

CSIE 5452 Project Report

李詠億, 何志宏, 方郁婷

TOTAL POINTS

26 / 30

QUESTION 1

1 Project Report 26 / 30

+ 30 pts Top 5% projects; significant efforts and contributions

+ 28 pts Top 20% projects; significant efforts and contributions

✓ + 26 pts *Significant efforts and contributions*

+ 24 pts Good efforts and contributions

+ 22 pts Reasonable efforts and contributions

+ 20 pts Basic score

+ 16 pts Only slides

1 Project Report 26 / 30

- + **30 pts** Top 5% projects; significant efforts and contributions
- + **28 pts** Top 20% projects; significant efforts and contributions
- ✓ + **26 pts** *Significant efforts and contributions*
- + **24 pts** Good efforts and contributions
- + **22 pts** Reasonable efforts and contributions
- + **20 pts** Basic score
- + **16 pts** Only slides

速限語音助理

R10922196 方郁婷 R10922115 何志宏 R11944059 李詠億

Abstract

Type: Implementation.

Contributions:

We study speed limit recognition approach, create new kinds of data sets and retrain models, and combine voice functions to become voice assistants.

Statement:

The report does not have any double assignment or any completed work before this semester.

Introduction

對於駕駛者來說，在平面道路駕駛時，除了會需要看四面八方的來車，還要注意紅綠燈，也要注意自己的儀表板。如果是在高速公路上行駛就要注意與前方車輛之間的距離、後方的來車等等。不管是哪一種，都還需要注意目前的速限是多少，一個駕駛要看的東西實在是太多了，這時候如果在車上加裝速限辨識系統，並用語音的方式來提醒目前限速，可以降低駕駛人分心的機率。

對於自駕車來說，如果能夠自動辨識速限標誌的話，就能將速度控制在速限以下，避免超速的危險以及收到罰單的情形。雖然GPS已經有把速限紀錄在地圖資料裡，但這些資料未必是最新的，使用者也不一定有去更新。另一個更嚴重的問題是我們都知道GPS無法準確的辨認車輛現在是行駛在高架道路或是平面道路上，而兩者的速限有一定程度上的落差，這時候車子如果能自動辨識目前的時速，會比GPS的效果來的好。

Proposed Approach

Dataset version 1:

1. 由於網路上資源並沒有以台灣速限標誌為主軸的資料集，且不同的國家速限標誌有不一樣的形式，如圖1所示。故我們產生模擬具速限標誌的台灣街景圖，以進行模型訓練。
2. 資料集是我們整合網路既有資源而產生，我們稱之為台灣速限標誌資料集。先使用傳統影像處理對速限號誌進行處理，其中速限標誌是採用台灣標準速限標誌，是從 *wikimedia*的 *Category: Speed limit road signs in Taiwan* 取得。而傳統影像處理的工作流程如圖2。
3. 使用 Copy-Paste 的方法將處理過的速限標誌圖片貼上我們所找到的街景圖資料集，街景圖資料集是採用 *Street View Dataset* 中的圖片。產生的結果如圖3，並同時產生標記座標，最後再以Roboflow進行資料標記，總共產生4000張圖片。

Dataset version 2:

1. 由於第一版的訓練資料集成效有限，主要原因在於一個街景圖只有貼上一個速限號誌，且一些亮度變化的案例無法被正確判別，還有一個速限標誌若是太邊緣或是太小，也容易造成誤判。
2. 我們在傳統影樣處理的工作流程中做了一些變更，在random size的步驟整體大小再調低，且新增了一個亮度調整的步驟如圖4。並且在同一張街景圖中貼上不只一張的速限標誌，結果如圖5所示。



圖1. Accudriver 是一個可以比較不同國家交通標誌的網站。

左邊為日本速限標誌，右邊則是美國的速限標誌。

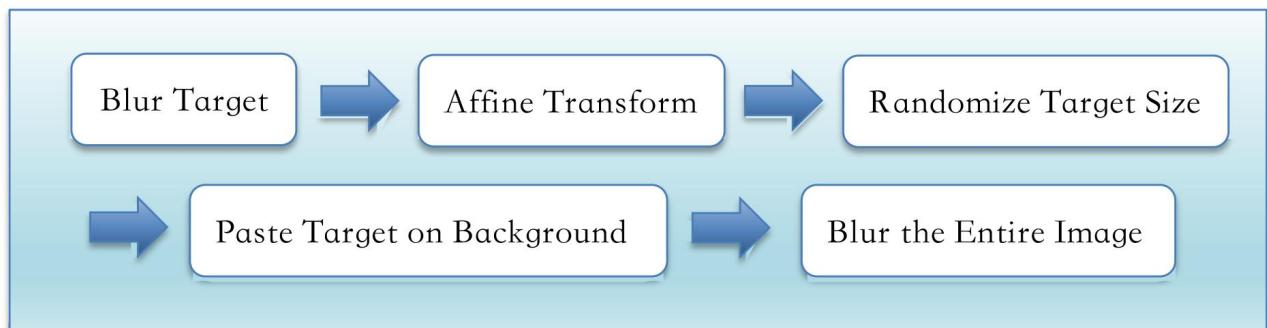


圖2. 對速限標誌進行傳統影像處理工作流程



圖3. Version 1 Dataset 台灣速限標誌資料集

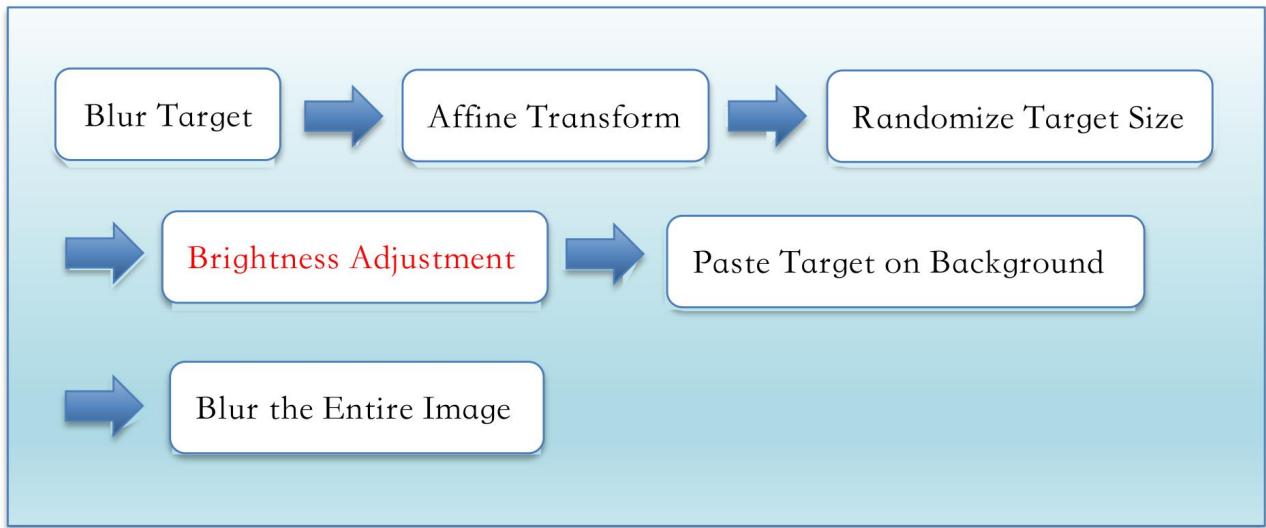


圖4. 對速限標誌進行傳統影像處理工作流程



圖5. Version 2 Dataset 台灣速限標誌資料集

YOLOv5s:

1. 我們選擇 YOLOv5 的原因是因為 YOLOv5 是一個精準且容易使用的開源模型，它是基於 PyTorch 框架，且在靈活性與速度上遠強於 YOLOv4。
2. 如圖4所示，YOLOv5 依照模型大小有不同的model。如圖5所示，橫坐標代表的是處理一張影像所需的時間，而縱座標可以大致理解成模型的效能，我們可以發現 YOLOv5n 和 YOLOv5s 所需的時間最少，但 YOLOv5s 所需時間只比 YOL

Ov5n 多一點，且效能有大幅度的提升，所以我們最後選擇的是 YOLOv5s 這個 model。

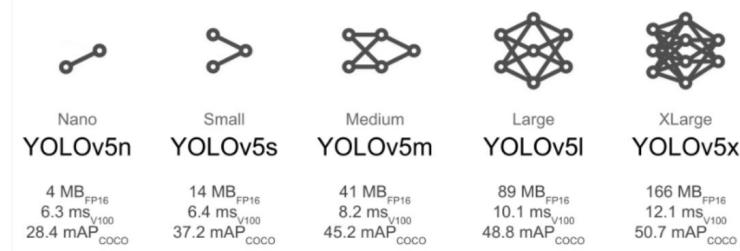


圖6. YOLOv5 Model Description

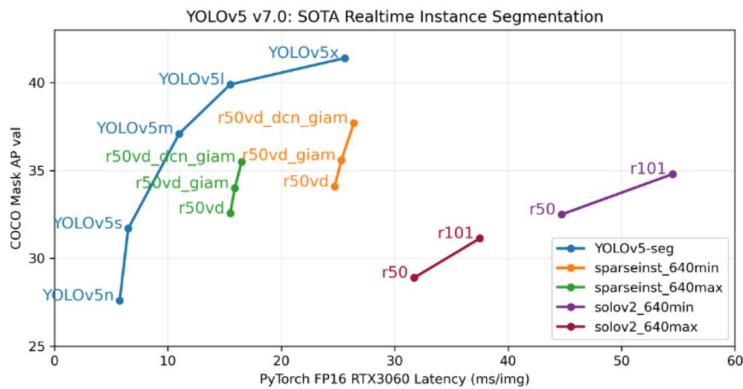


圖7. YOLOv5 Models Performance Chart

(圖6、圖7皆擷取自 <https://github.com/ultralytics/yolov5>)

Pyttsx3:

Pyttsx3 是一款文字到語音的轉換庫，支援多種 TTS 引擎 (sapi5、nsss、espeak)，可以非常方便的將文字轉換成語音，並能調整語速快慢、音量大小，也包含多種語音包可以做音色調整。我們的語音部分是即時產生的，為了避免手機晃動時誤判，語音只有在連續偵測出 5 張照片皆為相同的 label 時才會播出，也就是 Threshold = 5。

Experiment Results

我們的結果分為三部分：即時語音播報、街拍影像辨識、行車紀錄器辨識。

我們會針對兩種資料擴充的方法來對結果做統整與比較。

Dataset version 1: best1.pt

對於使用 Dataset version 1 訓練出來的model，我們命名為best1.pt（可以在github的yolov5資料夾裡找到），Dataset version 1是 將原本的時速標誌做 Scaling、Shiftting、Affine 和 Blur 所產生的資料集，對於原本的標準速限標誌來說只做了型變，以下是我們的結果：

1. 即時語音播報

我們使用筆電的相機鏡頭拍攝，再用 YOLOv5s model 對每張照片做辨識，框出速限標誌的位置，並標出相對應的 label，最後輸出到螢幕上，語音助理會根據偵測到的 label 播報出時速，結果如圖8所示，因為是使用標準時速標誌做測試，所以辨識正確率高達99.9%，幾乎不會有辨識錯誤的情況發生。



圖8. 即時語音播報結果圖

2. 街拍影像辨識

我們從網路上蒐集許多由台灣拍攝的速限標誌圖（通常是新聞拍攝），用來測試YOLOv5s model的效果，圖9為辨識成功案例，圖10為辨識失敗案例。成功的案例有點超出我們的預期，因為我們的訓練資料集並非拍攝現實環境，而是用標準速限標誌的圖片合成的。因此我們的best1.pt對於光照、霧、雜訊等現實環境具有robust的特性對我們來說是意外之喜，也許是我們將影像resize時順便將這些影響因子

降低了。失敗的案例顯示仍有許多圖片沒能辨識正確或有誤判，但大多數都至少有框出標誌的位置。



圖9. 街拍影像辨識成功案例



圖10. 街拍影像辨識失敗案例

3. 行車紀錄器辨識

汽車在行進過程中可能因速度過快，會出現在不同時間點判斷結果不一的情形，無法有很好的穩定性，結果如圖11。



圖11. 行車記錄器影片辨識

Dataset version 2: best2.pt

對於使用Dataset version 2訓練出來的model，我們命名為best2.pt（可以在github的yolov5資料夾裡找到），Dataset version 2是對Dataset做了更多的更動，除了原本的形變以外，還使用了類似 Mosaic data augmentation 這種在一張圖片放入多個標誌與 label 的技術。為了貼近真實情況，我們將標誌 scale 得更小，亮度變得更暗，以下是我們的結果：

1. 即時語音播報

如圖12所示，辨識結果與前一種方法有很大的落差，因為兩種資料集皆為4000張圖片，我們認為增加資料多樣性的同時可能造成單一種類資料變少，導致辨識的錯誤率很高。

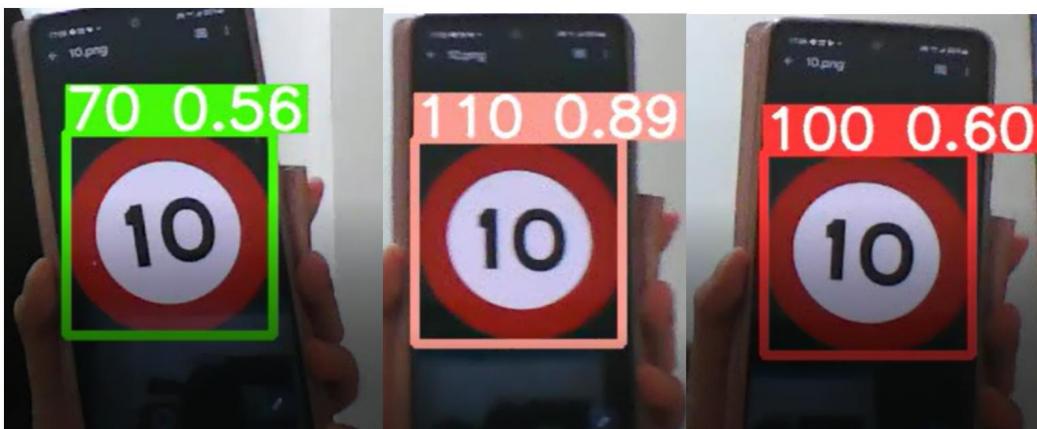


圖12. 即時語音播報（方法二）結果圖

2. 街拍影像辨識

大部分的結果都沒有 best1.pt 好，但還是有一些比 best1.pt 好的 case。如圖13所示，左邊為 best1.pt 的結果，錯誤或無法辨識。右邊則為 best2.pt 的結果，正確或至少能辨識 best1.pt 無法辨識的 case。



圖13. 街拍影像辨識（方法二）結果圖

3. 行車紀錄器辨識

在某些case能更快的偵測出結果，其餘結果差異不大（一樣有判斷結果不一的情形）。圖14（左）為best1.pt的結果，在第2秒時偵測到；圖14（右）為best2.pt的結果，在第1秒時偵測到。

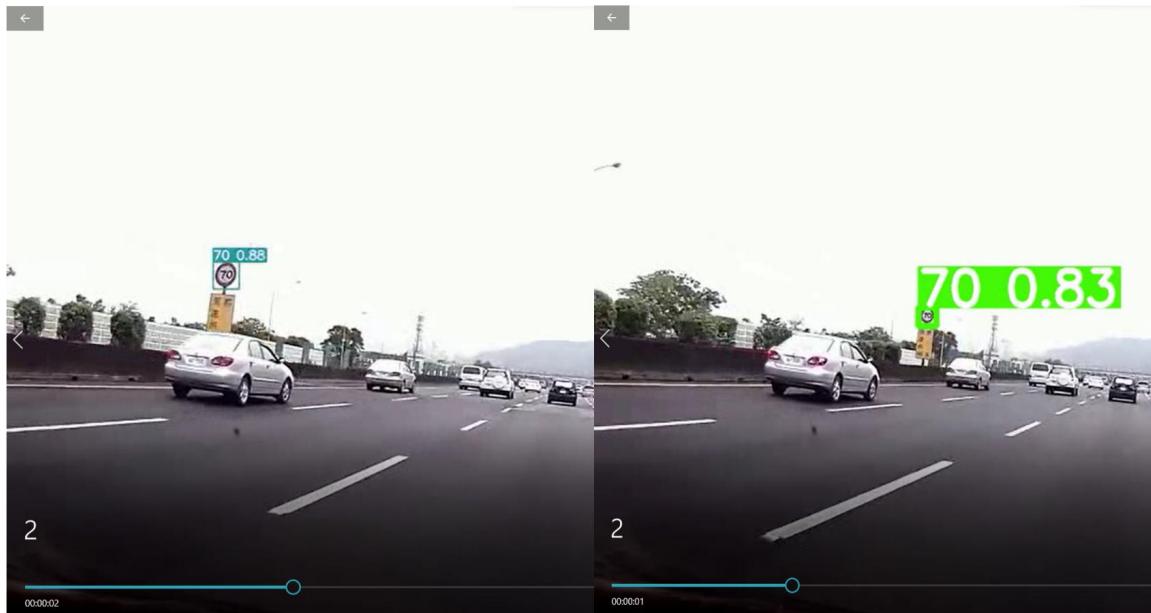


圖14. 行車記錄器辨識（方法二）結果圖

	即時語音播報	街拍影像辨識	街拍影像辨識
best1.pt	正確率: 99.9%	正確率: 64%	判斷結果不一
best2.pt	正確率: 10%	正確率: 46% 但對於某些best1.pt 判斷錯誤的case能 判斷正確	判斷結果不一，但 能在距離更遠的地 方開始判斷。

圖15. best1.pt與best2.pt比較表

Conclusions

1. We built a dataset of speed limit signs of Taiwan.
2. We use the YOLOv5s model to identify Taiwan speed limit signs.
3. Remind the driver of the recognition result by voice.

綜上所述，best2.pt確實有往我們希望的方向學習，比如說在標誌比較小或標誌比較暗的case下較能判斷正確。但或許是訓練資料的變動太大，與實際情況相差太多，效果反而不好。對於訓練資料的產生方式與訓練技巧還有很多可以繼續改進的地方。

Contributions of each member

方郁婷：模型及語音系統整合至實體應用

何志宏：產生訓練資料與模型訓練

李詠億：協助各項工作和整合成果與發表

Source Code

Traffic-Sign-Detection-Using-Yolov5

<https://github.com/YuTing-Fang1999/Traffic-Sign-Detection-Using-Yolov5>

Reference

https://commons.wikimedia.org/wiki/Category:Speed_limit_road_signs_in_Taiwan

<https://github.com/ultralytics/yolov5>

<https://app.roboflow.com>

http://www.cs.ucf.edu/~aroshan/index_files/Dataset_PitOrlManh/

https://pytorch.org/hub/ultralytics_yolov5/#:~:text=YOLOv5%20%F0%9F%9A%80%20is%20a%20family,size

Jocher, G. (2020). YOLOv5 by Ultralytics (Version 7.0) [Computer software]. <https://doi.org/10.5281/zenodo.3908559>