

## 應用 yolo 深度學習模型於影像辨識

專題編號：107-CSIE-S011

執行期限：106年第1學期至107年第1學期

指導教授：郭忠義

專題參與人員：104590452(學號) 李宇傑(姓名)

### 一、摘要

本題目主要的目的是對於現在使用於影像辨識主流之一的 CNN 神經網絡架構中，可以進行即時辨識的 yolo 網絡進行學習，及實作，在影像辨識的需求上目前應用上都是想追求以 real stream 來即時辨識，而是否可以達成在即時的反應下，仍有良好的辨識效果的關鍵就在於訓練時的樣本及參數的調整所產出的權重，因此在開始訓練前有必要了解 CNN 整體網絡架構，以及它進步的過程中採取甚麼技術，才能明白樣本的製作方法及其網絡架構中各參數意義，才有能力在實作中訓練出符合需求的權重

**關鍵詞：**深度學習，神經網絡，  
CNN，Yolo

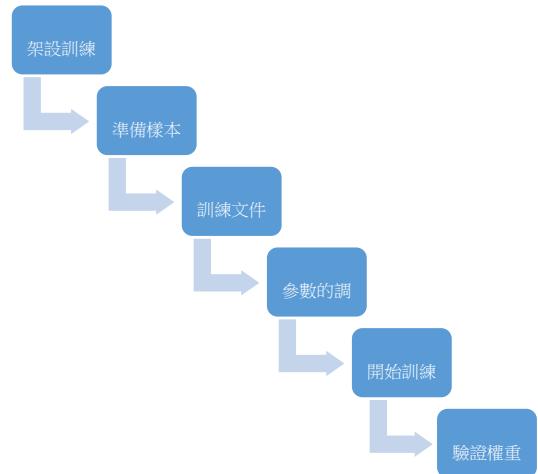
### 二、緣由與目的，

在資策會實習時，上司所交付的小組共同專案是與影像辨識相關的技術與應用，在工作時期間內開始對這個領域有了更多的了解與理解到這是一門值得了解且學習的技術，如果說未來這樣的技術很有可能會被廣泛地的使用，那對於現在來講就是學習與實作的最好理由

### 三、研究報告內容

首先從 CNN 網絡的架構了解深度學習中神經網絡的運作原理與內容，清楚明白其神經網絡的架構並了解樣本的 input 過程中的轉換與 output 的關係後，接者是與 CNN 網絡結合的相關技術及產物的理解，如 RCNN，Fast R-CNN，Faster R-CNN，最後是對於以上相關技術的結合與最後實際使用的神經網絡 yolo 的學習，要將神經網絡訓練成果運用於影像辨識上，

可以分為以下幾個步驟



#### 1. 架設訓練環境

在 Widows 上是藉由 visual Studio 進行建置後產生出 .exe 執行檔後，使用 cmd 進行操作

## 2. 準備樣本

準備自己的樣本，首先要收集需要的圖片資料，收集完成後需要以標記來標示自己想要訓練的辨識類別，標記完後會產生有標記圖片中想辨識物件資訊的 xml 檔案

## 3. 訓練文件的準備

在開始訓練前需要準備訓練中會使用到的文件，包括由 xml 轉為 Yolo 網絡所使用的 txt 檔案，樣本絕對路徑的 txt 檔案樣本名稱的 txt 檔案，驗證圖片如同上方兩種的 txt 檔案這些都準備完成後即可進入下一階段

## 4. 進行參數上的調整

選定要使用甚麼樣的架構後即可開始進行各種參數的調整，調整的依據在於樣本的數量與想辨識的類別等等，細部的參數調整則是根據學習中的輸出項目而變化

## 5. 開始訓練

準備就緒後，即可開始進行訓練一次的訓練根據樣本的數量及所選取的網絡層數而定，以我自己訓練中的經驗而言以五十，五十的數量訓練及驗證，至少要訓練四個至五個小時才能產生出可以使用的權重

## 6. 對權重進行效果驗證

進行訓練後的權重可以以圖片或是影片的方式來進行測試，藉由觀察測試中目標的辨識效果來判斷權重的使用可否使用

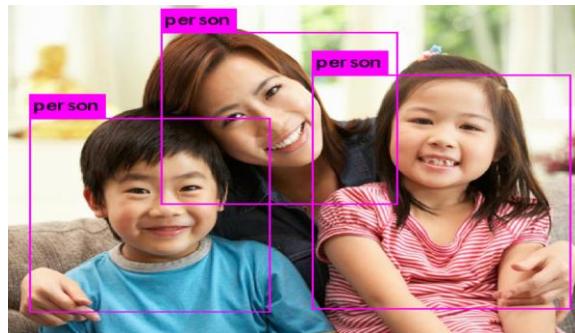
## 四、實驗結果

□ yolov3-voc_100.weights	2018/10/11 下午 ...	WEIGHTS 檔案	240,533 KB
□ yolov3-voc_200.weights	2018/10/11 下午 ...	WEIGHTS 檔案	240,533 KB
□ yolov3-voc_300.weights	2018/10/11 下午 ...	WEIGHTS 檔案	240,533 KB
□ yolov3-voc_final.weights	2018/10/11 下午 ...	WEIGHTS 檔案	240,533 KB

圖一. training 後權重結果

```
Total BFLOPS 65.290
Loading weights from yolov3-voc_final.weights...
| seen 64
Done!
people97.jpg: Predicted in 68.752000 milli-seconds.
person: 99%
person: 88%
person: 95%
```

圖二. 驗證權重時 cmd 輸出



圖三. 權重驗證結果

經由訓練後得到的權重，藉由驗證效果後即可投入到 Real stream 上進一步檢視及時辨識的效果

參考文獻:

[1] [Joseph Redmon](#): You Only Look

Once: Unified, Real-Time Object Detection.

<https://pjreddie.com/darknet/yolo/>