

Deep-Learning-for-Autonomous-Driving

Lab3 : Object Detection

ID : 310605007

Member : 鄭晴立

Department : 機器人學程

1 Introduction

首先，本次的時間較為緊迫，而目前 YOLOv4 在研究上已經有不錯的準確度及速度表現，因此我評估後認為，對於各種參數的修正或是 model 的改進皆難以達成顯著的效果。最終我採用 YOLOv7 去嘗試實作。

2 YOLOv4

2.1 Implement

由於沒辦法一一嘗試各項參數影響，因此在 YOLOv4 的實作中，基本上就是講文獻中提到的方法都有使用，例如：mosaic 為 YOLO 中很重要的 data augmentation；pretrained model 及 Cosine learning 可以大幅提升 training 的穩定度及速度。

表一 YOLOv4 參數設定

Parameter	
Backbone	Mobilenetv3
Optimizer	Adam
Learning rate	StepLR
Epoch	300
Loss function	CIoU
Batch size	16
Smooth label	0
mosaic	True
Cosine learning	True
pretrained	True
Unfreeze Epoch	50

2.2 Conclusion

如我所預估的 YOLOv4 已經有不錯的效果，即便我嘗試不同的 model 及參數，皆無法有效的提升，而在時間有限的情況下，我決定直接常是 yolo7，因此下一章節介紹，我使用 yolo7 的結果。



```
Auto 0.99528384 222 140 284 287
Auto 0.98059607 118 182 215 412
Car 0.992629 206 115 231 147
Car 0.9670038 199 129 226 173
```



```
Bus 0.99073875 196 169 248 284
LCV 0.9807065 106 192 135 246
LCV 0.8000286 139 144 163 184
```

圖一 YOLOv4 偵測結果，下方為其表現數據和方框座標

3 YOLOv7

由於上述原因我嘗試使用 YOLOv7，而由於 YOLOv7 是 Pytorch 架構，因此理論上可以直接用上，但後來在 Datasets 部分遇到問題所以無法實現。

3.1 Datasets

然而，YOLOv7 在 data 處理上有極大不同。在 YOLOv7 中所使用的 coco datasets 的 4 個座標分別是長、寬跟物體中心的座標，而助教所給的 highway datasets 則是 bounding box 左上角的座標及右下角的座標，如下圖。。

除 label 方式的差異外，YOLOv7 的匯入資料方式也有極大不同，可以看到 yolo7 是一個圖片對應一個 txt 檔，每一個 txt 檔裡面則記錄一個圖片的 label，如下圖。



圖二 yolo7 data



圖 三 以這張圖為例，在 YOLOv7 label 的座標格式下，左邊的人為(0.31, 0.4, 0.62, 0.97) ;領帶為 (0.35, 0.57, 0.09, 0.40)，以左上為原點，表示為 x、y、寬、高，對應於整體邊界的比例，可以看出其中心點為物體正中央



圖 四 以這張圖為例，其為 416x416，在 YOLOv4 label 的座標格式下，右邊的 Auto 為(222, 140, 284, 287) 、下面的 Auto 為 (118, 182, 215, 412) 、上面的 Car 為 (206, 115, 231, 147)、下面的 Car 為(199, 129, 226, 173)，以左上為原點，表示左上點的座標及右下點的座標。

3.2 Implement

由於 Datasets 的差異，沒有足夠時間調整 Data 的 labels，因此只好作罷，直接使用 Coco datasets train 出的 YOLOv7 去使用，可以看到，我使用 100 張圖片去偵測，僅花了 2.972 秒。

而實際辨識成果上，由於是用 Coco datasets train 出來的，可以看到會有很多不同的東西被偵測出來，以下面四張圖為例，他會很容易辨識到路人，較大的車輛比較容易辨識成功，而深色的車輛容易辨識成火車，聯結車則會偵測成公車。

而從 confusion matrix 可以看出 80 樣 coco 的分類都有很好的結果，而在上一次 lab 的 data 中也蠻不錯的，在日常生活的照片也有不錯的成果。

Done. (2.972s)

圖 五 yolo7 偵測 100 張圖片所花的時間



圖 六 yolov7 偵測成果

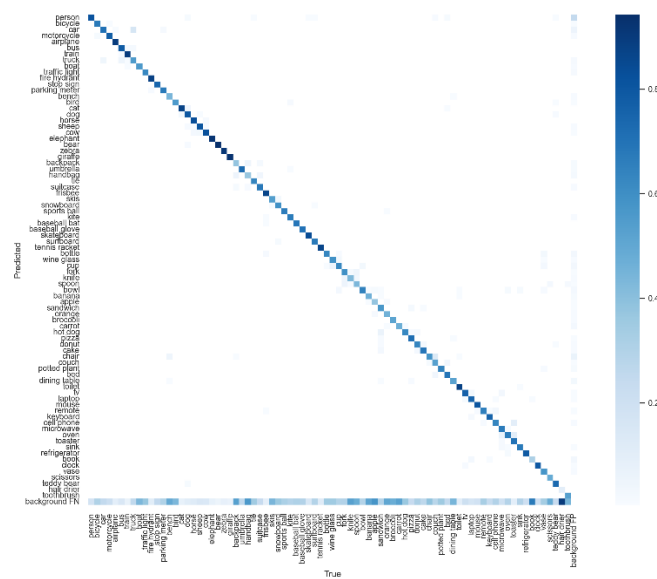


圖 七 yolov7 的 confusion matrix

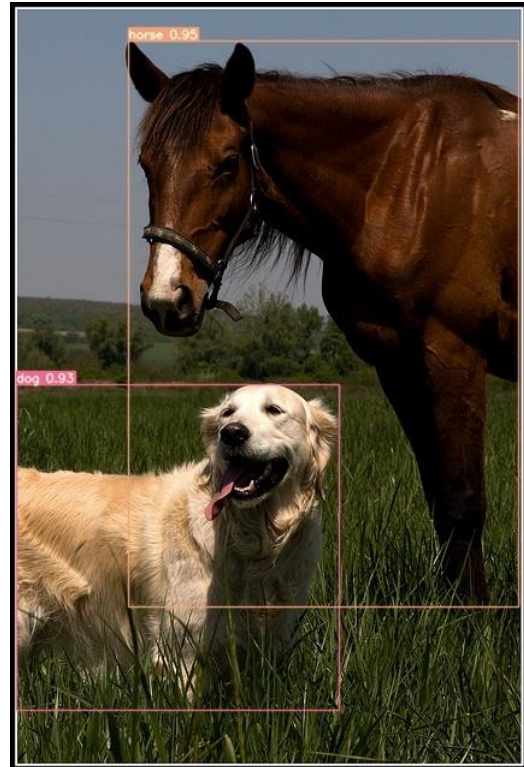


圖 八 YOLOv7 的其他偵測成果，左下為我房間，右下為上一次 lab 的 data

4 Discussion

4.1 原先的結果已經不錯，難以有顯著提升

YOLOv4 的精準度已經有不錯的成果，而本次實驗的 data 也較為單純，因此不做任何修正就可以有不錯的表現，改變 backbone 或是任何參數修正，皆無法達到顯著的效果，甚至會拖慢其辨識速度，而 yolo 本身的特色在於其便是速度，因此，我認為太著重於精確的調整是沒有意義的，所以我才會直接以 yolo v7 作為參考去研究。