

影像處理期中報告心得：使用 YOLOv4 進行鐘面時間影像識別

數三甲 S1022030 張云亭

從獲知這項課題開始，我就在思索如何把課堂上學習到的影像處理內容，結合我數學系身分在數學領域的相關知識，製作出一份具有挑戰性和關聯性的主題。在瀏覽 Kaggle 網站上的公開資料集後，我很迅速地鎖定了影像識別鐘面指針的課題方向。

在深度學習模型方面，我選擇使用助教在課堂所教導的 YOLOv4-tiny。YOLO (You Only Look Once) 是目前輕量化目標檢測網路效果最好的結構之一，而 YOLOv4 版本是一個具備更為簡單、快速、高效等優點的深度學習模型。在單純從乾淨的鐘面影像資料集中識別出指針時間結果上，YOLOv4-tiny Darknet 是一個非常適宜的選擇。

在觀察我所選擇的資料集後，能發現內部圖檔處理的還是比較完整的，已經清晰的裁切好鐘面影像，所以我意識到，少了耗時且繁瑣的手動標記圖像過程，我能使用 224 X 224 X 3 張圖像在 train 上，以及 224 X 224 X 3 張圖像在 valid 上。因此，我所面臨到的第一個難題，即撰寫一個腳本來批量產出 YOLO 所需的圖像標籤。運用迴圈與檔案路徑名，我以三層迴圈架構完成了共 224 X 224 X 6 個圖像標籤 TXT 檔。

然而，我迅速面臨了第二個難題。YOLOv4 的設置和訓練需要高性能的 GPU，而我只有有限的計算資源，這迫使我不得不使用雲端計算資源，利用 google colab 似乎能解決這項難題，但我希望能識別出共 144 項類別，在不購買 colab pro 的條件下，我最終成功的取得了共 23890 筆訓練，雖然距離原先期望的 288000 筆有段距離，但在乾淨的資料集訓練下，識別正確率是非常高的。

雖然在這項課題中遇到了許多困難，但我也在逐步解題中學習到了許多有價值的知識。首先，我學會了如何設計和訓練一個 YOLOv4 模型。這包括選擇適當的超參數、設置訓練流程以及優化性能。我也學到了如何使用不同的數據擴增技術來改進模型的泛化性能。

另外，我還深入瞭解了深度學習模型的工作原理，包括卷積神經網絡 CNN 的運作方式和 YOLO 的特點。這些知識對於深度學習和影像處理領域的未來研究和應用非常有幫助。

在翻閱別人的訓練模型後，我也學習到在這項課題後，加入霍夫變換即能識別出更為精細的時間。先以目標檢測鐘面定位，再加入霍夫變換等特徵識別技術識別出指針，運用角度差即可獲得更為細緻的結果。

總結來說，這次的報告讓我獲益匪淺。我克服了挑戰，學到了許多新知識和技能，並在影像處理和深度學習領域取得了一些成就。這個經驗對我的學術和興趣鼓舞都有所幫助，我期待著未來繼續深入研究和學習影像處理。