IMAGE TAGGING AND ROAD OBJECT DETECTION

Guide : Prof. Anoop M. Namboodiri

Mentor : Sangeeth Reddy

- Richard Deepak
- Rahul Juluru
- Sridhar Deshpande
- Kshama Pandey

5

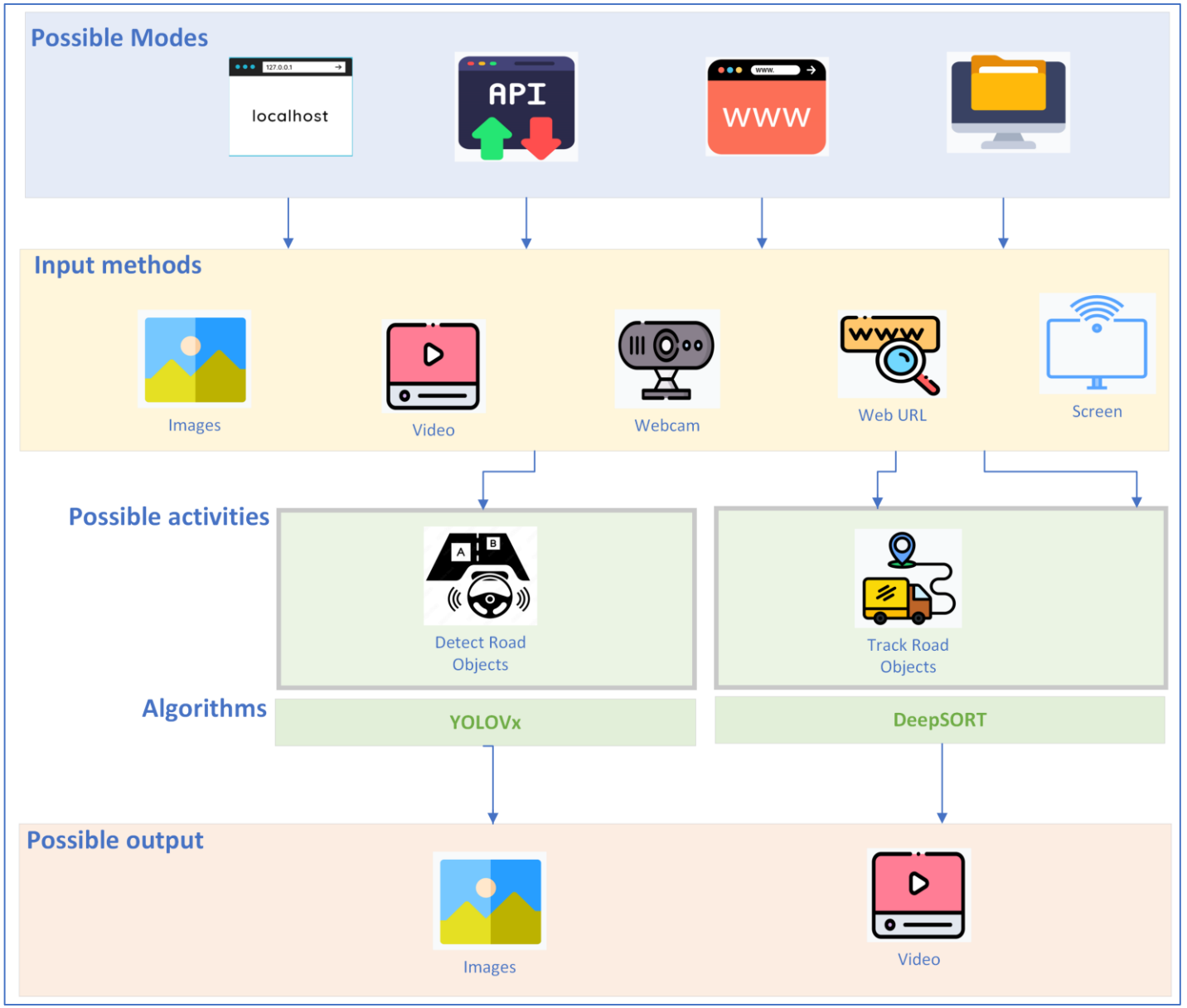
Agenda

- Objective
- Methodology and design considerations
- Outcome in stages with final outcome
- Challenges
- Application Demo

Objective

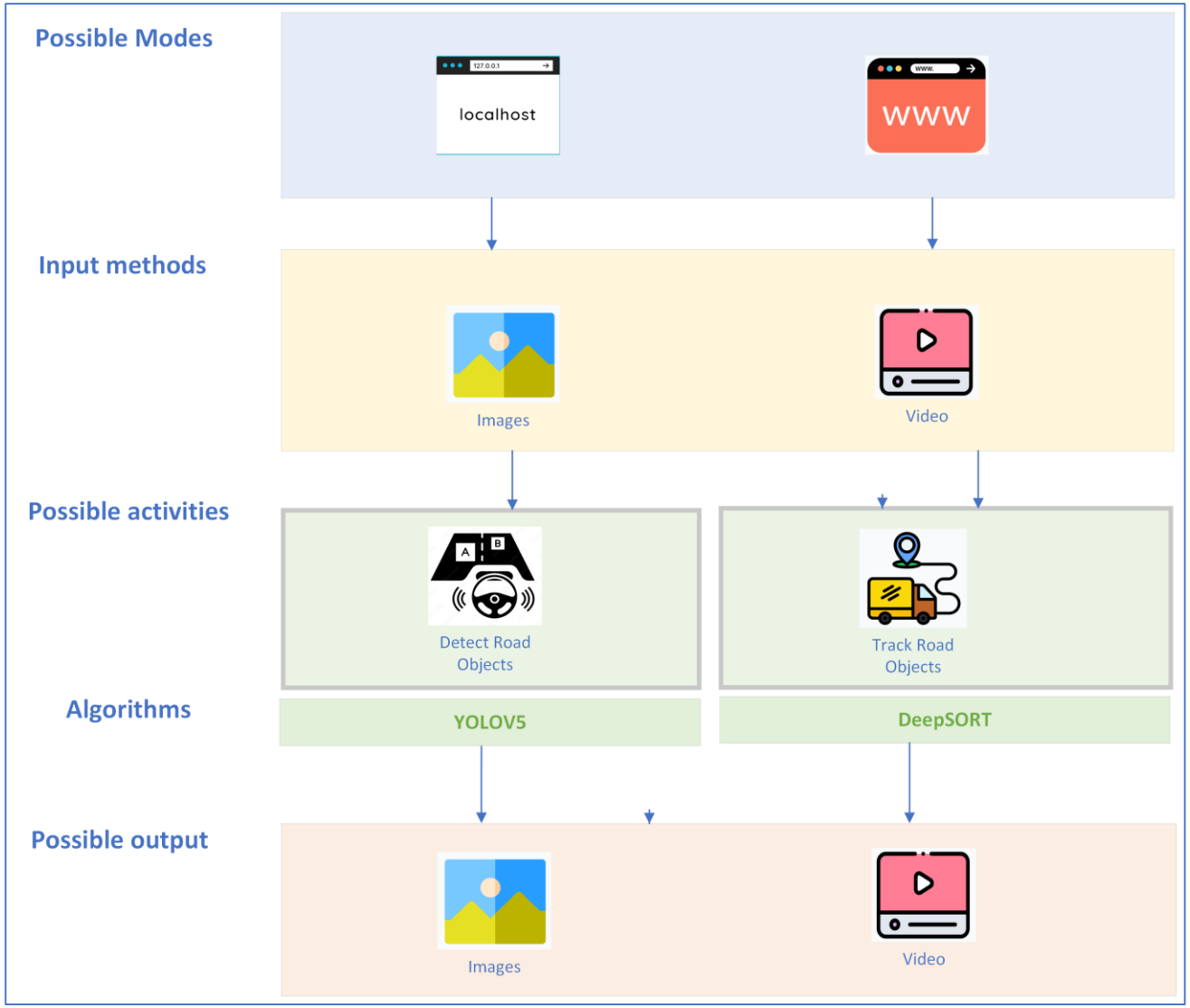
Build an application to detect multiple objects, tag and tracking in a video.

Methodology and design considerations



Methodology and design considerations

Project Scope



Outcome in stages

Understand the data

Learn about BDD100K dataset and download the images from website. Load these images with labels in Google Drive for further processing

Identify the object detection algorithm

Understand the repository with detailed analysis on the way algorithm works and its efficiency. We have chosen YOLOv5 for this purpose and understood steps for download its repository

Identify the tracking algorithm

We have chosen Deep SORT algorithm as our tracking algorithm as it has greater efficiency and support community.

Process images & videos for object detection

Train , validate and test the BDD100K images & videos with YOLOv5 algorithm. Obtain the final model weights generated from the epoch which can be the base for object detection.

Process tracking algorithm for videos

Test the BDD100K video samples downloaded form website to understand the tracking efficiency after downloading their repo and run the required files. Customize the solution to import essential packages / libraries.

Build the streamlit application and deploy the solution

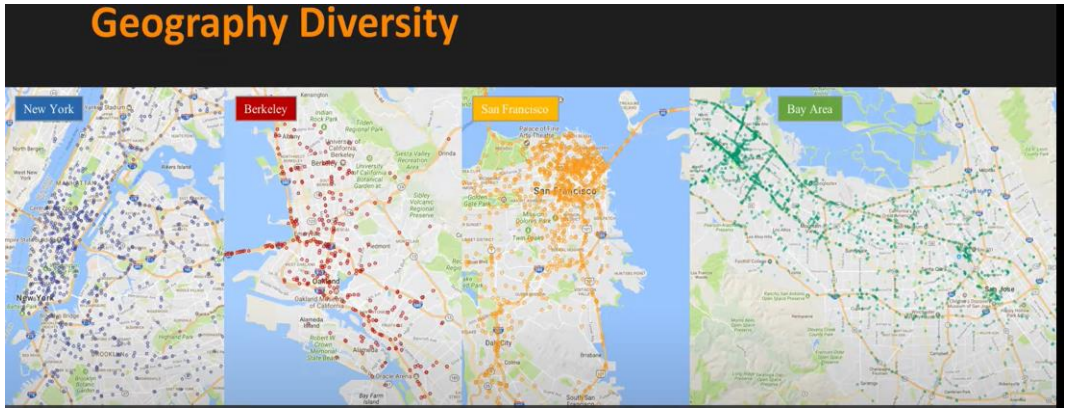
Develop the streamlit application to use the model weights for object detection and tracking. Deploy the entire solution to Streamlit Cloud after pushing entire solution to GIT repository.

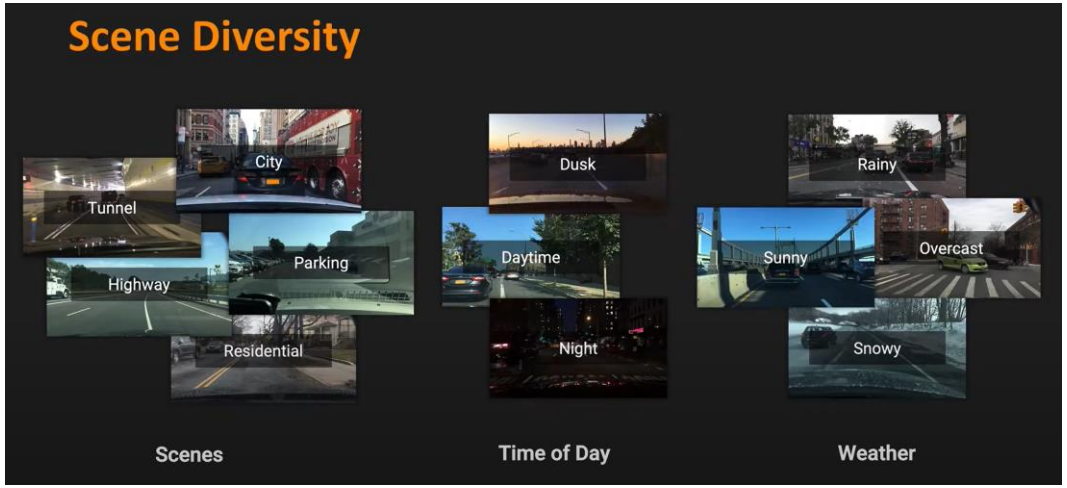
Outcome in stages

Stage 1 – Understand the data

BDD100K Dataset

- Consists of 100,000 videos.
  Each video is about 40 seconds long, 720p, and 30 fps
- Geographically diverse
- Scene Diversity : Scene = City , Tunnel , Highway , Parking , Residential
- Time of the day = Dusk , Daytime , night
- Weather = Rainy , Sunny , Overcast , Snowy
- There are >10 objects per image with pixel annotation and tracking





Outcome in stages

Stage 1 – Understand the data

Original Image

BDD100k dataset bounding box JSON format

dd100k_labe...ges_val.json*

```
[
    {
        "name": "000d4f89-3bcbe37a.jpg",
        "attributes": {
            "weather": "overcast",
            "scene": "city street",
            "timeofday": "daytime"
        },
        "timestamp": 10000,
        "labels": [
            {
                "category": "traffic sign",
                "attributes": {
                    "occluded": false,
                    "truncated": false,
                    "trafficLightColor": "none"
                },
                "manualShape": true,
                "manualAttributes": true,
                "box2d": {
                    "x1": 1000.698742,
                    "y1": 281.992415,
                    "x2": 1040.626872,
                    "y2": 326.91156
                },
                "id": 0
            },
            {
                "category": "traffic sign",
                "attributes": {
```

Image name

One object complete details in an image

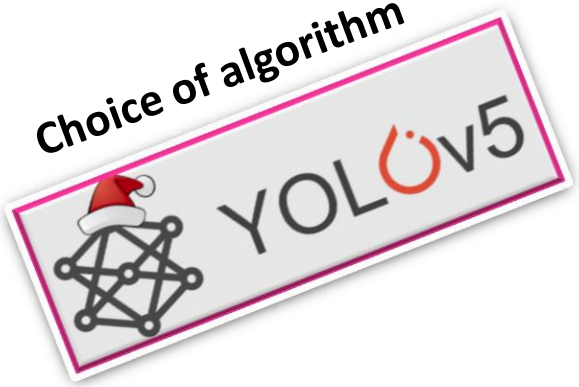
Object class

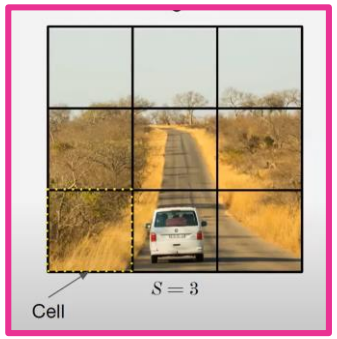
Bounding box dimensions

Outcome in stages

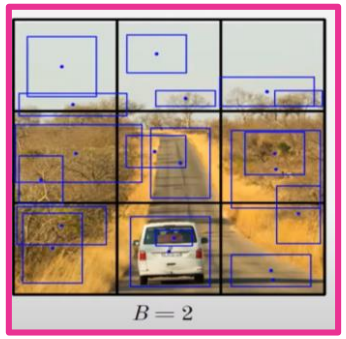
Stage 2 – Identify the object detection algorithm

Choice of algorithm

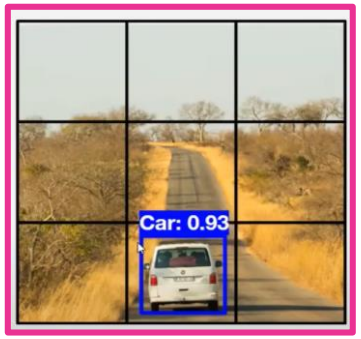




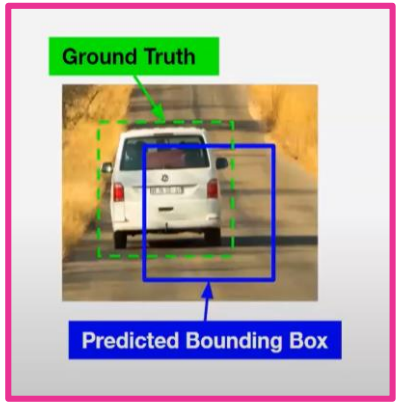
Divide the image to cells



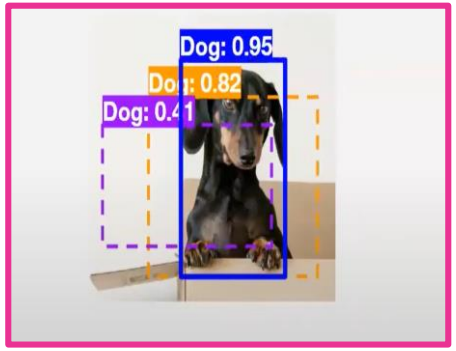
Predicting bounding boxes within each cell



Selection of bounding box which is above confidence threshold



Intersection over Union



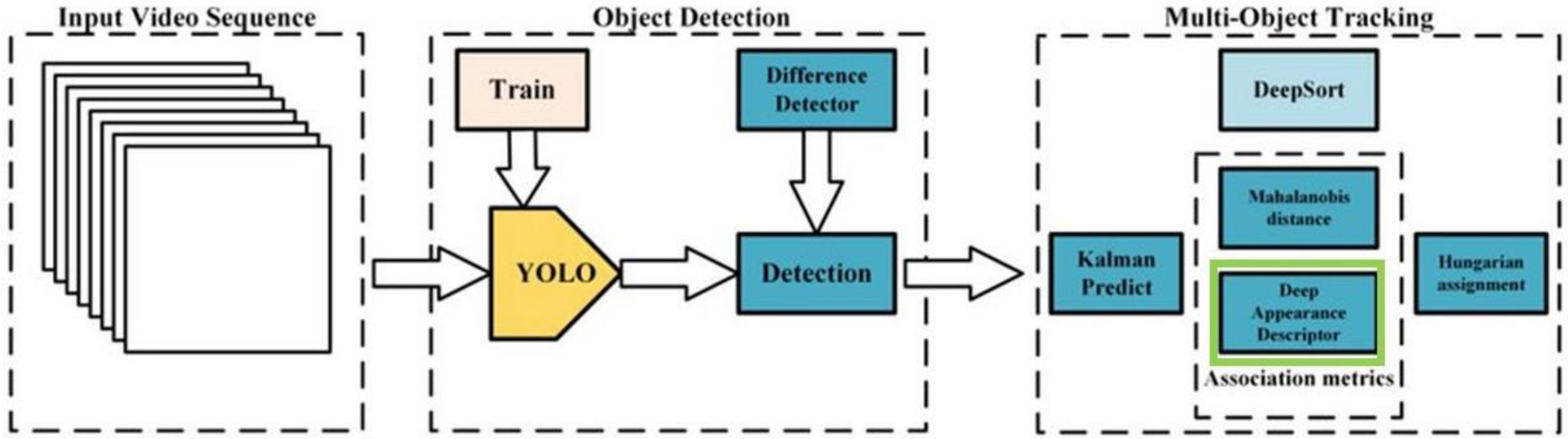
Non max suppression

Outcome in stages

Stage 3 – Identify the tracking algorithm

Choice of algorithm

DeepSORT



Outcome in stages

Stage 4 – Process images and videos for object detection

Original Image

```
bdd100k_labe...ges_val.json"
[
    {
        "name": "000d4f89-3bcbe37a.jpg",
        "attributes": {
            "weather": "overcast",
            "scene": "city street",
            "timeofday": "dadaytime"
        },
        "timestamp": 10000,
        "labels": [
            {
                "category": "traffic sign",
                "attributes": {
                    "occluded": false,
                    "truncated": false,
                    "trafficLightColor": "none"
                },
                "manualShape": true,
                "manualAttributes": true,
                "box2d": {
                    "x1": 1000.698742,
                    "y1": 281.992415,
                    "x2": 1040.626872,
                    "y2": 326.91156
                },
                "id": 0
            },
            {
                "category": "traffic sign",
                "attributes": {
```

Image name

One object complete details in an image

Object class

Bounding box dimensions

BDD100k dataset bounding box JSON format

000d4f89-3bcbe37a - Notepad

File Edit Format View Help

bounding box x1 ,y1 ,x2 ,y2 normalized

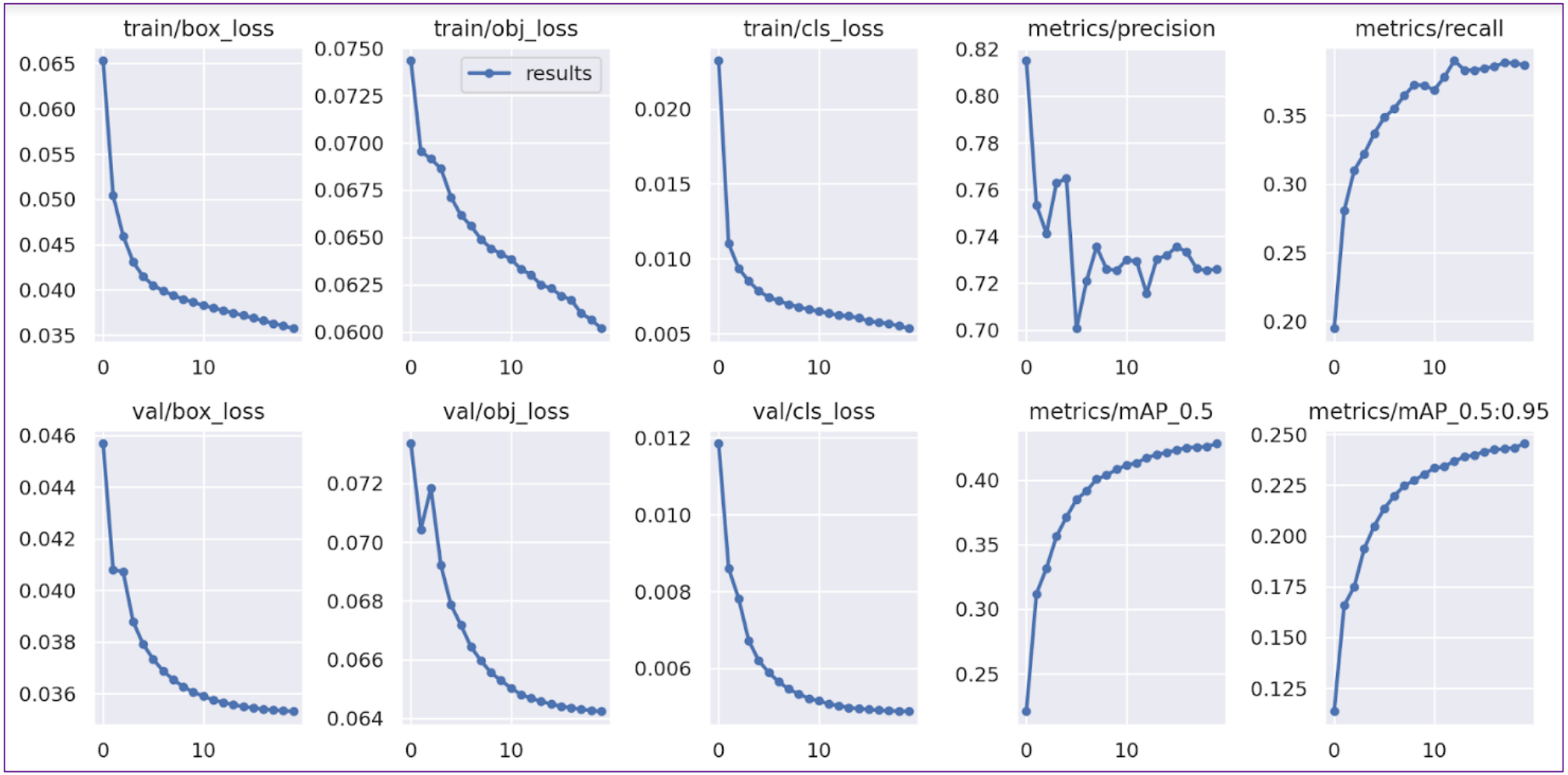
```
2 0.32104250507812504 0.49959173124999995 0.03926019609374998 0.0440816236111111
2 0.2588552611822023 0.4901953158935303 0.062384705760595335 0.05467654567594956
2 0.23528997304687502 0.49408152638888886 0.06612243515625002 0.06244896666666667
2 0.1738167703125 0.48760738138567267 0.1456759875 0.07154148499356758
2 0.058102510156249995 0.4720407166666667 0.03202805625 0.025714280555555608
2 0.015361844921875001 0.4674488784722222 0.030723689843750002 0.060612231944444435
2 0.46025830133564527 0.5259084653235868 0.05424690826620946 0.05630693620272922
```

Class Label

Converted to TXT file with required class and Bbox with mapping data and labels

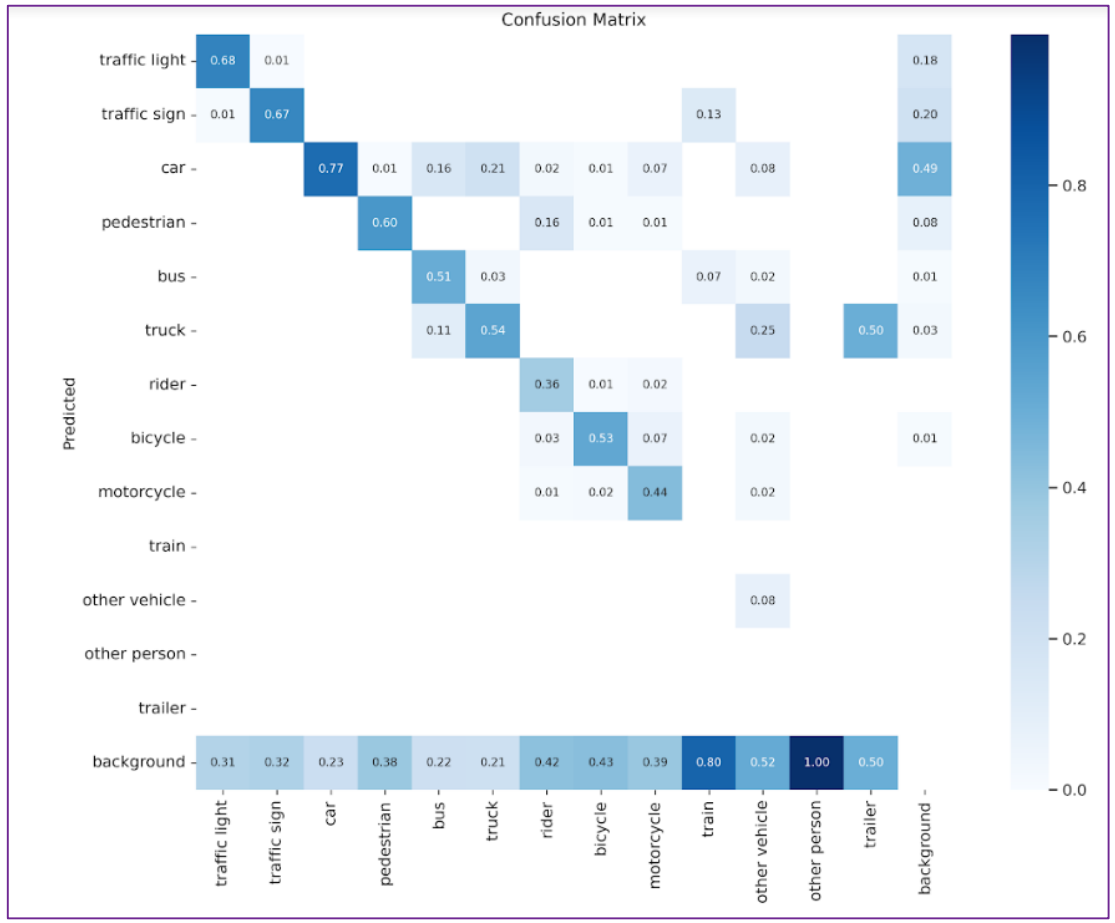
Actual Outcomes

Preliminary results



Actual Outcomes

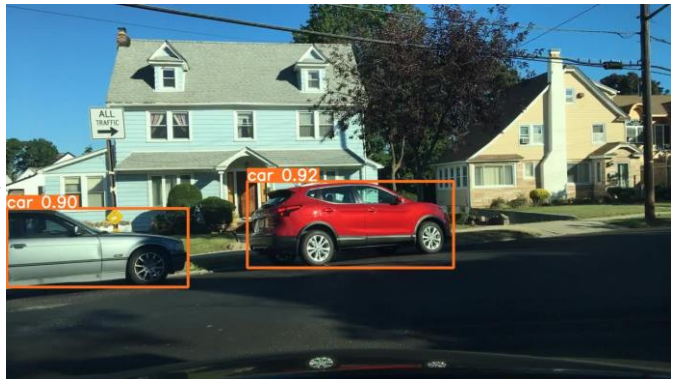
Preliminary results – Confusion Matrix

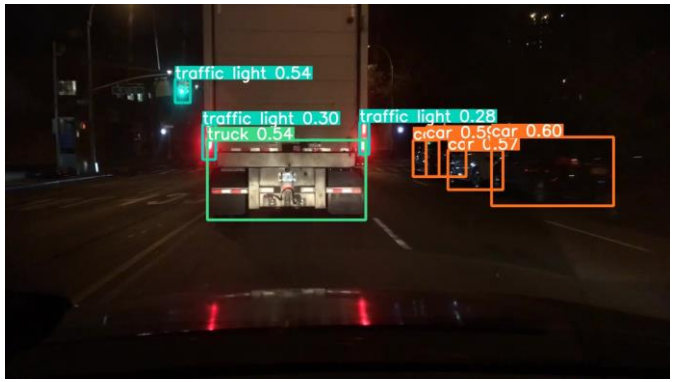


Actual Outcomes

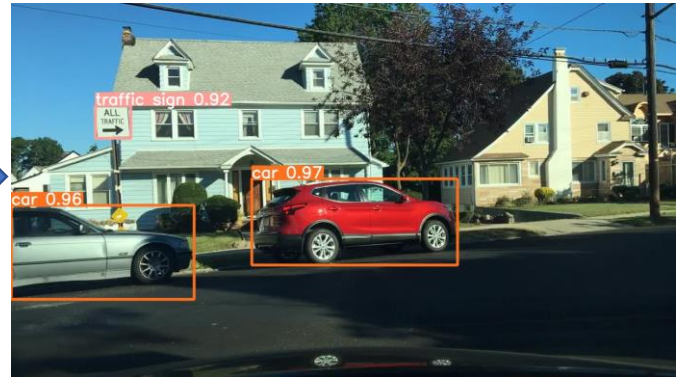
YOLOv5 Object Detection

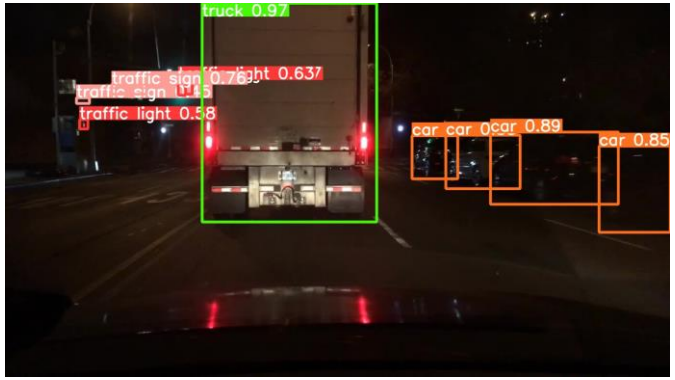
Pre-trained COCO dataset





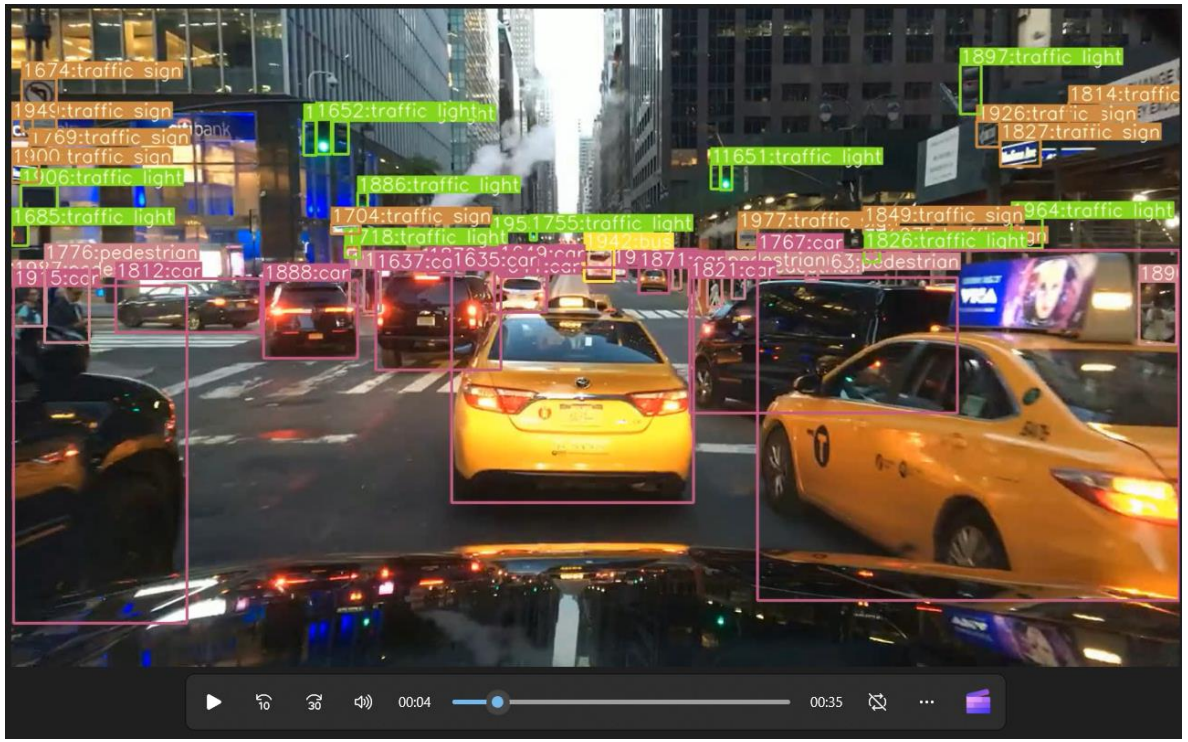
After BDD100K dataset training





Outcome in stages

Stage 5 – Process videos for object tracking



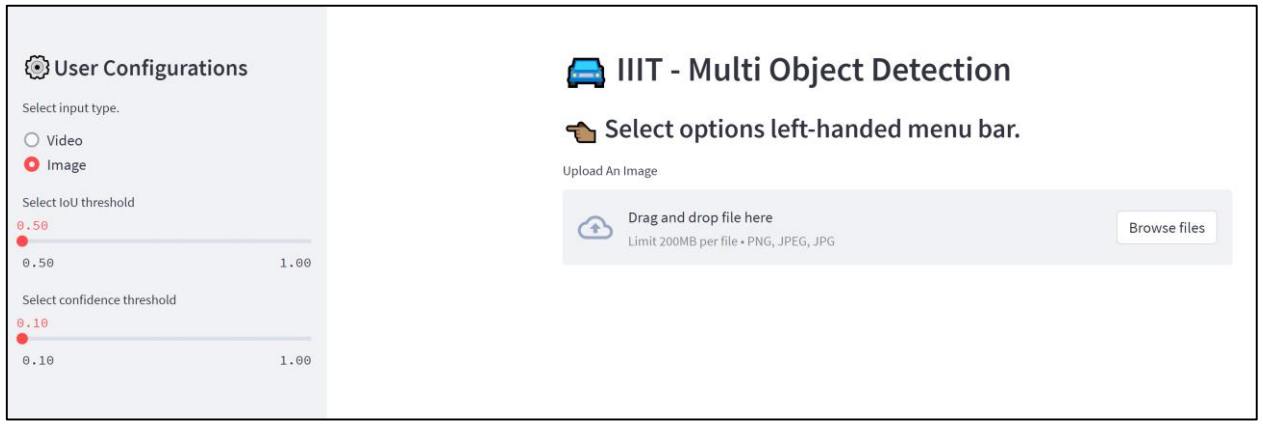
Actual Outcomes

YOLOv5 + Deep SORT tracker output

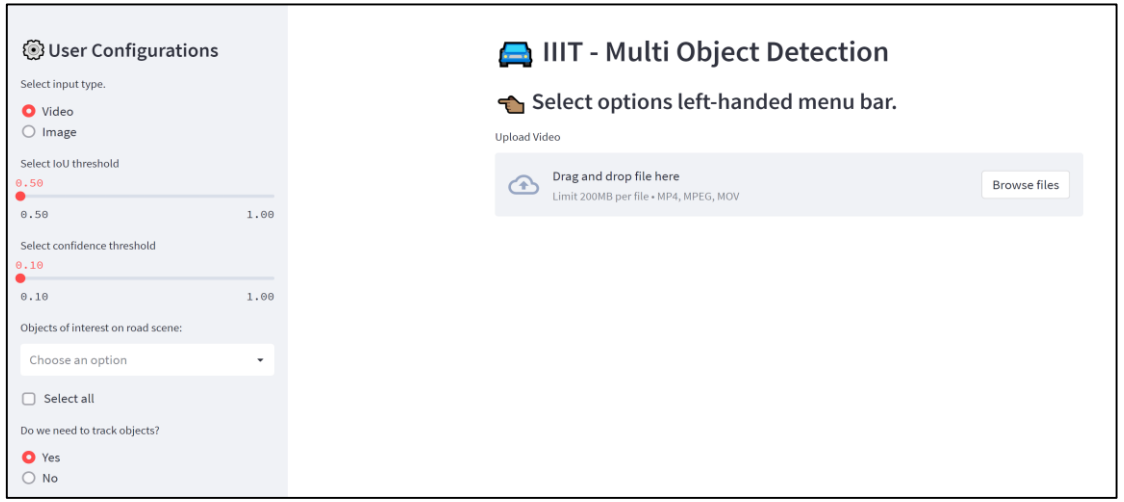
Tracked GIF via YOLOv5 + DeepSORT

Outcome in stages

Stage 6 – Develop streamlit application and deploy onto Streamlit Cloud



Application for Image detection



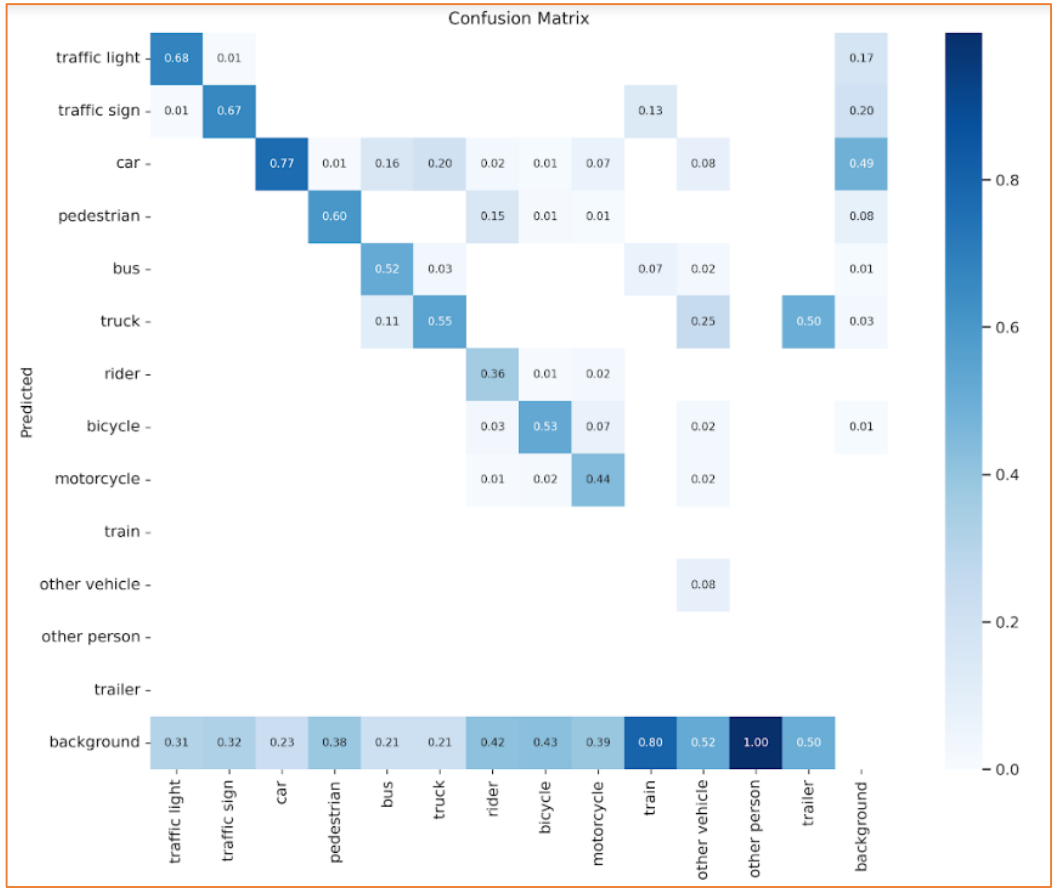
Application for Video detection and tracking

Challenges

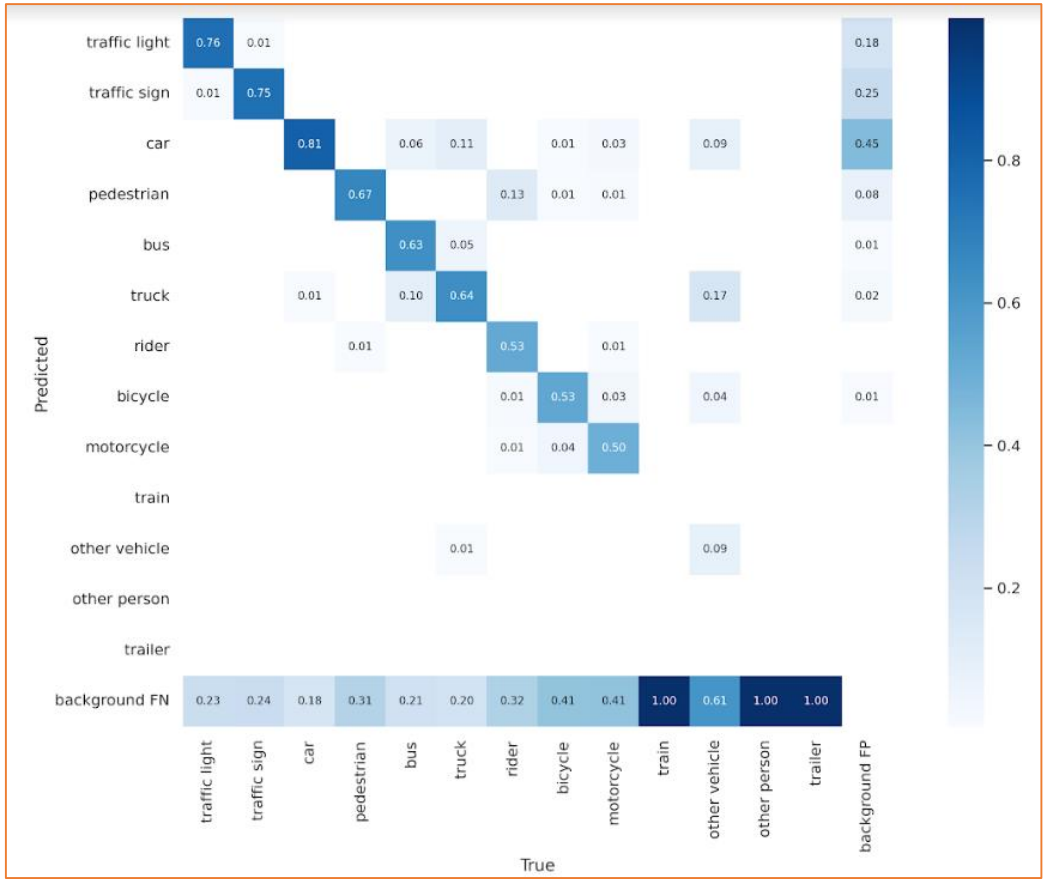
1. Missing labels for some of the images of BDD100k data.
2. Downloading of BDD100k videos were time consuming.
3. Resource constraint – GPU causing disconnection of runtime in colab.
4. Sharing GIT Code repositories
5. Understanding cloud constraints
6. Understanding installation of packages in streamlit cloud.
7. How to store the models so as to be loaded in the cloud environment.
8. Understanding Codec formats compatible with browsers.
9. Resource constraints while testing the web-UI locally.

Stretch Goals

YOLOv5 vs YOLOv7



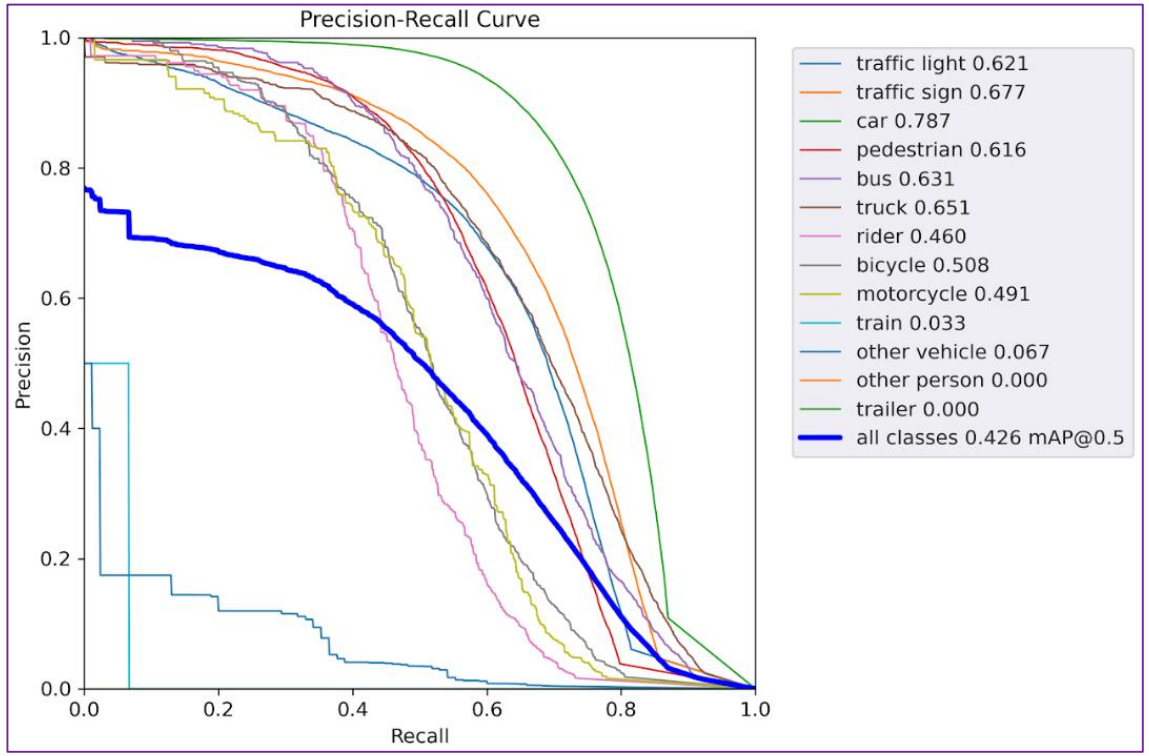
V5



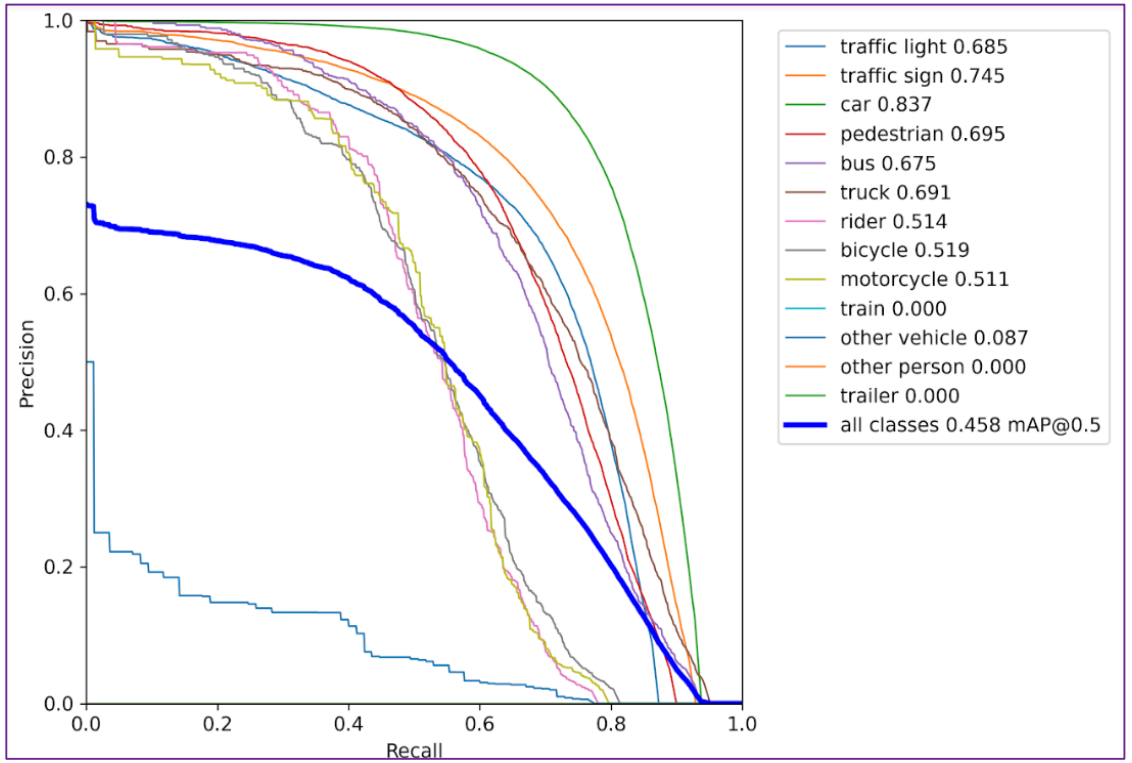
V7

Stretch Goals

YOLOv5 vs YOLOv7



V5



V7

Stretch Goals

YOLOv5 vs YOLOv7

Class	Images	Instances	P	R	mAP50	mAP50-95:
all	10000	186033	0.727	0.387	0.426	0.244
traffic light	10000	26884	0.676	0.601	0.621	0.234
traffic sign	10000	34724	0.728	0.623	0.677	0.36
car	10000	102837	0.796	0.722	0.787	0.505
pedestrian	10000	13425	0.719	0.553	0.616	0.307
bus	10000	1660	0.698	0.551	0.631	0.488
truck	10000	4243	0.691	0.594	0.651	0.477
rider	10000	658	0.733	0.388	0.46	0.231
bicycle	10000	1039	0.601	0.477	0.508	0.247
motorcycle	10000	460	0.667	0.446	0.491	0.244
train	10000	15	1	0	0.0326	0.0261
other vehicle	10000	85	0.143	0.0706	0.0672	0.0507
other person	10000	1	1	0	0	0
trailer	10000	2	1	0	0	0

V5

Class	Images	Labels	P	R	mAP@.5	mAP@.5:.95:
all	10000	186033	0.742	0.422	0.458	0.258
traffic light	10000	26884	0.708	0.668	0.685	0.258
traffic sign	10000	34724	0.727	0.701	0.745	0.396
car	10000	102837	0.833	0.76	0.837	0.522
pedestrian	10000	13425	0.756	0.617	0.695	0.344
bus	10000	1660	0.752	0.586	0.675	0.517
truck	10000	4243	0.704	0.63	0.691	0.505
rider	10000	658	0.611	0.491	0.514	0.257
bicycle	10000	1039	0.611	0.497	0.519	0.249
motorcycle	10000	460	0.73	0.457	0.511	0.251
train	10000	15	1	0	0	0
other vehicle	10000	85	0.218	0.0824	0.0873	0.0519
other person	10000	1	1	0	0	0
trailer	10000	2	1	0	0	0

V7

Stretch Goals

YOLOv5 vs YOLOv7

V5

V7

Application Demo

Thank You