

ARI3129 Assignment

# Road Objects Detection

Ryan Vella  
Sean Diacono  
Andrew Aquilina



L-Università  
ta' Malta

# Contents

- ① Data Pre-Processing & Annotation
- ② Background & Foreground Estimation
- ③ YOLOv4
- ④ Mask R-CNN
- ⑤ Haar-Cascades
- ⑥ Evaluation
- ⑦ Video Analysis

# Data Pre-Processing & Annotation

- Trimming video clips (3 minutes total).
- Total number of frames is 2,544, each video having 15 FPS.
- Any noise present was omitted.



cloudy.avi



cloudy2.avi



night.avi



sunny.avi

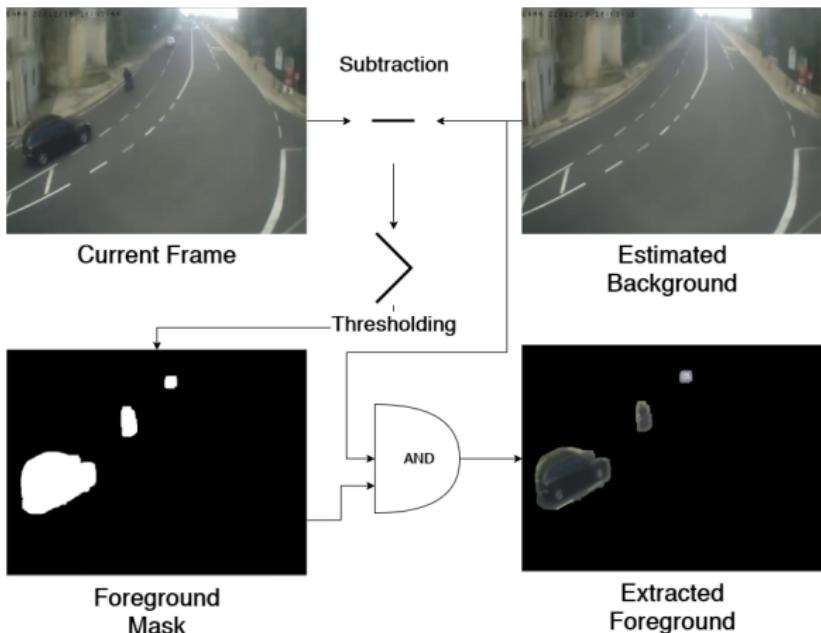


rainy.avi

- To annotate our data, we have made use of EVA, due to its object tracking capabilities.

# Background & Foreground Estimation

- The background is estimated by calculating the **median** pixel values.



# YOLOv4

- YOLOv4 by Bochkovskiy et al. based on YOLOv3 by Redmon et al.
- Improvement both in terms of speed and accuracy.

## Open-Source Implementation

An Open-Source Implementation of YOLOv4 using Darknet developed by Alexey Bochkovskiy was used.

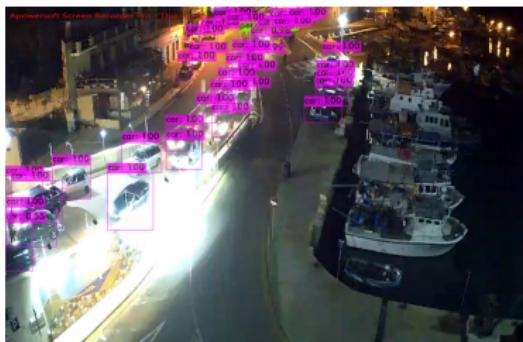
- Trained using Google Colaboratory Notebooks.
- 3 Labels: Car, Bike and Person.
- Trained for 6000 stages with a learning rate of 0.001 and a weight decay of 0.0005.

# YOLOv4

## Example Inferences



(a) Cloudy



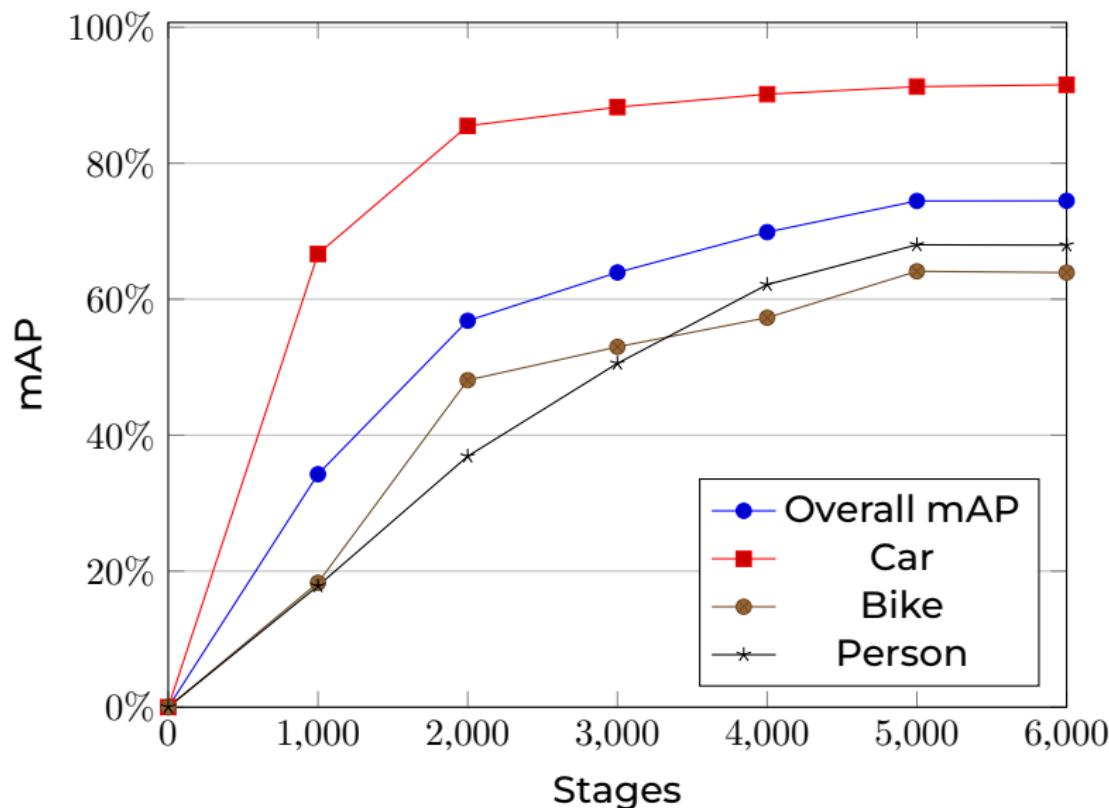
(b) Night



(c) Rainy



(d) Sunny



# Mask R-CNN

## Open-Source Implementation

An open-source implementation of the Mask R-CNN model was obtained from GitHub.

- Trained for 80 epochs with a learning rate of 0.001 and a weight decay of 0.0001.
- Data augmentation was also performed.



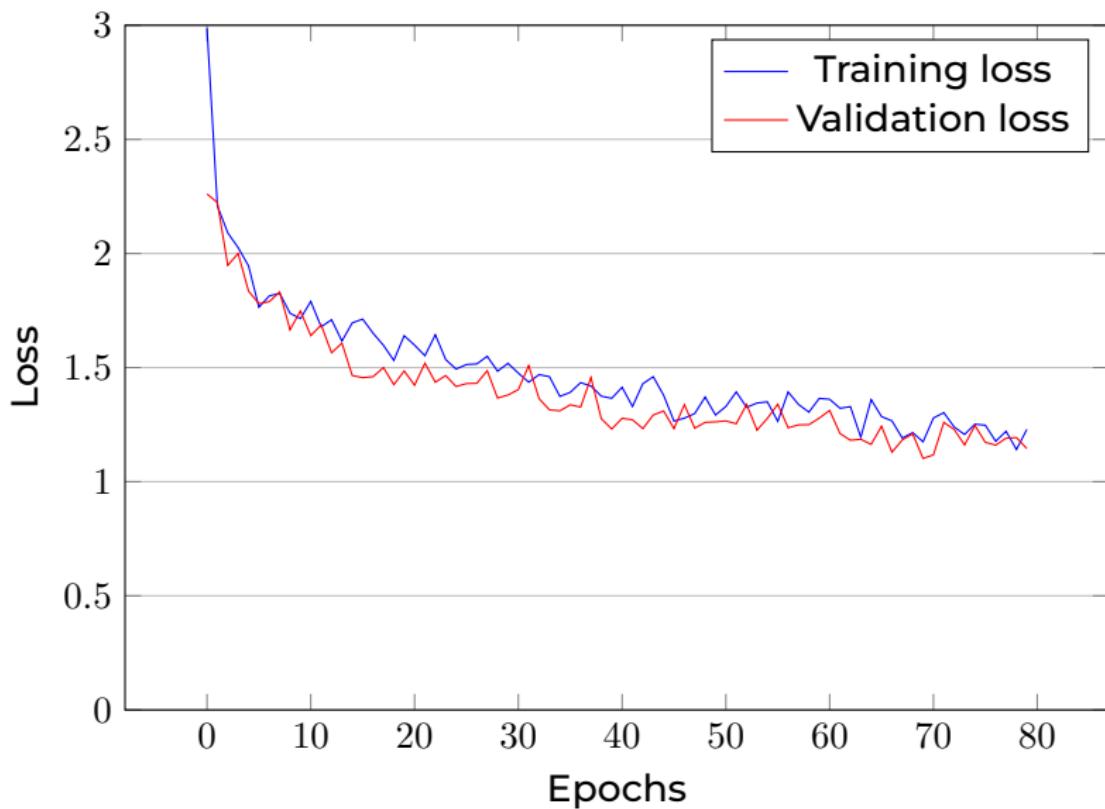
Training sample



Augmented sample

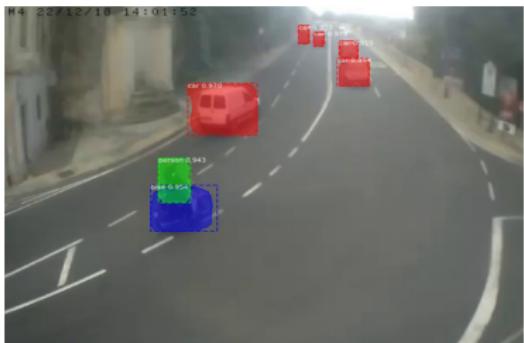
# Mask R-CNN

## Loss by Epochs Plot



# Mask R-CNN

## Example Inferences



(a) Cloudy



(b) Night



(c) Rainy



(d) Sunny

# Haar-Cascades

- An open-source cascading classifier for the car label was used.
- The Scale Factor and Minimum Neighbours parameters of the OpenCV function `detectMultiScale()` were set to 1.3 and 5 respectively.



# Haar-Cascades

## Example Inferences



(a) Cloudy



(b) Night



(c) Rainy



(d) Sunny

# Evaluation

Model	mAP	Car	Bike	Person	FPS
YOLOv4	74.45%	91.26%	64.10%	68.00%	44.50
Mask R-CNN	32.30%	65.30%	12.70%	19.00%	1.40
Haar-Cascade	0.50%	0.50%	N/A	N/A	21



(a) YOLOv4

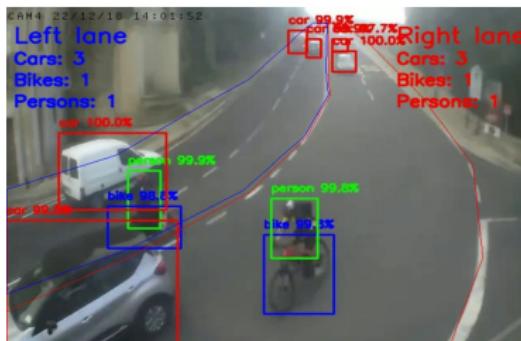


(b) Mask R-CNN

Unexpected Input

# Video Analysis

## Object Counting



Vehicle counting example on cloudy.avi.

Model	Cloudy	Night	Sunny	Rainy
YOLOv4	0.03	0.21	0.11	0.26
Mask R-CNN	0.21	6.58	0.35	1.92
Haar-Cascades	0.69	44.01	11.64	16.96

MSE values differentiated by weather conditions.

# Video Analysis

## Object Tracking



Vehicle tracking example on `cloudy.avi`.

Model	Cloudy	Night	Sunny	Rainy
YOLOv4	33	88	46	34
Mask R-CNN	71	45	117	35
Haar-Cascades	47	39	24	9
Ground Truth	28	57	36	26

Number of unique vehicles detected in each video.

# Thanks! Any Questions?

