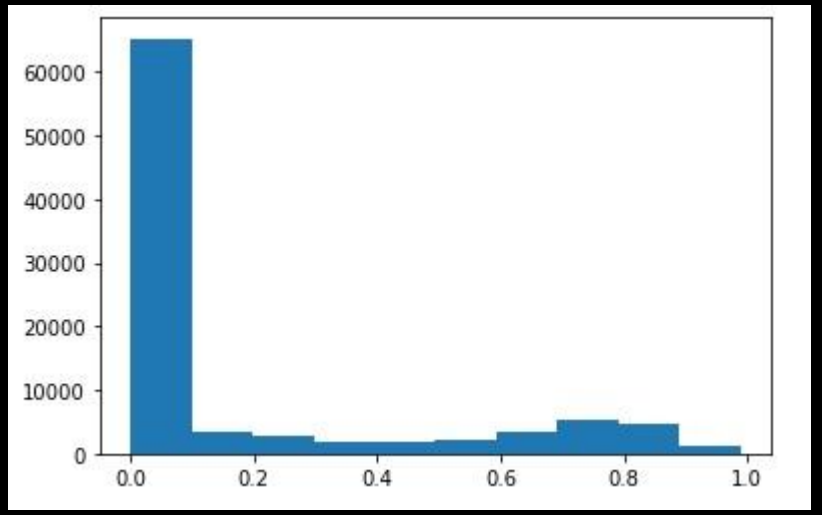
Predicted versus Labeled

- YOLO detects more cars in the image than are labeled.
- A2D2 does not have obscured or very distant cars labeled

	Boxes Known	**Boxes Predicted**
Mean	2.2	5.6
Max	17	29
Min	0	0

IOU Score

- Intersection over Union calculated for matching boxes
- Zero indicates no intersection(no cars)

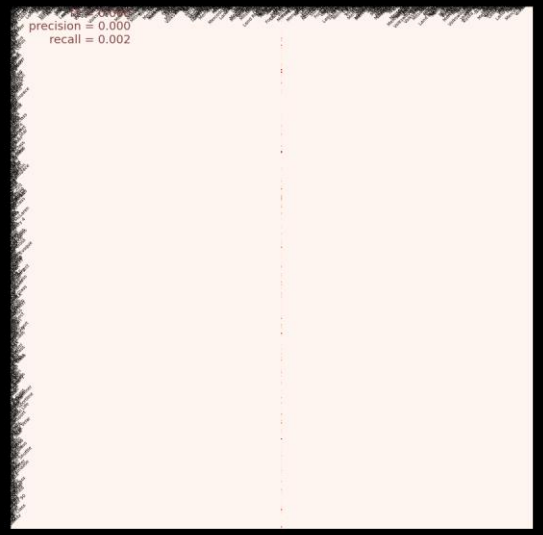


Classification Model

InceptionNet v3 architecture was used with pretrained weights
Output layers trained on DMV-Car dataset
Two separate models trained

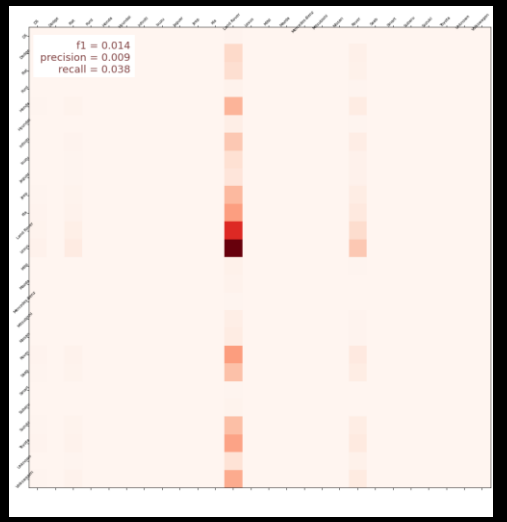
Make and Model

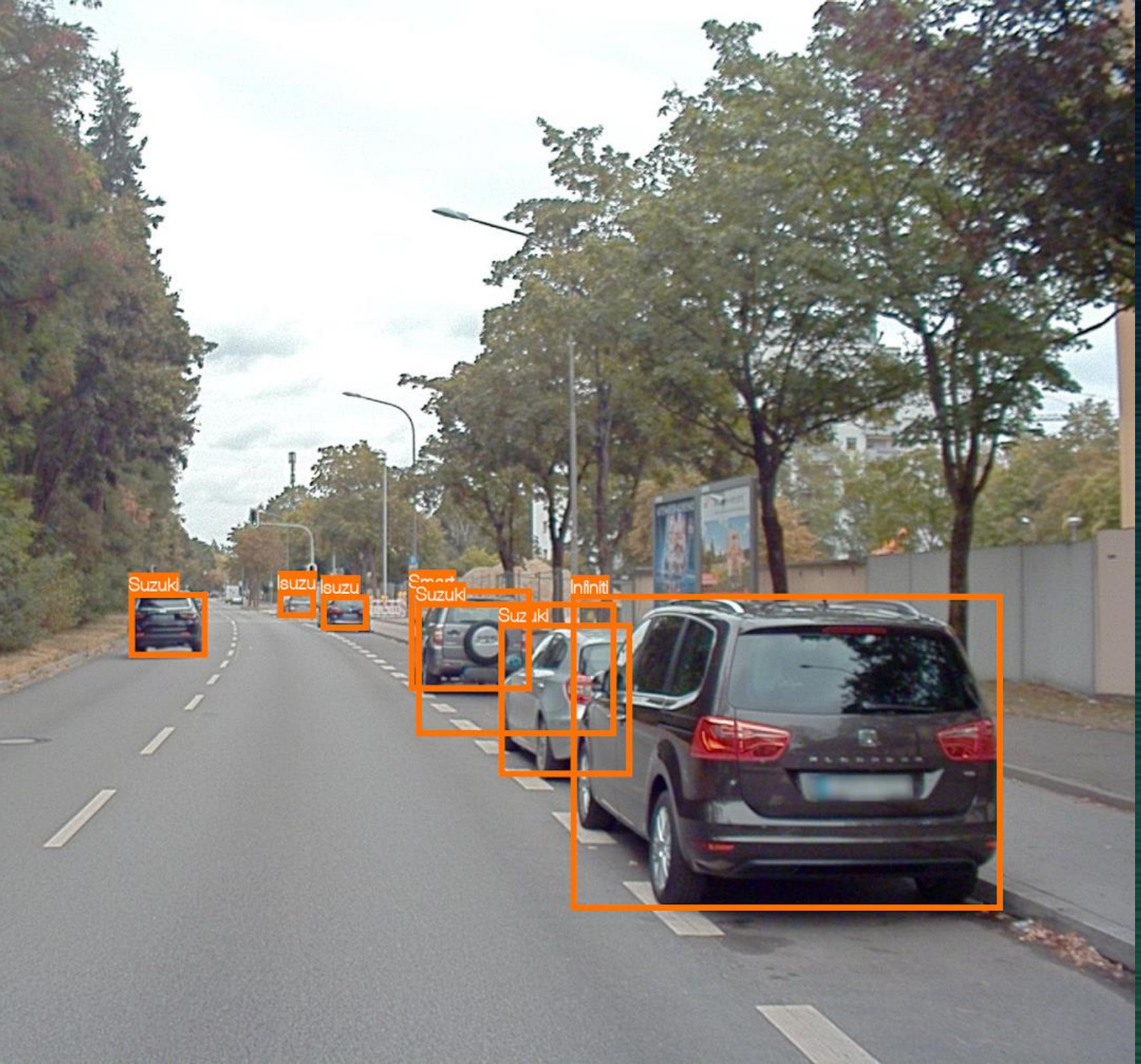
456 classes
Accuracy 0.2%
Precision 0.0%
Nearly all predictions Ford Focus



Make only

26 classes
Accuracy 23%
Precision 0.9%
Nearly all Land Rover and Rover





Putting it together

- Image passes through detection model and bounding boxes are predicted
- Bounding boxes are used to crop the cars
- The car images are resized and put through the classification model
- Car classes are returned
- Final image is returned with bounding boxes and new labels

Recommendations and Future Work

- The system currently should only be used to count vehicles in the image
- The classification phase only has 0.9% precision
- More time should be allotted to increase the performance of the classification phase

Future Work

- Alter the function so the class the YOLO model is searching for can be changed
- Optimize the model to work in real time
- Train other classification models to use
- Package the entire project in a user friendly application