

111 智慧型系統專題成果

輕軌鐵道障礙物偵測

408415005 吳致廣

408415032 潘昱文

408415008 林坤翰

指導教授：林惠勇 教授

Department of Electrical Engineering
National Chung Cheng University
Chiayi 621, Taiwan

中華民國 一百一十一年七月八號

目錄

一、摘要	P.3
二、研究動機	P.3
三、輕軌鐵道障礙物偵測	P.3
3.1 系統流程	P.3
3.2 YOLOv5	P.3
3.3 Yolact Edge	P.5
3.4 異物資料集	P.5
3.5 軌道資料集	P.6
四、評估與實驗結果	P.7
4.1 YOLOv5 結果與評論	P.7
4.2 Yolact Edge 結果與評論	P.8
4.3 成果展示	P.9
五、結論	P.10
5.1 研究結論	P.10
5.2 未來展望	P.10
5.3 預期困難	P.10
六、參考資料	P.11

一、摘要

透過在車頭配戴攝影機，獲得第一人稱視角行駛影像。先利用 YolactEdge 去將鐵軌的 ROI 分割出來，再利用 YOLOv5 對畫面中可能是障礙物的物體作辨識及定位並稍作拓寬，總共有 8 個類別(person、bike、car、train、rider、bus、truck、motor)，偵測障礙物是否有與鐵軌重疊，如果重疊會將影像顯示不同顏色以示警告。

二、研究動機

近年來，隨著台灣不少輕軌系統已通車或即將通車，輕軌在平面道路上與一般行車發生事故的機率會越來越高。

此外，我們也對於機器視覺領域有相當高的興趣，我們希望能夠透過專題研究，利用機器視覺來減少輕軌相關的事故，希望透過這次的專題能夠更加了解關於深度學習以及機器視覺的基本知識，並且能夠活用模型去達成我們的目的。

三、輕軌鐵道障礙物偵測

3.1 系統流程

我們將圖片的軌道辨識與異物偵測分成兩個部分去處理，從物件偵測的結果取得異物所在的位置，將異物位置稍作拓寬再去判斷偵測到的異物是否可能遭受撞擊，最後用顏色去警示駕駛軌道上的情況，紅色表示軌道上有可能會撞擊到的異物，綠色則否。(流程圖如圖 1.)

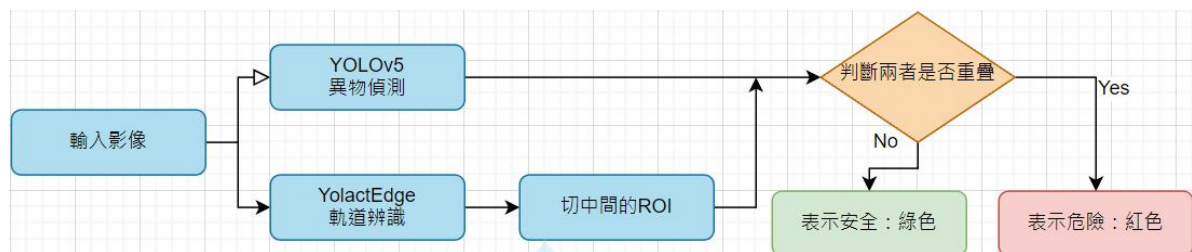


圖 1.系統流程圖

3.2 YOLOv5

YOLO 系列在物件偵測上擁有非常快的速度和不錯的準確度，在考量到軌道上的異物入侵多屬於突發性的狀況，因此選擇 YOLO 系列中由 Ultralytics 開發最新的 v5 模型作為我們物件偵測的模型。

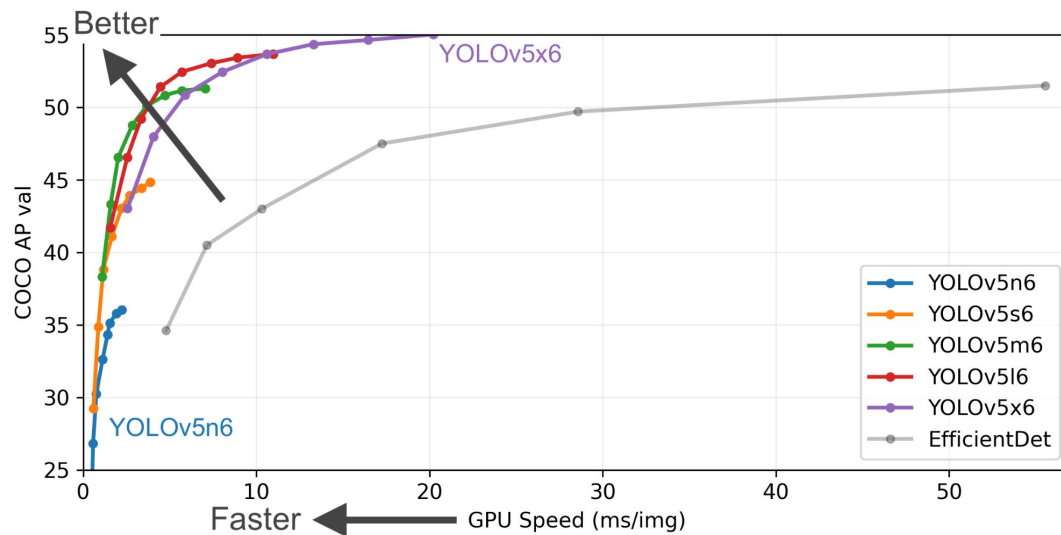


圖 2.為 YOLOv5 中各種模型的準確性與速度比較圖

YOLOv5 有非常多的模型，不同模型的速度跟準確性都會不同(如圖 2.)，如果準確性越高的話速度就相對地比較慢，因此我們所採用的是 s model，因為我們希望速度快一點，避免司機反應時間過短，而失去作用。

同時，我們也試過使用 YOLOv5 去偵測軌道，但當軌道用 bounding box 去切割就會有很大一部份是在軌道外面，會導至假警報的狀況會非常嚴重，因此我們決定要犧牲一點速度去增加我們的準確性，進而將實例分割模型整合到我們的系統之中。

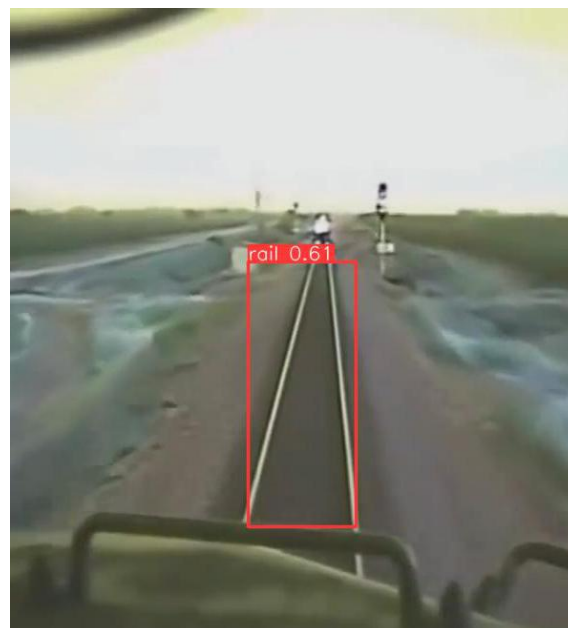


圖 3.為 YOLOv5 所框出來的軌道範圍

3.3 YolactEdge

是 Haotian Liu 在 2020 提出的一個實例分割方法，可以在 RTX2080 上以 FPS170 運行，且 mAP 高達 44 以上(如圖 4.)，在該領域是速度上的佼佼者，因此我們選用了此模型。

Method	Backbone	mask AP	box AP	RTX FPS
Mask R-CNN [32]	R-101-FPN	43.1	47.3	14.1
CenterMask-Lite [2]	V-39-FPN	41.6	45.9	34.4
BlendMask-RT [5]	R-50-FPN	44.0	47.9	49.3
SOLOv2-Light [3]	R-50-FPN	46.3	–	43.9
YOLACT [1]	R-50-FPN	44.7	46.2	59.8
YOLACT [1]	R-101-FPN	47.3	48.9	42.6
Ours				
YolactEdge (w/o TRT)	R-50-FPN	44.2	45.2	67.0
YolactEdge (w/o TRT)	R-101-FPN	46.9	47.8	61.2
YolactEdge	R-50-FPN	44.0	45.1	177.6
YolactEdge	R-101-FPN	46.2	47.1	172.7

圖 4.YolactEdge 模型表現

3.4 異物資料集

輕軌主要穿梭在城鎮道路之間，我們列出以下常見到的可能入侵軌道的異物：汽車、公車、卡車、機車、腳踏車、行人... 等等，由於臺灣的機車族群數量非常多，機車密度為世界第一，因此資料集須包含充裕的機車類別。

Udacity dataset 是由市區街景影片以幀為單位而組成，提供 404,916 幀用於訓練以及 5614 幀用於測試，標籤包含：汽車、行人、卡車、交通號誌等等...，在資料集平臺 Roboflow 上，有基於 *Udacity dataset* 去增加與整理標籤的自駕車資料集，總共包含 11 種類別及 15000 張圖片(如圖 5.)。

Class Balance

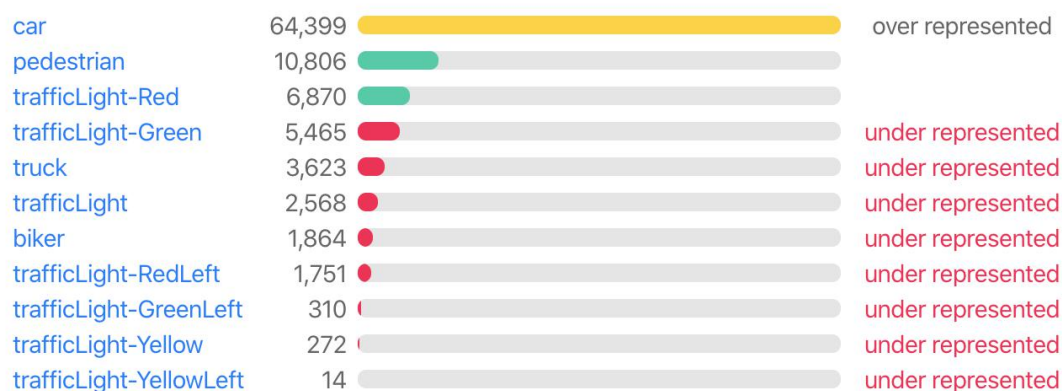


圖 5.Roboflow 上自駕車資料集

BDD100K 是柏克萊大學 2018 年發布的駕駛影像資料集，具有地理、環境和天氣等多樣性，並且帶有時間資訊。每個影像大約 40 秒長、幀率 30fps、解析度為 720p，附有手機記錄的 GPS/IMU 資訊。研究者們在每個影片的第 10 秒取關鍵幀，並提供標註，其中邊界框標註(Bounding Box)為經常出現在道路上的物件進行標註(圖 6.)



圖 6.BDD 資料集 Bounding Box 標註類別

3.5 軌道資料集

Railsem19：關於軌道的資料集我們是從 railsem19 利用 roboflow 去手動以 polygon(多邊形)標示軌道(如圖 7.)，我們總共大概標示了 4207 張圖片，我們在框的時候會稍微超出軌道的原因是因為這樣在做訓練的時候軌道的邊緣會是一個重要的特徵。



圖 7.為我們手動標示的範例

四、評估與實驗結果

4.1 YOLOv5 結果與評論

由於是需要應用在輕軌行駛上，我們在參考架構上主要會以速度為主要考量，因此選用了 YOLOv5，我們實際測試的速度可以達到，100 fps，mAP 也能達到 48.2。(如圖 8.)

```
Speed: 0.7ms pre-process, 5.0ms inference, 2.1ms NMS per image at shape (1, 3, 640, 640)
Results saved to runs/detect/exp23
```

圖 8.此數據為我們時測出來的速度，換算成 fps 大約等於 128 fps

下表為我們分別用不同 dataset 所做出來的 mAP 數據比較，可以發現 BDD100K 的表現大多都優於 Udacity(表 1. 圖 9.)，關於機車、腳踏車的 labels 數量都較少，mAP 的數值雖較不理想，但平均仍優於 Udacity。

Udacity	car	person	truck	biker				
lables	21909	2052	1197	348				
mAp	81.4	39.3	46.7	37.5				
BDD100K	car	person	truck	rider	bike	motor	bus	train
lables	102506	13262	4245	649	1007	452	1597	15
mAp	76.8	58	60.7	41.4	46.6	42.2	59.3	0.7

表 1.Udacity 與 BDD100K 在 YOLOv5 上的表現



圖 9.為我們利用 YOLOv5 所偵測出來的障礙物

4.2.YolactEdge 結果與評論

實測部分，我們因為設備和環境限制，沒有使用 TensorRT 加速，FPS 大約落在 25 到 30 之間(圖 10.)，mAP 也不太理想(表 2.)。

	all	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95
box	17.46	24.42	24.40	24.16	23.77	23.33	21.35	18.75	10.83	3.25	0.34
mask	19.58	24.38	24.00	23.93	23.57	23.28	22.77	22.08	19.69	11.04	1.07

表 2.目前最佳的訓練結果

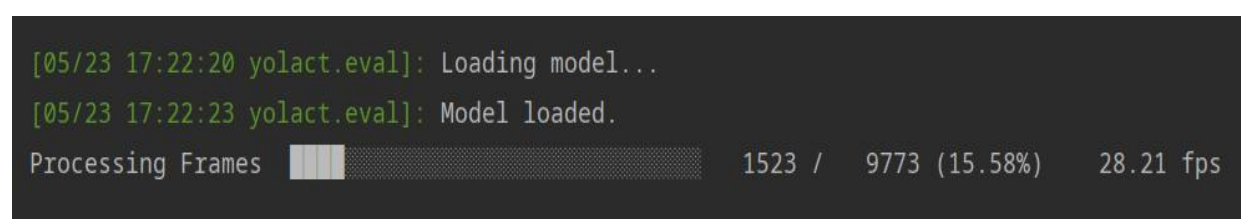


圖 10.影片偵測的速度



圖 11.輸出的圖片

由輸出圖片(圖 11.)可以看出，它極有可能在一些有平行線的地方（尤其是天線）將其誤判為軌道，我們目前的處理方式是，在輸出 mask 之後，將其切出一個中央區域的 ROI 再做障礙物是否重疊的判斷，此方法仍有很大的改善空間。

4.3 成果展示

如果有障礙物在軌道上的話就會將整張圖片變成紅色以示警惕，沒有則是綠色，同時將偵測到的障礙物用紅色框框標出來，而軌道則是用橘色區塊去標示。(如圖 12.)

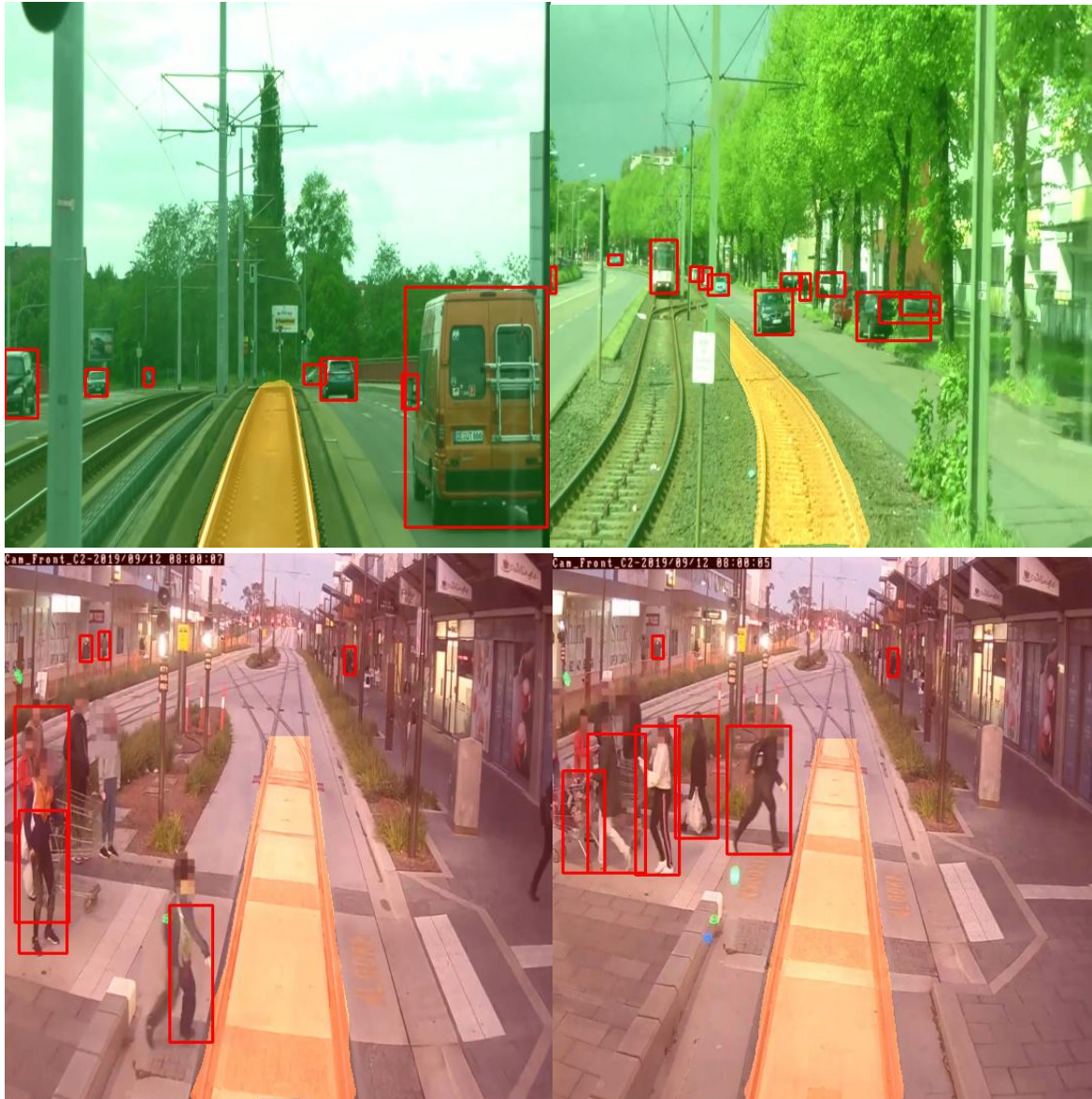


圖 12.成果展示

五、結論

5.1 結語

就觀察結果而言，我們可以將障礙物很好的分辨出來並且可以初步判斷是否有障礙物在軌道上，不過關於軌道的部分精準度還需要再加強，因為仍會有出現一些不該出現的誤判。

而隨著輕軌的普及，如果可以成功的應用這個技術的話，就可以在要撞到東西之前給予駕駛員提醒，進而來減輕駕駛員的負擔，並且也能降低人為疏失的機率，讓整個大眾運輸更安全。

5.2 未來展望

(1)蒐集符合臺灣交通環境的物件偵測資料集：

由於輕軌障礙物偵測相關研究較多以國外的交通環境去做研究，在交通物件偵測的資料集中包含機車族群的比較少，因此我們認為之後可以試著去整合更多針對機車族群的資料集，得出符合臺灣交通環境的物件偵測資料集。

(2)改用效果更好的模型：

本次專題模型架構的使用上，由於我們鐵道分割標註的資料集只有將近 5000 張，訓練出來的影像分割效果沒有預期的好，未來希望改成精度更高且針對少量資料集的影像分割模型，另外 YOLO 系列近期推出 v6 模型，物件偵測的部分也會嘗試使用看看 v6 模型。

5.3 預期困難

目前在鐵軌切割的部分還相當不穩定，以及障礙物資料集仍然缺少台灣的機車、電動代步車等常見交通工具，因此我們會需要自己去蒐集台灣常見的摩托車車種然後自己標註。

另一個是萬一行車速度過快，會導致煞停時間跟距離增加，因此如果只有鏡頭前面的部分可以偵測到的話容易會有來不及的情況發生，因此在軌道末端需要在做的更精確，如果攝像頭的畫素品質不好也容易影響判斷。

最後是特殊天氣的話也容易偵測錯誤，這點我們或許可以參考自駕車是如何解決像是下雨天或是其他等特殊天氣。

五、參考資料

- [1] Ultralytics. YOLOv5
- [2] Haotian Liu, Rafael A. Rivera Soto, Fanyi Xiao, Yong Jae Lee, "YolactEdge: Real-time Instance Segmentation on the Edge," Published in IEEE International Conference on Robotics and Automation (ICRA), 2021
- [3] Fisher Yu, Haofeng Chen, Xin Wang, Wenqi Xian, Yingying Chen, Fangchen Liu, Vashisht Madhavan, Trevor Darrell, "BDD100K: A Diverse Driving Dataset for Heterogeneous Multitask Learning," Published at IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2020
- [4] Oliver Zendel, Markus Murschitz, Marcel Zeilinger, Daniel Steininger, Sara Abbasi, Csaba Beleznai, "RailSem19: A Dataset for Semantic Rail Scene Understanding," Published in IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2019
- [5] Shervin Minaee, Yuri Boykov, Fatih Porikli, Antonio Plaza, Nasser Kehtarnavaz, and Demetri Terzopoulos, "Image Segmentation Using Deep Learning: A Survey," in Proceedings of IEEE Transactions on Pattern Analysis and Machine Intelligence (Volume: 44, Issue: 7, July 1 2022)