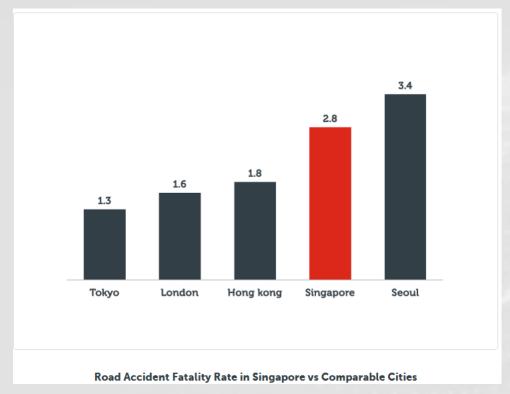


"Total number of traffic accidents resulting in **injuries or fatalities increased by 4.9 %,** from 3,012 in the first half of 2021 to 3,159 in the same period this year."

"The total number of **injured people** and fatalities also **increased by 3.3 %,** from 3,669 in the first six months of 2021 to 3,791 in the same period in 2022.

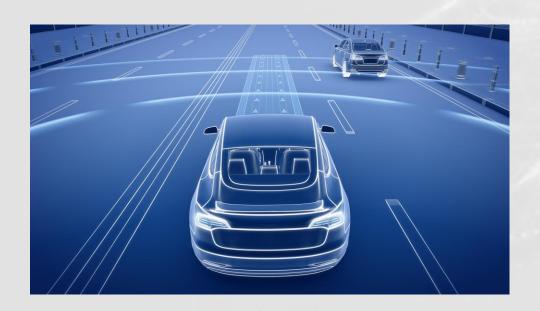
~By CNA(8<sup>TH</sup> Sept 2022)

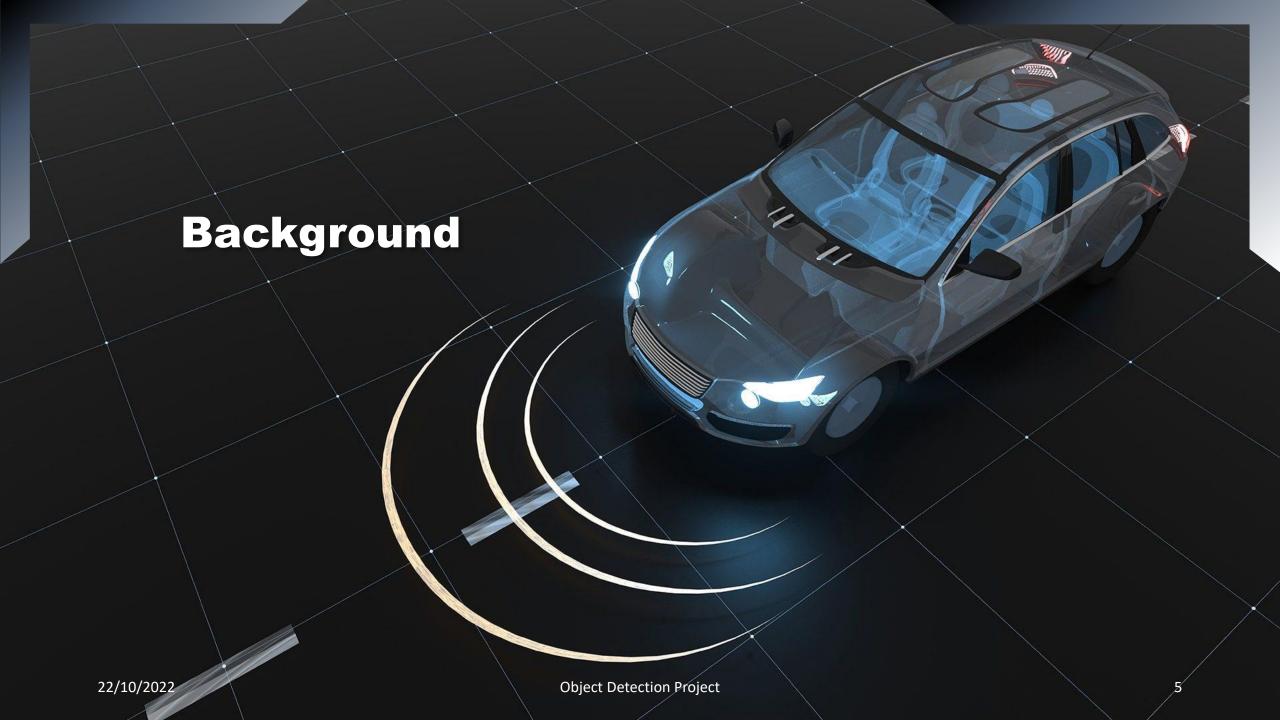


Road accident statistics in Singapore 2022 (budgetdirect.com.sq)

How can we use technology to help us to solve the problem? **Automated Vehicle** (**AV**) can help us to reduce the number of vehicles on the road and reduce the number of accidents through eliminating human errors. However, One of the main challenges in AV is to detect the road objects accurately to plan the driving path safely.

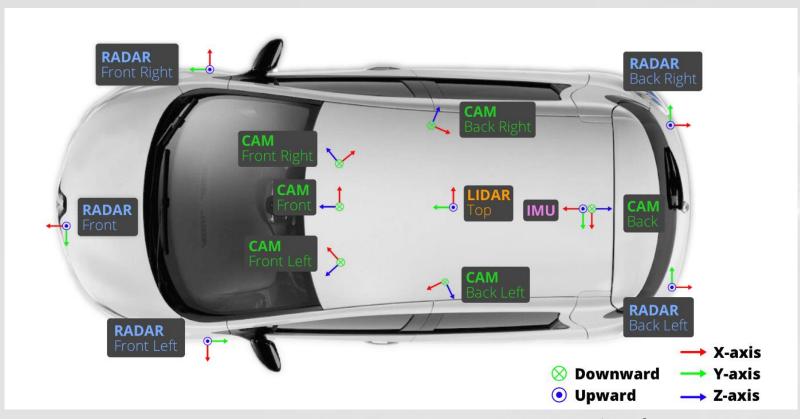
This project aim is to **develop a road object detection model** to help AV to detect road objects using Computer Vision (CV). It will also be trained to detect specific road objects like road barriers and traffic cone to help AV to drive safely.



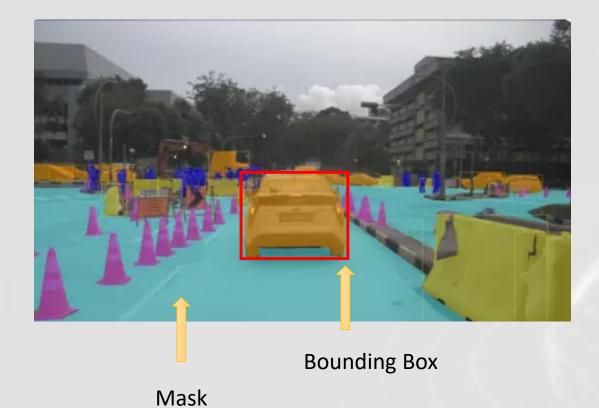


There are 3 types of object detection method in modern AV

- A. Lidar light (Object location and 3D shape)
- B . Radar radio wave (Distance and location)
- **C**. Camera human vision (Object type and segmentation)

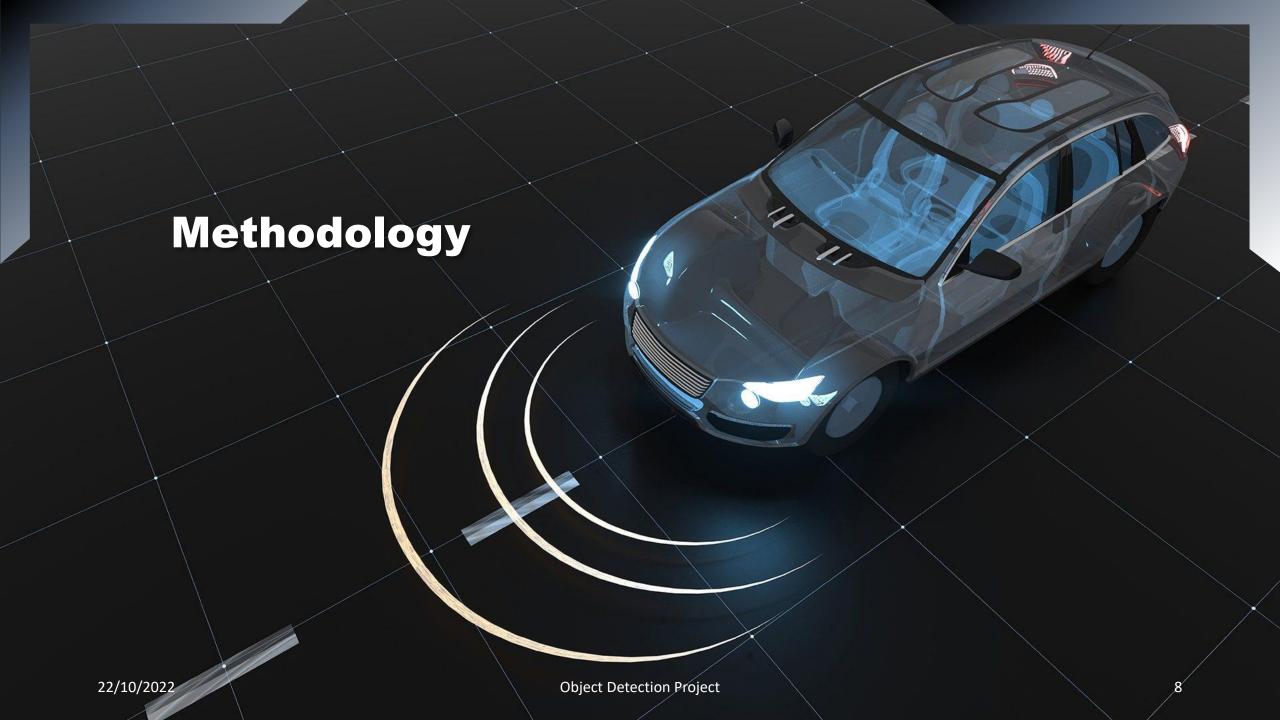


Pic taken from : nuScenes



#### **Object Detection and Segmentation**

- Driving route planning
- Barrier detection
- Locate moving object distance and size
- Dangerous object alarm system



#### **Data Science Work Flow**

- Data Preparation
- EDA
- Data Pre-processing
- Data sampling
- Model Training
- Parameter Tuning
- Result Evaluation

• Demo and deployment

#### **Data Preparation**

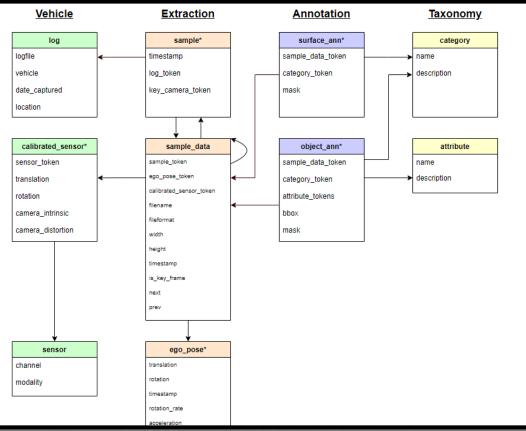
NUSCENES by Motional https://www.nuscenes.org/

Public large-scale dataset for autonomous driving, annotated images.

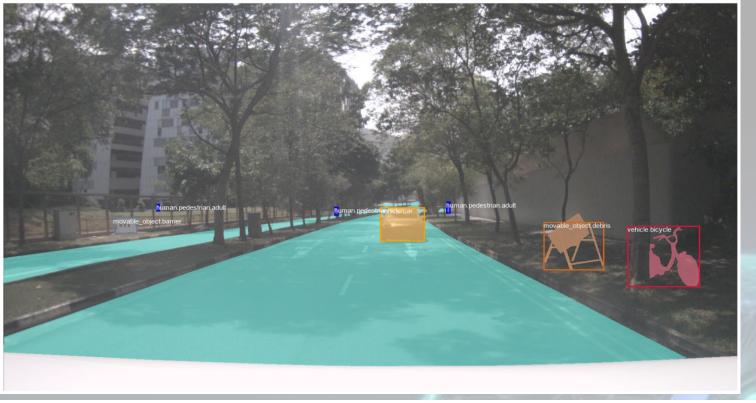
- 1500h of driving data from multiple cities including Singapore.
- 93,000 images with instance mask and 2d boxes with 800k foreground objects.

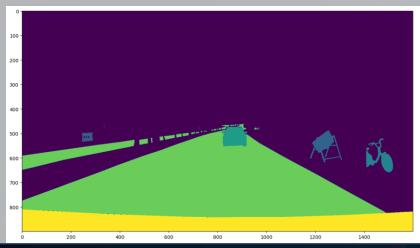
#### Data Extraction

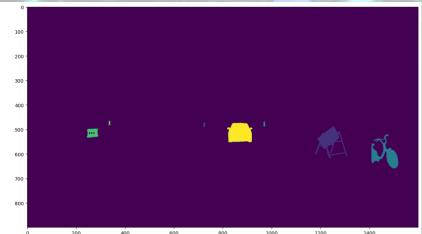
- > Explore multiple data files
- Decipher encoded information like mask into useful data
- ➤ Merge the folders and customize them into standard COCO annotation format



EDA







#### **Pre-processing**

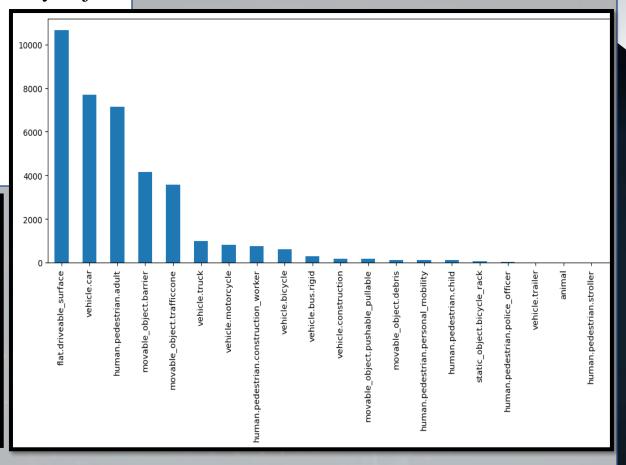
- 1. Select images from Singapore road
- 2. Combine annotation files and key files
- 3. Select 10% of overall dataset (Downsize sample)
- 4. Decode and creating bounding box for every object
- 5. Select top 8 classes.
- 6. Final dataset:

Number of images: 4801

Number of objects: 35774

7. Train test split (0.8,0.2)

#### COCO Segmentation format:



#### **Data Imbalance**

Class	Weight
Drivable Surface	29.8%
Vehicle Car	21.5%
Pedestrian	20%
Barrier	11.6%
Traffic Cone	10.0%
Truck	2.8%
Motorcycle	2.3%
Construction Worker	2.1%



#### **Modelling – Transfer Learning**



Open-source object detection API built with Pytorch

Model used: Mask RCNN (Resnet50) FPN

Image Input

**FPN** Res 2 Res 3 Res 4 Res 5 **Backbone Network** 

This backbone network is trainable by freezing different Resnet layers.

Region Proposal Network(RPN)

ROI Pooler

**Bounding Box ROI aligner layers** 

**Mask ROI Aligner layers** 

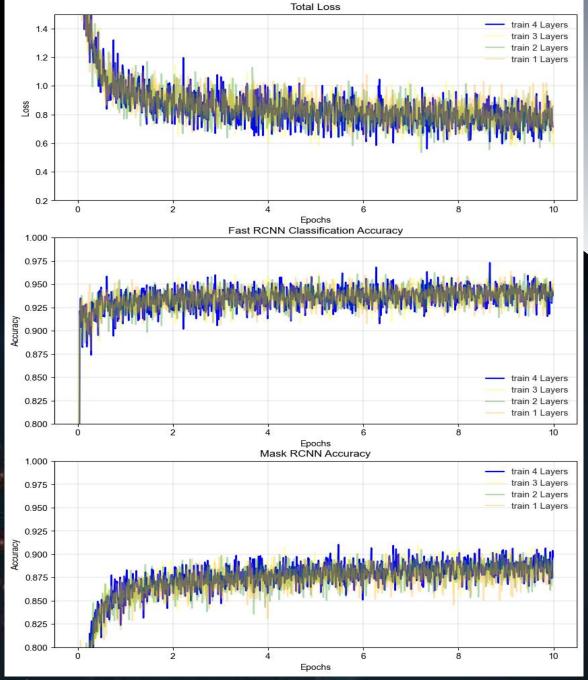


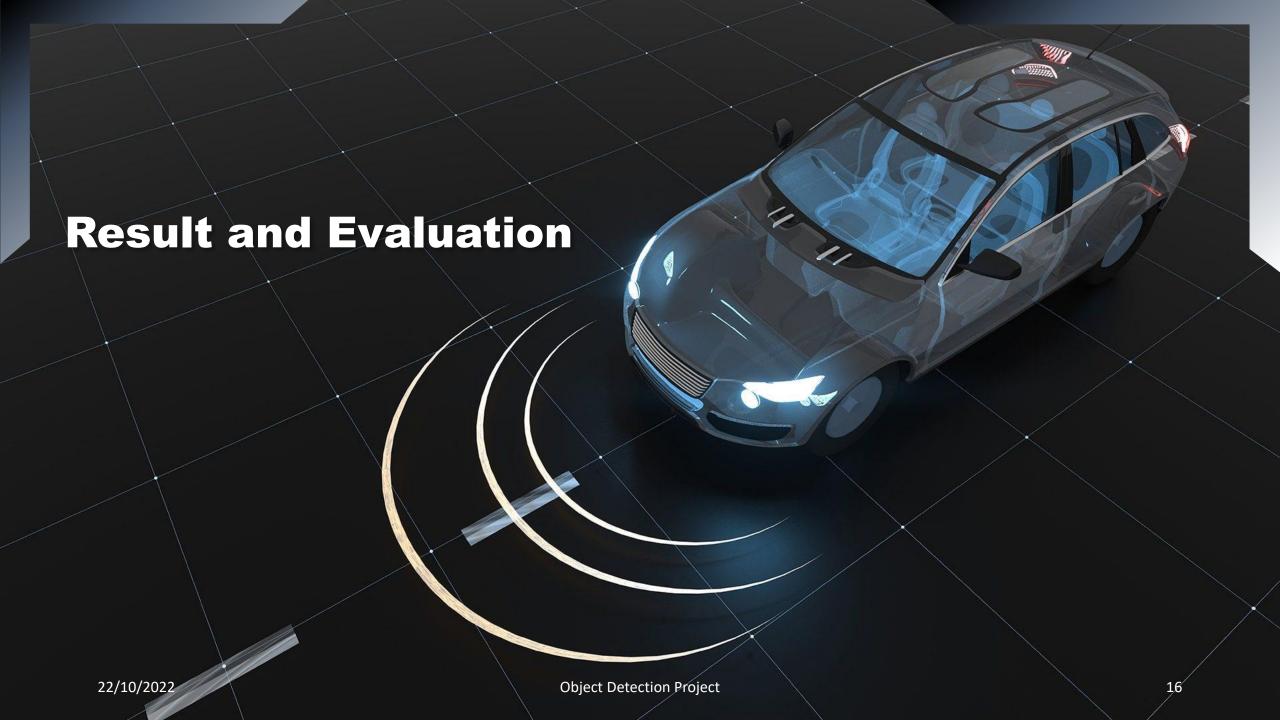
Softmax - Classification

#### **Modelling – Training Result**

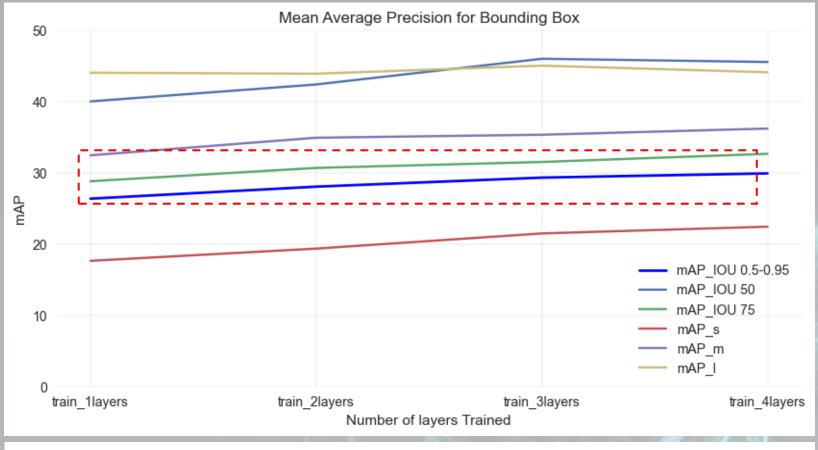
Model	Epoch	Training Time		
Freeze 1 <sup>st</sup> layer	10	3.54 hr		
Freeze 2 layers	10	3.32 hr		
Freeze 3 layers	10	3.10 hr		
Freeze 4 layers	10	2.48 hr		

- Loss is decreasing with epoch
- Accuracy is increasing with epoch
- Project resource consideration (Time and Hardware)



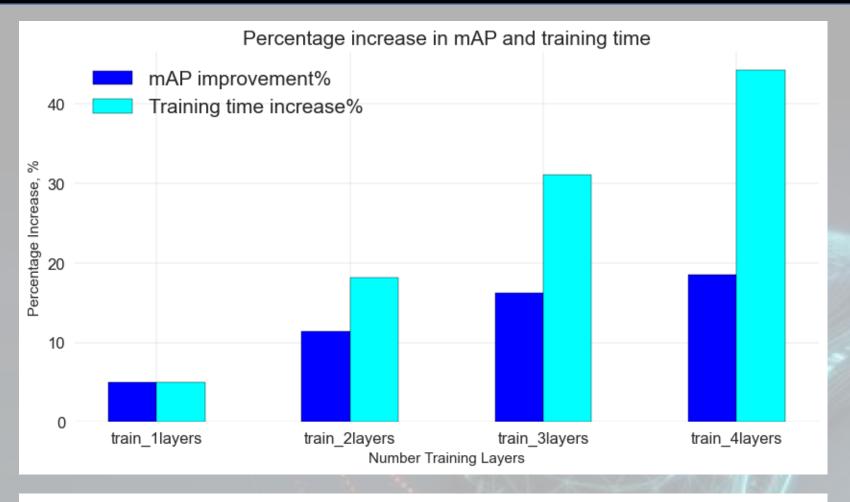


#### Mean Average Precision increase by training more RESNET layers



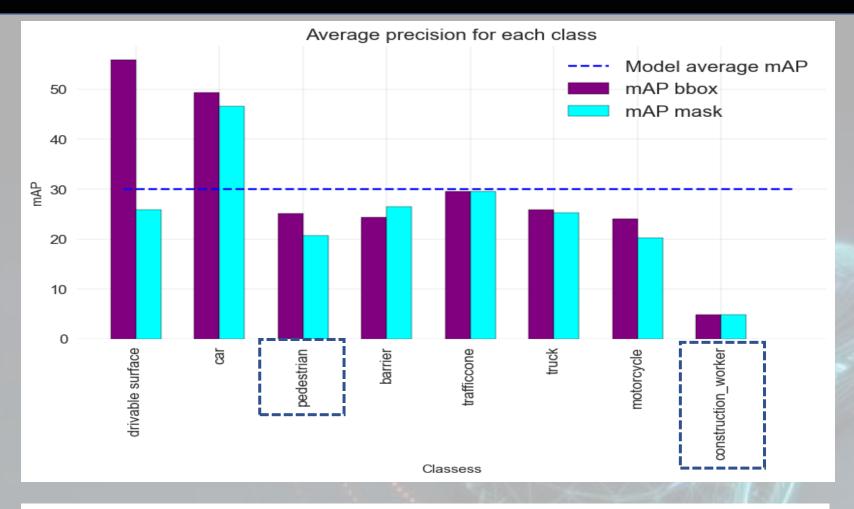
- mAP with IOU 0.75 is closer to average mAP, can represent the model general performance
- Most common threshold used in the industry
- Large object has better mAP
- Overall map result is not good (<0.5)

#### It is not worth to train more layers if project time is limited



- However, it is subjective to stakeholder's decision to get better mAP
- Model trained with 4 layers will be used as the final model because mAP is important to this project

#### Minority classes have poorer performance



- Construction worker image pixels is very close to pedestrian
- Relatively smaller object size, smaller object tends to have less accurate prediction.

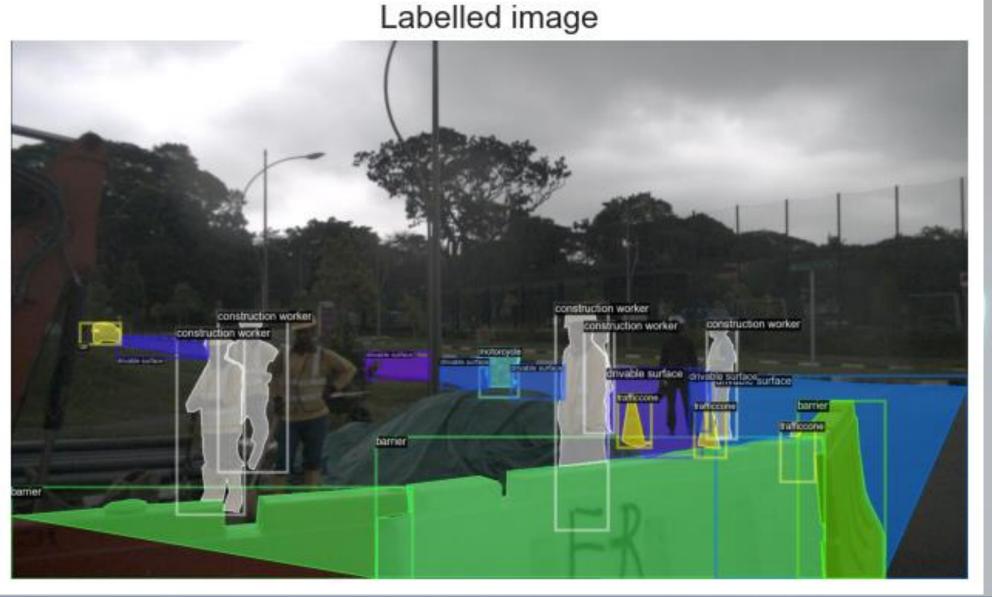
### **Recall for Construction worker class is only 0.25**

Г	Confusion Matrix								
Actual	drivable surface	1344	1	0	0	5	0	0	0
	car	0	930	0	0	0	2	0	0
	pedestrian	0	0	579	0	0	0	0	0
	barrier	0	0	0	310	3	0	0	0
	trafficcone	0	0	0	0	302	0	0	0
	truck	0	10	0	0	0	69	0	0
	motorcycle	0	1	1	0	0	0	56	0
	construction worker	0	0	18	0	0	0	0	6
	drivable surface car pedestrian barrier trafficcone truck motorcycle construction work  Predicted								

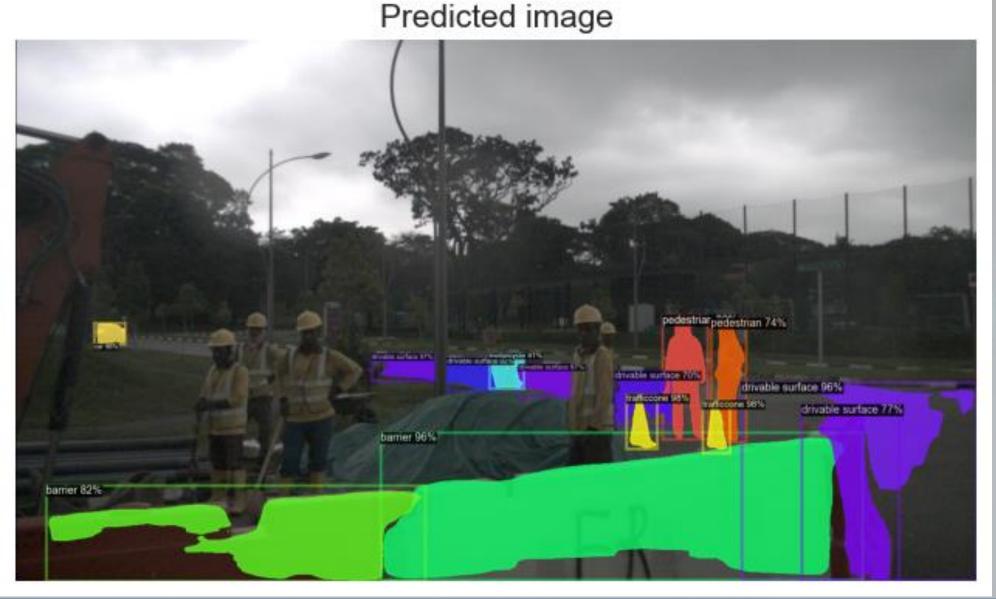
IOU Threshold = 0.75

Object Prediction threshold = 0.7

#### Miss annotation also causing object to has low mAP

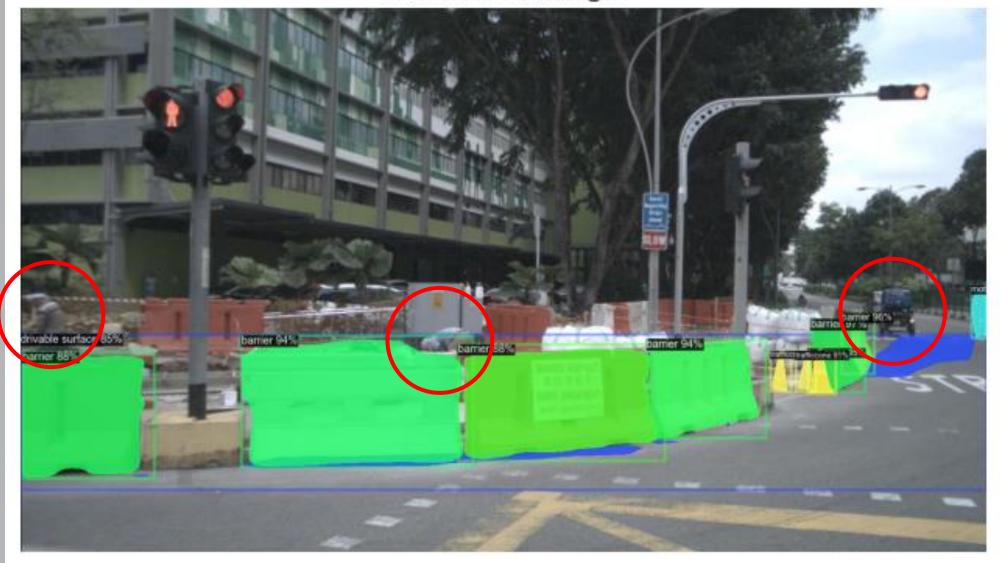


#### Worker that do not wear safety vest misclassified as pedetrian



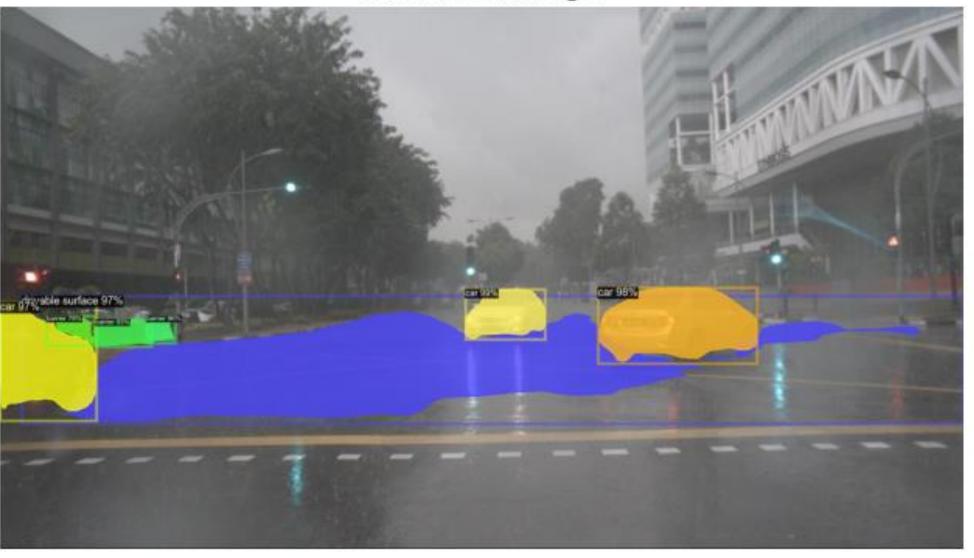
## Small size object is not detected

## Predicted image



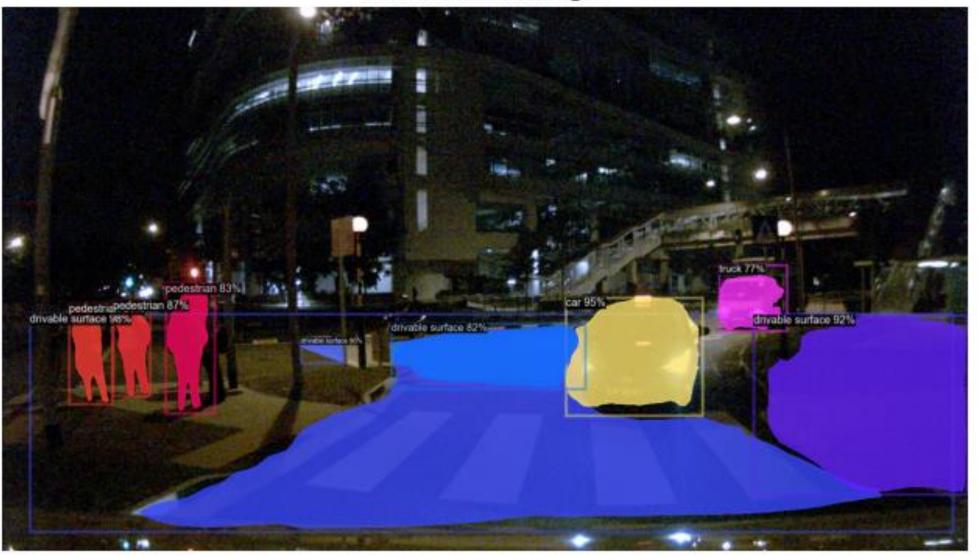
## Model does well for image in raining day

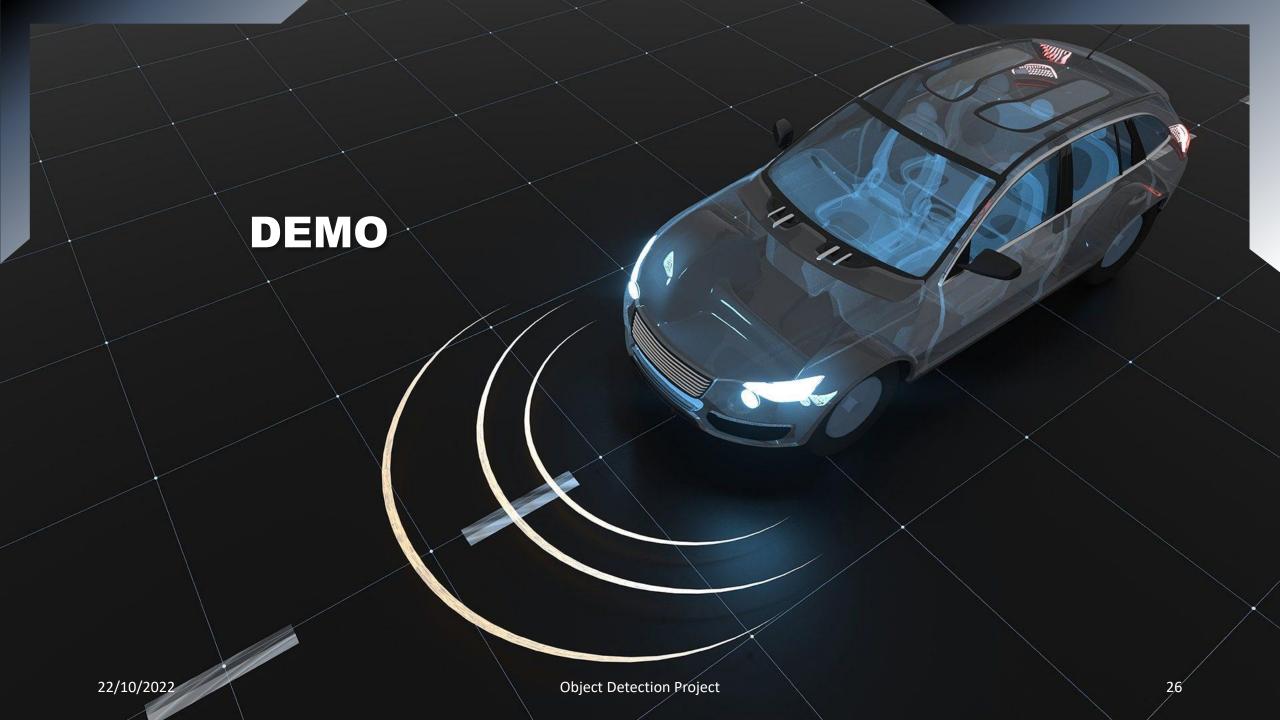
### Predicted image



## Model does well for image in night

#### Predicted image



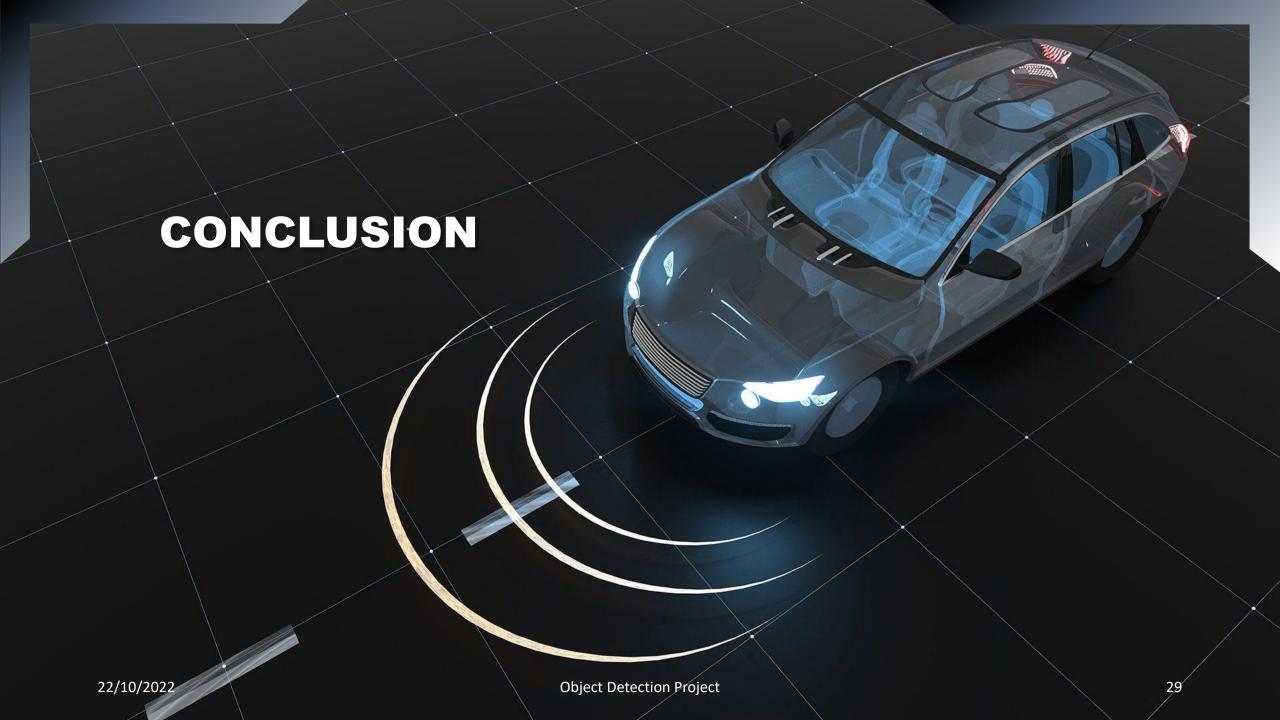


## DEMO 1



# DEMO 2





#### **Conclusion**

- Good performance in detecting the large objects.
- Good performance in detecting the object in the raining weather and night.

#### Weakness

- 1. Does not do a good job in distinguishing the **pedestrian and construction worker** classes.
- 2. Failed to detect the **object that is too small** that even human eye cannot easily spot it.
- 3. **mAP of the model is not very high** which might not meeting the industry requirement.
- 4. The low **mAP** is also due to wrong annotation in training data

#### Recommendation

- 1. Use more data to train the model.
- 2. Train the model with more epochs to achieve better performance.
- 3. Increase the target classes to build a complete object detection model.
- 4. Collect and annotate the data for the minority class like construction worker class.

