

National Sun Yat-sen University
Department of Electrical Engineering

112-2 機器學習系統設計實務與應用

HW7Data Augmentation Evaluation

組別：第 8 組你說的隊

組員：B103012034 黃麗穎

B103012041 黃詣辰

B103015006 胡庭翊

B103015001 林佳明

一. 有無 augmentation 的差別



使用參數:epoch=32 ; input size :608

- 說明：augmentation 是一種機器學習中常使用的技術，在訓練過程中擴展 dataset，用來提升模型的泛化能力。而在上課時我們學到不少 augmentation 的方式，例如：水平翻轉、加噪、飽和度、亮度、對比度、銳度、旋轉、模糊等等。augmentation 的目的是增加 dataset 的多樣性，讓模型可以更好地學習特徵，也同時減少過擬合的風險。



● 訓練結果

i. 相片預測結果

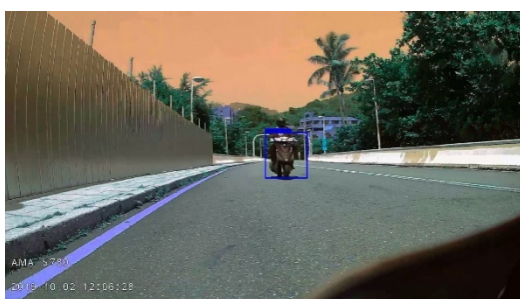
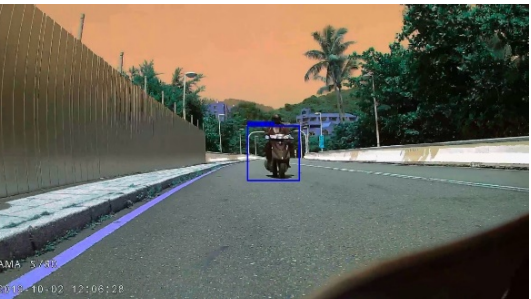
✧ 29.jpg

有 augmentation	無 augmentation
	
左框: Motor (0.62) ; 右框: car(0.79)	左:car(0.79) 中 car(0.80) 右: car(0.29)

✧ 186.jpg

有 augmentation	無 augmentation
	
左框:car(0.70) ; 右框: car(0.51)	左框:car(0.79) ; 右框: car(0.56)

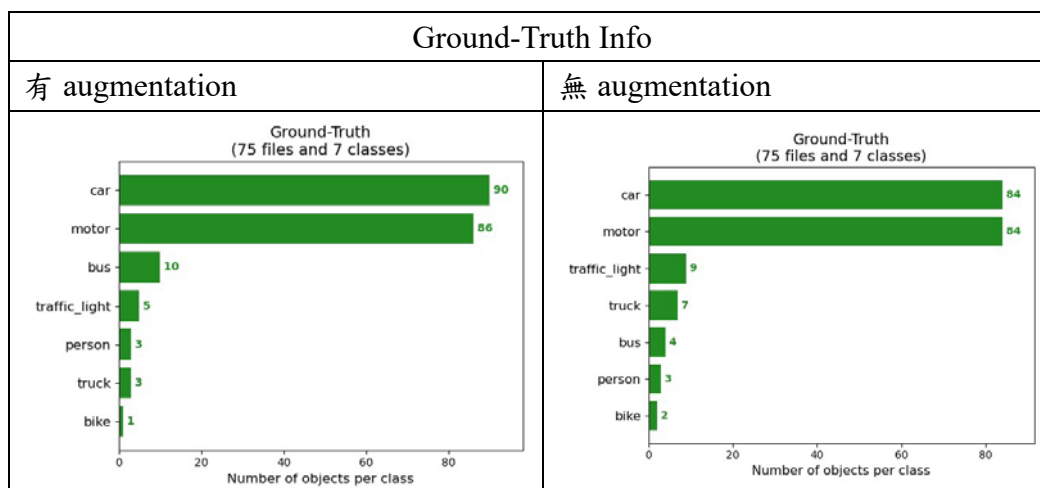
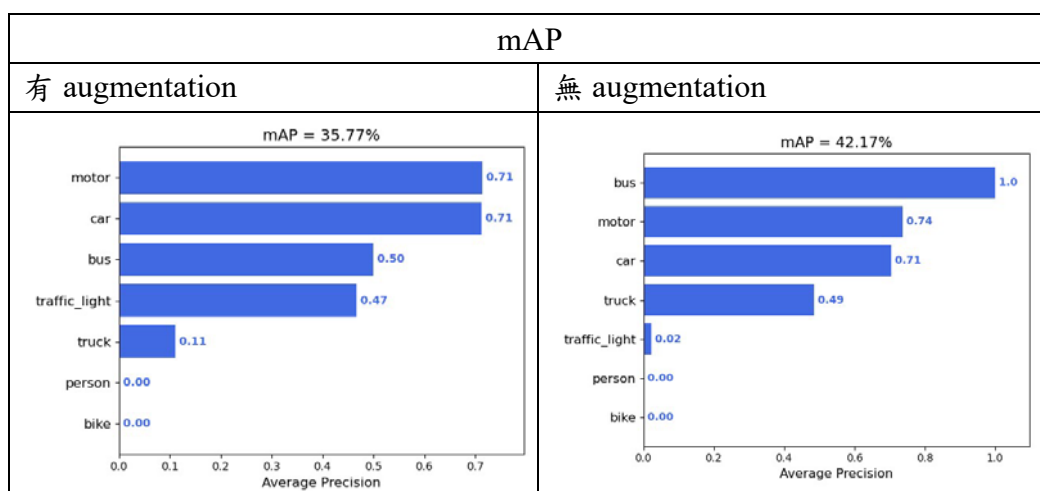
✧ 218.jpg

有 augmentation	無 augmentation
	

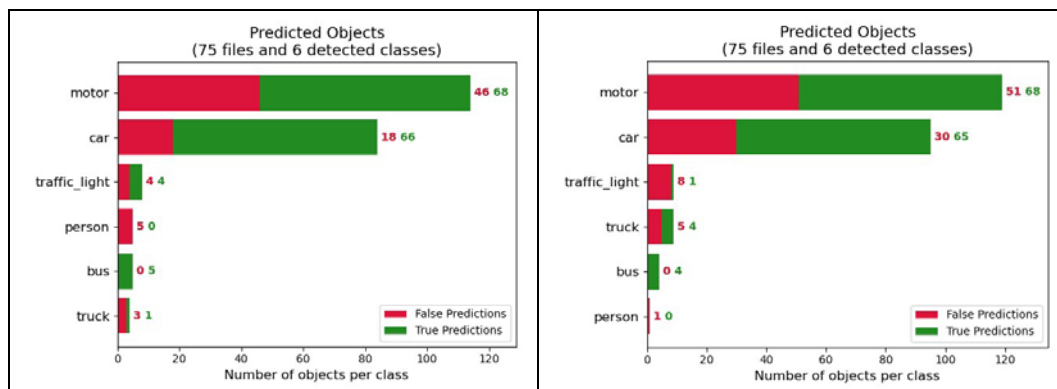
Motor(0.97)	Motor(0.93)
-------------	-------------

分析:在這幾張照片中，沒有使用資料增強的情況下，模型對 car 的信心指數較高，motor 較低。有使用資料增強的照片中，則是對於偵測 motor 的信心指數較高，像是 29.jpg 中，有使用資料增強的能偵測到機車，但對於右邊的車子就沒有辨識出來，而無資料增強的辨識不出左邊機車，但右邊兩輛汽車可以辨識出。這表示使用 augmentation 後，不一定每個類別的信心指數都會上升，也可能受到拆分資料集 7:3，訓練集資料不同，或是使用 augmentation 後沒有完整把不適用影像的刪除，特徵選取上造成辨識差異。

ii. 圖表結果



Predicted Objects Info	
有 augmentation	無 augmentation



分析:

mAP 及各類別精度: 無 augmentation 的模型在 mAP 和部分類別的 AP 表現更好, 特別是 bus 類別, 在無 augmentation 時, 平均精度達到 1.00, 而有 augmentation 時只有 0.50。在沒有資料擴充的情況下, 模型的整體效能更好, 表示資料擴充除了增加數據多樣性, 但也增加了模型學習的難度。但如果看 motor 和 car 類別, 其實 AP 沒有差距太大。而在 Traffic Light 類別中, 有 augmentation 後, AP 是大幅提高。

FP 與 TP: 在 motor 和 car 類別中, 有 augmentation 的 TP 大致維持不變, 但減少了 FP。Traffic Light 類別, 有 augmentation 提高了 TP, 可能是原本圖片中 Traffic Light 物件很小, 在擴充時放大圖片, 讓模型能夠分辨到更多特徵。

根據分析推測, 資料擴充在增加資料多樣性的同時, 有些預測會提高準確度, 但也可能導致模型在部分類別上的學習變得更困難, 這說明了為什麼有些類別在有 augmentation 時的表現反而比較差。

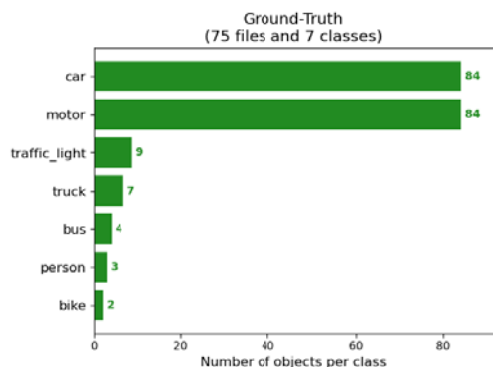
二. 設定 config 參數與成效

在固定資料集的情況下, 我們調整 config 內的各個參數進行討論與分析:

```

1  #!/usr/bin/env python
2  # coding=utf-8
3  from easydict import EasyDict as edict
4  _C = edict()
5  # Consumers can get config by: from config import cfg
6
7  cfg = _C
8
9  # YOLO options
10 _C.YOLO = edict()
11
12 _C.YOLO.CLASSES = ["./data/classes/train.names"]
13 _C.YOLO.ANCHORS = [12,16, 19,36, 40,28, 36,75, 76,55, 72,146, 142,110, 192,243, 459,401]
14 _C.YOLO.ANCHORS_V3 = [10,13, 16,30, 33,23, 30,61, 62,45, 59,119, 116,90, 156,198, 373,326]
15 _C.YOLO.ANCHORS_TINY = [23,27, 37,58, 81,82, 81,82, 135,169, 344,319]
16 _C.YOLO.STRIDES = [8, 16, 32]
17 _C.YOLO.STRIDES_TINY = [16, 32]
18 _C.YOLO.XYSCALE = [1.2, 1.1, 1.05]
19 _C.YOLO.XYSCALE_TINY = [1.05, 1.05]
20 _C.YOLO.ANCHOR_PER_SCALE = 3
21 _C.YOLO.IOU_LOSS_THRESH = 0.5
22
23
24 # Train options
25 _C.TRAIN = edict()
26
27 _C.TRAIN.ANOT_PATH = ["./data/dataset/train.txt"]
28 _C.TRAIN.BATCH_SIZE = 2
29 # _C.TRAIN.INPUT_SIZE = [320, 352, 384, 416, 448, 480, 512, 544, 576, 608]
30 _C.TRAIN.INPUT_SIZE = 608
31 _C.TRAIN.DATA_AUG = True
32 _C.TRAIN.LR_INIT = 1e-3
33 _C.TRAIN.LR_END = 1e-6
34 _C.TRAIN.WARMUP_EPOCHS = 2
35 _C.TRAIN.FIRST_STAGE_EPOCHS = 0
36 _C.TRAIN.SECOND_STAGE_EPOCHS = 32
37
38
39
40 # TEST options
41 _C.TEST = edict()
42
43 _C.TEST.ANOT_PATH = ["./data/dataset/test.txt"]
44 _C.TEST.BATCH_SIZE = 2
45 _C.TEST.INPUT_SIZE = 608
46 _C.TEST.DATA_AUG = False
47 _C.TEST.DECTECTED_IMAGE_PATH = ["./data/detection/"]
48 _C.TEST.SCORE_THRESHOLD = 0.25
49 _C.TEST.IOU_THRESHOLD = 0.5
50
51

```

1. Size: 208 vs 416 vs 608
Epoch=32

288	416	608																																																															
<p>Predicted Objects (75 files and 6 detected classes)</p> <table><tr><th>Class</th><th>False Predictions</th><th>True Predictions</th></tr><tr><td>motor</td><td>70</td><td>57</td></tr><tr><td>car</td><td>40</td><td>62</td></tr><tr><td>traffic_light</td><td>17</td><td>0</td></tr><tr><td>person</td><td>6</td><td>0</td></tr><tr><td>truck</td><td>4</td><td>2</td></tr><tr><td>bus</td><td>2</td><td>4</td></tr></table>	Class	False Predictions	True Predictions	motor	70	57	car	40	62	traffic_light	17	0	person	6	0	truck	4	2	bus	2	4	<p>Predicted Objects (75 files and 6 detected classes)</p> <table><tr><th>Class</th><th>False Predictions</th><th>True Predictions</th></tr><tr><td>motor</td><td>51</td><td>68</td></tr><tr><td>car</td><td>30</td><td>65</td></tr><tr><td>traffic_light</td><td>8</td><td>1</td></tr><tr><td>truck</td><td>5</td><td>4</td></tr><tr><td>bus</td><td>0</td><td>4</td></tr><tr><td>person</td><td>1</td><td>0</td></tr></table>	Class	False Predictions	True Predictions	motor	51	68	car	30	65	traffic_light	8	1	truck	5	4	bus	0	4	person	1	0	<p>Predicted Objects (75 files and 6 detected classes)</p> <table><tr><th>Class</th><th>False Predictions</th><th>True Predictions</th></tr><tr><td>motor</td><td>64</td><td>71</td></tr><tr><td>car</td><td>32</td><td>69</td></tr><tr><td>traffic_light</td><td>7</td><td>1</td></tr><tr><td>truck</td><td>4</td><td>3</td></tr><tr><td>bus</td><td>0</td><td>4</td></tr><tr><td>person</td><td>3</td><td>0</td></tr></table>	Class	False Predictions	True Predictions	motor	64	71	car	32	69	traffic_light	7	1	truck	4	3	bus	0	4	person	3	0
Class	False Predictions	True Predictions																																																															
motor	70	57																																																															
car	40	62																																																															
traffic_light	17	0																																																															
person	6	0																																																															
truck	4	2																																																															
bus	2	4																																																															
Class	False Predictions	True Predictions																																																															
motor	51	68																																																															
car	30	65																																																															
traffic_light	8	1																																																															
truck	5	4																																																															
bus	0	4																																																															
person	1	0																																																															
Class	False Predictions	True Predictions																																																															
motor	64	71																																																															
car	32	69																																																															
traffic_light	7	1																																																															
truck	4	3																																																															
bus	0	4																																																															
person	3	0																																																															
<p>mAP = 35.40%</p> <table><tr><th>Class</th><th>Average Precision</th></tr><tr><td>bus</td><td>1.0</td></tr><tr><td>car</td><td>0.65</td></tr><tr><td>motor</td><td>0.59</td></tr><tr><td>truck</td><td>0.24</td></tr><tr><td>traffic_light</td><td>0.00</td></tr><tr><td>person</td><td>0.00</td></tr><tr><td>bike</td><td>0.00</td></tr></table>	Class	Average Precision	bus	1.0	car	0.65	motor	0.59	truck	0.24	traffic_light	0.00	person	0.00	bike	0.00	<p>mAP = 42.17%</p> <table><tr><th>Class</th><th>Average Precision</th></tr><tr><td>bus</td><td>1.0</td></tr><tr><td>motor</td><td>0.74</td></tr><tr><td>car</td><td>0.71</td></tr><tr><td>truck</td><td>0.49</td></tr><tr><td>traffic_light</td><td>0.02</td></tr><tr><td>person</td><td>0.00</td></tr><tr><td>bike</td><td>0.00</td></tr></table>	Class	Average Precision	bus	1.0	motor	0.74	car	0.71	truck	0.49	traffic_light	0.02	person	0.00	bike	0.00	<p>mAP = 43.11%</p> <table><tr><th>Class</th><th>Average Precision</th></tr><tr><td>bus</td><td>1.0</td></tr><tr><td>motor</td><td>0.78</td></tr><tr><td>car</td><td>0.75</td></tr><tr><td>truck</td><td>0.43</td></tr><tr><td>traffic_light</td><td>0.06</td></tr><tr><td>person</td><td>0.00</td></tr><tr><td>bike</td><td>0.00</td></tr></table>	Class	Average Precision	bus	1.0	motor	0.78	car	0.75	truck	0.43	traffic_light	0.06	person	0.00	bike	0.00															
Class	Average Precision																																																																
bus	1.0																																																																
car	0.65																																																																
motor	0.59																																																																
truck	0.24																																																																
traffic_light	0.00																																																																
person	0.00																																																																
bike	0.00																																																																
Class	Average Precision																																																																
bus	1.0																																																																
motor	0.74																																																																
car	0.71																																																																
truck	0.49																																																																
traffic_light	0.02																																																																
person	0.00																																																																
bike	0.00																																																																
Class	Average Precision																																																																
bus	1.0																																																																
motor	0.78																																																																
car	0.75																																																																
truck	0.43																																																																
traffic_light	0.06																																																																
person	0.00																																																																
bike	0.00																																																																

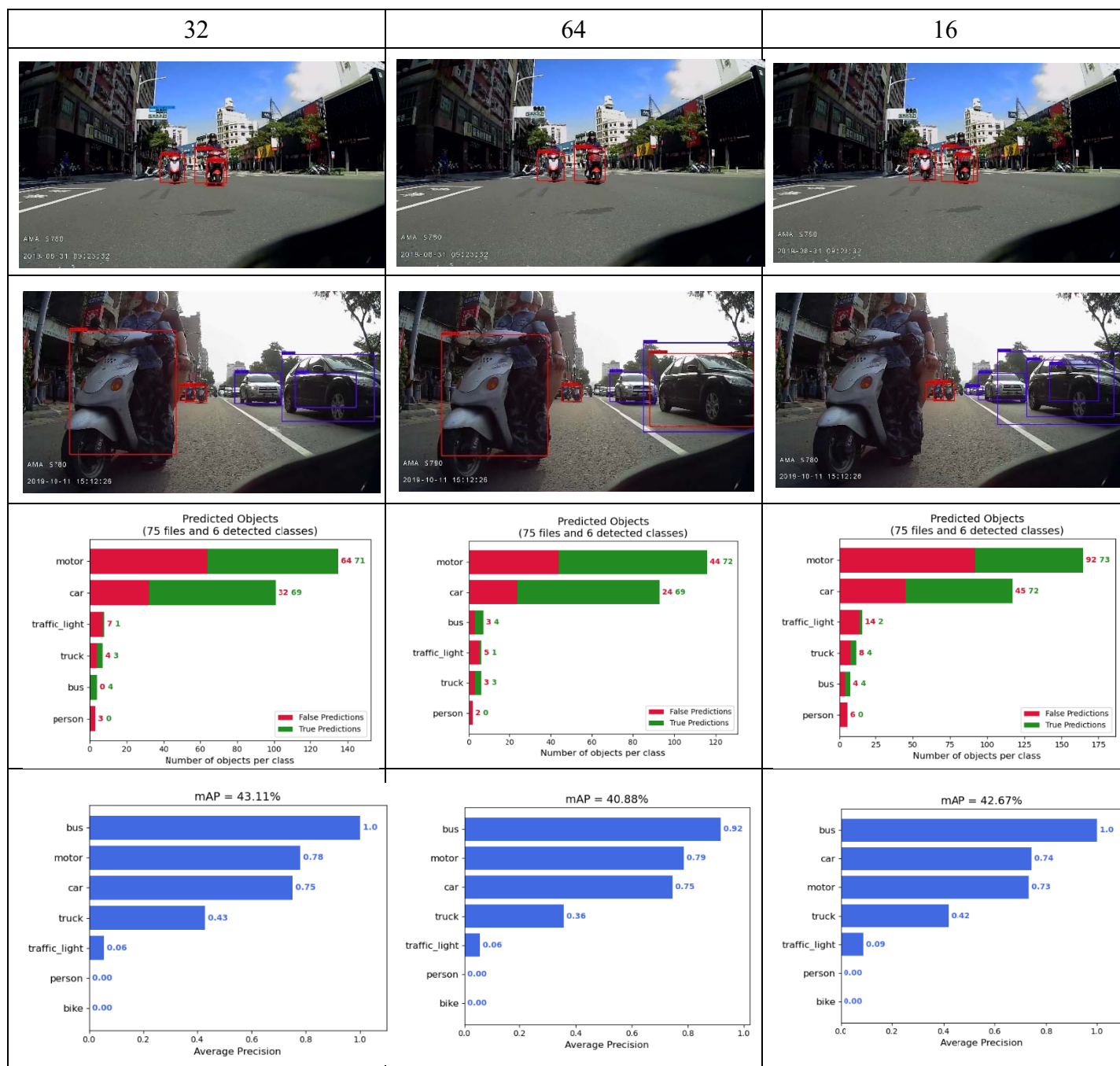
成效分析:

Size 指的是輸入的圖片大小，越大的影像維度因為擁有更多像素，所以可以保留更多細節，有助於提高 True Prediction 的數量，更精確的辨識樣本，較大的 size 能幫助模型提取更多有效特徵使得 mAP 提高。實際操作的結果我們可以看到隨著 size 變大，True Precisions 跟著變大，False Predictions 變小，

mAP 增加。

2. Epoch: 32 vs 64 vs 16

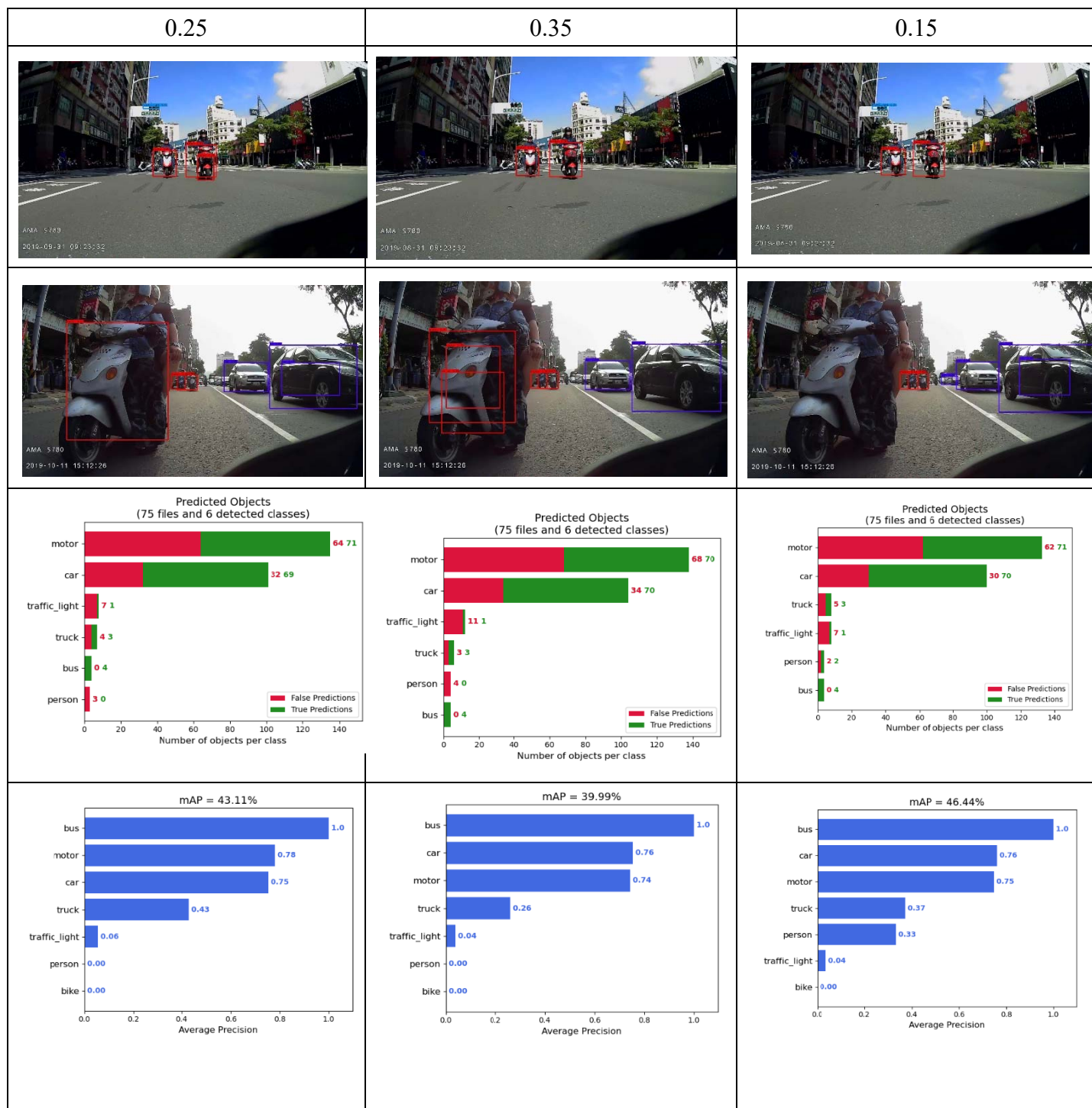
Size = 608



成效分析:

Epoch 是訓練週期，通過多次的迭代訓練，有助於改善模型的性能以及泛化能力。然而，epoch 數並不是越大越好，如果使用過多的 epoch，模型可能會過度擬合導致在新的資料上實測表現不佳。每種資料集最適合的 epoch 數不同，因此在實際應用中通常須經測試才能找到合適的 epoch 數。由上表可以觀察到當訓練次數增加 True Predictions 沒有明顯變化，False Predictions 數量減少，mAP 訓練次數最多反而下降，可能是因為產生過度擬合的情況使得在訓練數據上表現很好，但在測試數據上表現不佳。

3. Test score threshold: 0.25 vs 0.35 vs 0.15



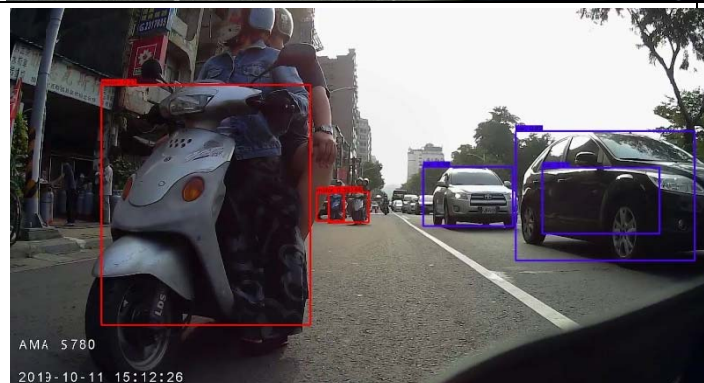
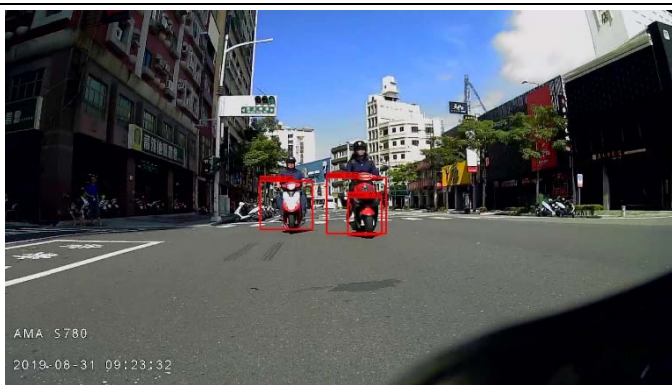
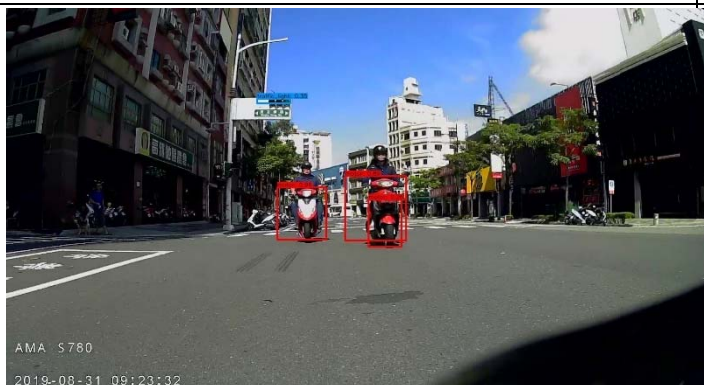
成效分析:

Test score threshold 指的是預測的信賴分數閾值，調整這個值會改變模型在預測中必須具備的可信度等級，用以將某個預測資料視為陽性。也會改變在預測中能容忍多少偽陽性和偽陰性結果。當 test score threshold 提高，因為模型只有在非常有信心的情況下才會將樣本分類為正類，因此 TP 會下降，TN 會增加，FP 會下降，FN 會增加，mAP 則取決於 Precision 和 Recall 的平衡。實驗結果為當 test score threshold 變大，True Predictions 沒有太大的變化，False Predictions 變越大，mAP 則變小。

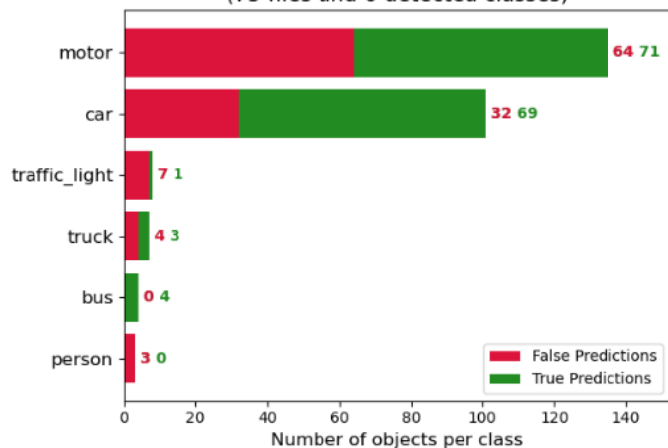
4. Batch size: 2 vs 1

2

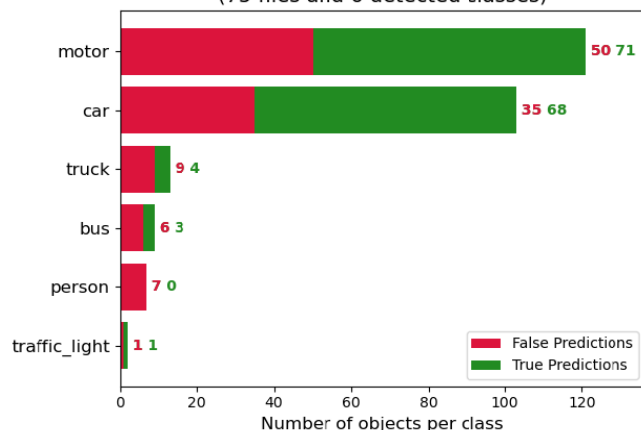
1



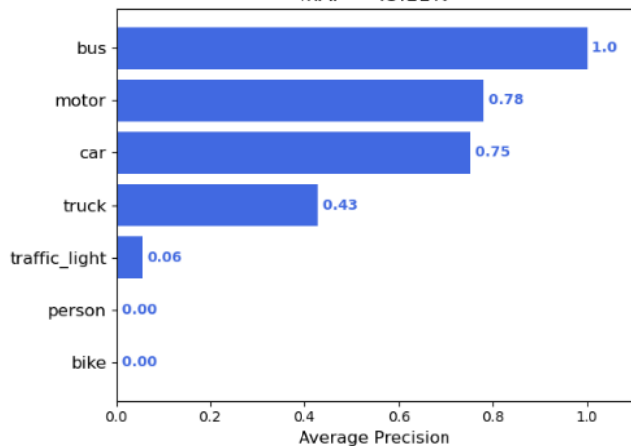
Predicted Objects
(75 files and 6 detected classes)



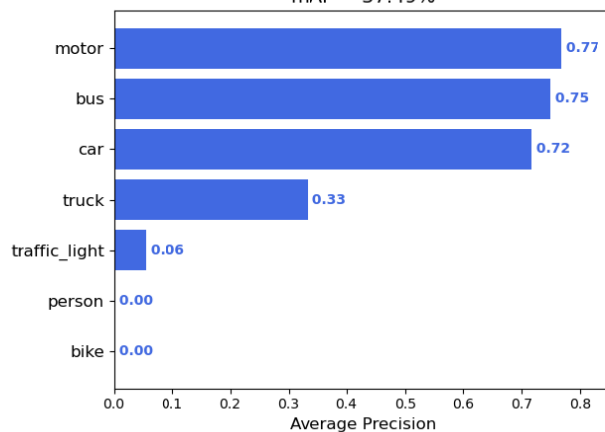
Predicted Objects
(75 files and 6 detected classes)



mAP = 43.11%



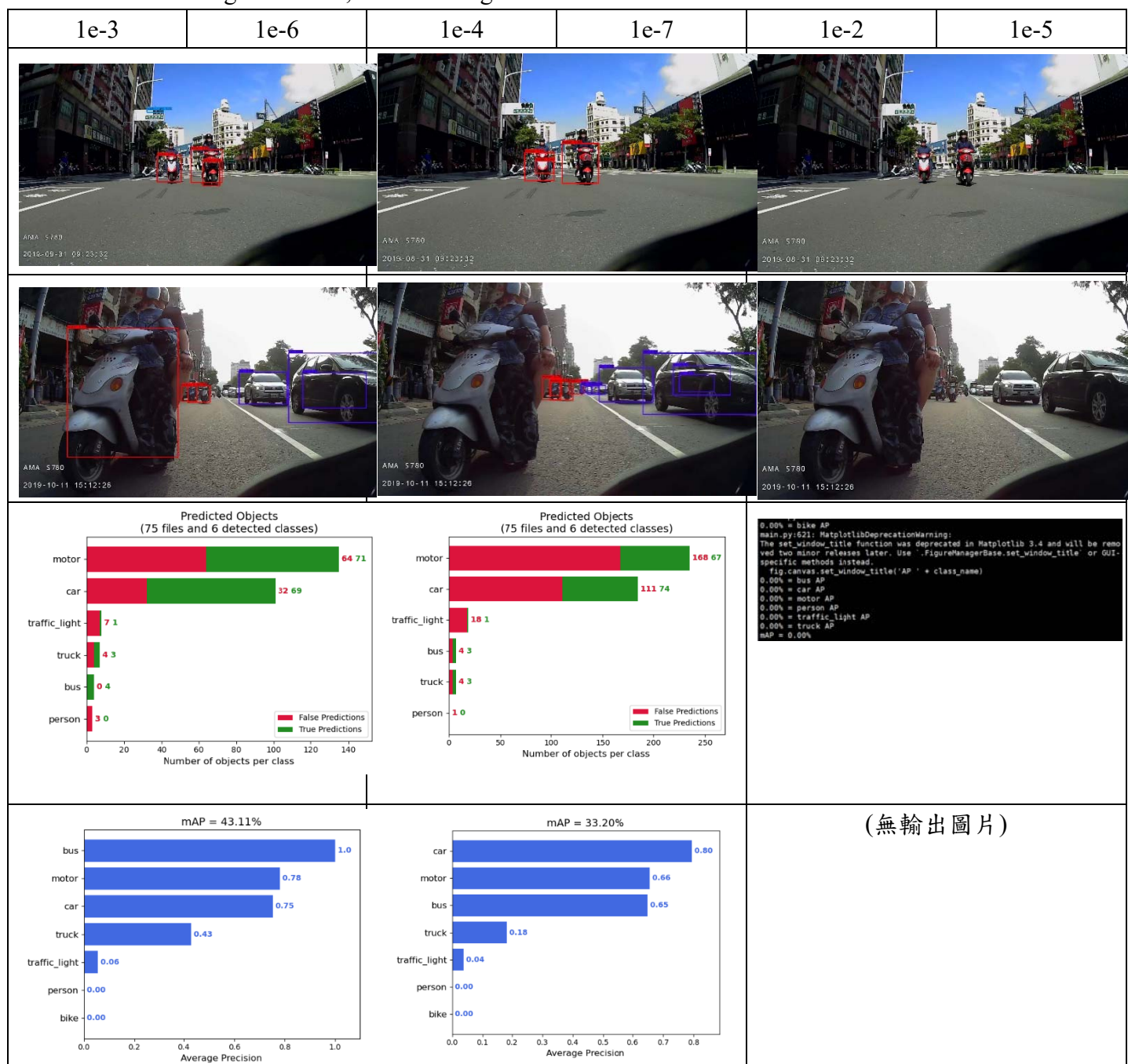
mAP = 37.49%



成效分析:

Batch size 指的是一次訓練所抓的訓練樣本數。Batch Size 的大小影響模型的最佳化程度和速度，在合理範圍內，一般來說 batch size 越大，其確定的下降方向越準確，引起的訓練震盪會越小，然而 batch size 的增加也代表跑完一次 Epoch 的迭代次數減少，要想達到相同的精度，其所花費的時間的時間會因此大幅增加。不過因為記憶體空間有限，我們只能夠執行 batch size<2 的訓練，觀察實作結果，batch size 變大 True Predictions 沒有太大的變化，False Predictions 則是在 motor 的類別中較顯著的減少，可能是因為較大的 batch size 可以減少每個梯度更新中的噪音。較大的 batch size 也能提高訓練過程的穩定性和收斂速度，有助於在訓練數據上有更好的效能，使得 mAP 提高。

5. Train learning rate initial, train learning rate end

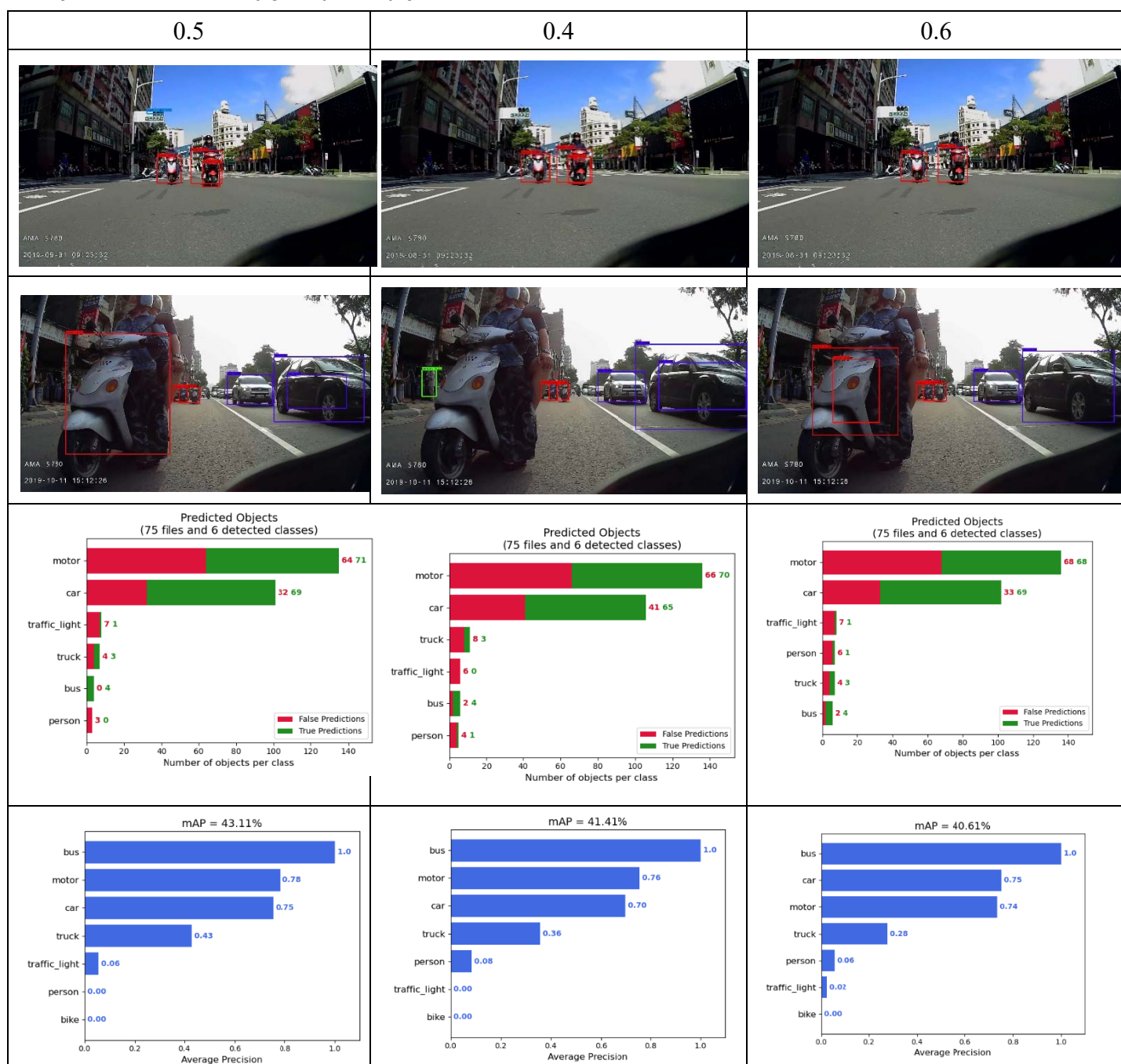


成效分析:

TRAIN.LR_INIT 與 TRAIN.LR_END 分別代表的是學習率 learning rate 的初始值與結束值。學習率為控制模型中梯度下降的速度，也有人稱為步長。公式為：新權重= 舊權重- 學習率* 梯度。合適的學習率可以幫助模型快速收斂，常見的調整方法是在訓練初期時給定較大的學習率，隨著模型的訓練逐漸調低學習率。當學習率的初始值越大，模型在訓練早期可以快速收斂減少錯誤預判數量，因此 False

Predictions 會減少，mAP 提高，但是需要慎防學習速率結束值過大無法收斂導致 mAP 下降的可能。

6. IOU threshold: 0.5 vs 0.4 vs 0.6



成效分析:

TEST.IOU_threshold 指的是對 IoU 計算設定的閾值。在影像物件偵測領域上，都是先選出物件候選人，然後再於物件候選人中判斷是不是物件，但有可能一個物件被很多候選框給選到，這時就要利用 Non-Maximum Suppression 的方法去消除多餘物件框找到最佳的框。而在 NMS 流程中，如果 IoU 閾值設定太高，可能會造成物件重複偵測的問題，顯示的預測框會很多；IOU 閾值低，顯示的預測框會變少。當 IOU 閾值越大時，模型需要預測的 bounding box 與真實 bounding box 之間的重疊面積越大才能被視為正確預測，導致 True Predictions 數量減少，False Predictions 數量增加。mAP 考慮了精度和召回率的平衡。隨著 IoU 閾值的增加，雖然精度可能有所提高，但召回率的下降會更顯著，通常導致 mAP 降低。

分工比例表:

組員	黃麗穎	黃詣辰	胡庭翊	林佳明
工作比例	25%	25%	25%	25%
簽名	黃麗穎	黃詣辰	胡庭翊	林佳明

參考資料:

1. <https://chih-sheng-huang821.medium.com/%E6%A9%9F%E5%99%A8-%E6%B7%B1%E5%BA%A6%E5%AD%B8%E7%BF%92-%E7%89%A9%E4%BB%B6%E5%81%B5%E6%B8%AC-non-maximum-suppression-nms-aa70c45adffa>
2. https://www.cupoy.com/qa/club/ai_tw/0000016E62FABB7A000000016375706F795F72656C656173654B5741535354434C5542/0000016F7D7BD5D1000001916375706F795F72656C656173655155455354
3. https://docs.aws.amazon.com/zh_tw/machine-learning/latest/dg/step-4-review-model-and-set-cutoff.html