

國立雲林科技大學

資訊管理系研究所

機器學習

專案作業三

M11223033 鍾季衡

M11223036 魏冠宇

M11223047 范棣雲

指導教師：許中川

中華民國 113 年 5 月

摘要

本研究旨在探討利用物件偵測和文字辨識技術對貨櫃號碼進行自動辨識的方法。使用了貨櫃號碼資料集進行模型訓練，並將其應用於圖片和影片資料集中，以檢測貨櫃號碼並進行文字辨識。通過探索不同的深度學習模型，評估了各個模型在貨櫃號碼辨識上的表現。實驗結果表明，本研究方法能夠有效地檢測和識別貨櫃號碼，並且在不同模型之間進行性能比較。這項研究對於提高貨櫃管理效率和安全性具有重要意義，尤其是在貨物運輸和物流領域。

關鍵字:YOLOV5、R-CNN、貨櫃號碼、自動辨識

第一章 緒論

第一節 研究動機

貨櫃號碼在國際貿易和物流運輸中扮演著重要的角色，但目前對於貨櫃號碼的辨識仍然主要依賴人工操作，存在著效率低下和容易出錯的問題。因此，利用物件偵測和文字辨識技術開發自動化的貨櫃號碼辨識系統具有重要意義。透過這樣的系統，可以提高貨櫃管理的效率和準確性，減少人為錯誤的風險，同時加速貨物流通，促進貿易發展。

第二節 研究目的

本研究旨在開發一種基於物件偵測和文字辨識的自動化貨櫃號碼辨識系統，以提高貨櫃管理的效率和安全性。具體目標包括：

1. 探索不同的深度學習模型，如 YOLO、Faster R-CNN 等，尋找最適合貨櫃號碼辨識任務的模型。
2. 建立訓練集、驗證集和測試集，進行模型的訓練、驗證和測試。
3. 評估不同模型在貨櫃號碼辨識上的性能，並進行性能比較。
4. 研究影片資料集的貨櫃號碼辨識方法，實現對影片內貨櫃號碼的自動辨識和文字提取。
5. 最終構建一個準確、高效的貨櫃號碼辨識系統，為物流運輸行業提供更好的服務。

第二章 方法

第一節 程式架構

研究採用了物件偵測和文字辨識技術相結合的方法來自動識別貨櫃號碼。在程式架構中，首先使用了先進的物件偵測模型，如 YOLO、Faster R-CNN，來定位圖片中的貨櫃號碼位置。這些模型通常基於深度學習技術，具有良好的準確度和泛化能力。接著，將定位出的貨櫃號碼區域送入文字辨識模型進行識別，文字辨識模型是基於深度學習的端到端模型，為傳統的光學字符識別（OCR）模型，用於將圖片中的文字轉換成可識別的字元序列。最後為提高識別準確度，引入了貨櫃號碼公式判斷模塊，根據貨櫃號碼的前 10 位計算出第 11 位，以進一步校正識別結果。

第二節 執行方法

首先安裝所需的執行環境、深度學習框架、圖像處理庫和文字辨識庫。接著使用 Python 程式語言編寫程式，根據程式架構中的各部分，按照步驟依次執行物件偵測、文字辨識和號碼判斷等功能，並根據需求進行訓練或使用預訓練模型。再將訓練集、驗證集和測試集導入到程式中，用於模型的訓練和測試。最後將影片資料集讀取並進行每幀圖片的處理，使用物件偵測模型和文字辨識模型對圖片進行處理和分析，最終得到辨識結果。

第三章 實驗

第一節 資料集

1. 貨櫃號碼資料集 (Container Number Dataset) :

- 這個資料集包含了訓練、驗證和測試三個部分的圖片集，以及每個圖片對應的標註檔 (XML)。
- 資料集共有六個資料夾，分別是「訓練集 (train)」、「驗證集 (val)」和「測試集 (test)」的圖片和標註檔。
- 這個資料集的用途是訓練物件偵測模型，以便辨識圖片中的貨櫃號碼。
- 每張圖片都被標註了貨櫃號碼的位置和相關信息，這些信息可用於模型的監督式學習。

2. 貨櫃影片資料集 (Container Video Dataset) :

- 包含測試用的 10 部影片和 1 部範例影片。
- 資料集的用途是對影片進行處理，並使用物件偵測模型和文字辨識模型對圖片進行分析和處理。
- 影片中的每一幀圖片都將被提取出來，然後進行貨櫃號碼的偵測和辨識。
- 通過將影片轉換為圖片序列，可以對影片中的貨櫃號碼進行更準確和有效的辨識。

第二節 前置處理

YOLOv5 模型中，本研究在訓練時先將圖片大小調整為 640*640，使影片和測試集中讀取地每幀圖片調整為適當大小，並使用 `imgaug.augmenters` 進行多種資料增強，例如水平翻轉 50% 的圖片、隨機剪裁、調整對比度、調整亮度、隨機旋轉和剪切，以及加入高斯雜訊。此外，於測試檔案中，使用 `ImageEnhance.Contrast` 增加圖片對比度，並使用 `ImageOps.grayscale` 將圖片轉換為灰度圖片，有助於提高後續 OCR 的辨識準確性。本研究將灰度圖片進行二值化處理，以便於 Tesseract OCR 能更好地辨別文字。

RCNN 模型中，調整圖片大小，確保輸入到模型中的圖片尺寸一致，以便於模型處理和特徵提取，圖片從 RGB 轉換為灰度圖片，減少計算量，只保留亮度資訊，增加文字辨別準確性，以及將圖片對比度提高，使文字區域與背景區別，再將灰度圖片進行二值化處理，將低於某個閾值的像素設為黑色，高於閾值的像素設為白色，對 OCR 的精確辨識有所幫助。

第三節 實驗設計

YOLOv5 模型進行先前前置處理所描述的資料增強和預處理，參數設定分別為 `epochs=200`、`batch size=32`、`img size=640`、`patience=30`，訓練過程中使用上述設定的參數，透過多輪訓練優化模型性能，進行前向傳播、計算損失、反向傳播和參數更新，每輪訓練結束後，紀錄損失值並進行早停檢查，訓練完成後進行驗證並提取評估指標：`mAP`、`Recall`、`Precision`、`F1-score`，以及 OCR 辨識對每幀圖片中偵測到的物件進行裁剪和預處理，使用 Tesseract 進行文字辨識。

RCNN 模型進行先前描述之前置處理，提高模型的泛化能力，防止過擬合。參數設定分別為 epochs=1、Iteration=1050、batch size=16、learning rate=0.007，在每個訓練周期結束後，使用驗證集進行評估，計算 mAP、Recall、Precision、F1-score 指標，了解模型在不同評估標準下的性能。將最優模型權重保存到文件中，以便後續測試和應用。

第四節 實驗結果

YOLOv5 衡量模型在不同 IoU 閾值下的平均精度為 0.9235，能夠正確偵測到的目標佔所有目標的比例為 1.000，能偵測出的目標中正確的比例為 1.000，以及最後的 F1-score 為 1.000。RCNN 衡量模型在不同 IoU 閾值下的平均精度為 1.000，能夠正確偵測到的目標佔所有目標的比例為 1.000，能偵測出的目標中正確的比例為 0.9949，以及最後的 F1-score 為 0.9974。在訓練過程中，紀錄每輪訓練的損失值，監控模型的收斂情況，達到令人滿意的 mAP、Recall、Precision、F1-score，顯示出模型在物件偵測任務中的高性能，使模型在不同環境和光照條件下依然保持穩定的偵測辨識。

圖 1 模型各項指標績效

| | mAP | Recall | Precision | F1-score |
|--------|--------|--------|-----------|----------|
| YOLOv5 | 0.9235 | 1.000 | 1.000 | 1.000 |
| RCNN | 1.0000 | 1.000 | 0.9949 | 0.9974 |

使用 Tesseract 進行 OCR 辨識，對偵測到的物件進行文字辨識，並計算 OCR 準確率，這些步驟包括對每幀圖片進行對比度增強和二值化處理，以提高 OCR 的辨識準確性，結果顯示大部分圖片與影片中的貨櫃號碼能準確被辨識出，少數情況下，由於圖片、影片畫面品質和光照條件影響，會出現部分文字辨識錯誤，但模型仍然能準確地進行物件偵測和文字辨識。

圖 2 模型之 OCR 辨識準確率

| | Image_test_OCR | Video_test_OCR |
|--------|----------------|----------------|
| YOLOv5 | 0.87 | 0.90 |
| RCNN | 0.83 | 0.90 |

第四章 結論

綜合以上研究結果，本研究成功展示了基於物件偵測和文字辨識技術的貨櫃號碼自動辨識系統的有效性和實用性。透過對貨櫃號碼資料集的訓練和測試，建立了一套高效的辨識系統，能夠在圖片和影片資料中準確地偵測和識別貨櫃號碼。實驗結果驗證了該系統在提高貨櫃管理效率和確保貨物運輸安全方面的重要作用。未來不僅可以減少了人力成本，還提高了辨識的準確性和速度，為物流行業帶來了巨大的便利和效益。未來，將繼續優化模型性能，探索更多先進的圖像處理和文字辨識技術，以滿足不斷發展的實際應用需求，並推動自動辨識技術在物流管理領域的廣泛應用。

參考文獻

- 伊凡 . (2021, June 17). [論文筆記]Faster R-CNN. 聰明的伊万.
<https://ivanfang.coderbridge.io/2021/06/17/fang-faster-rcnn/>
- 凌龍英. (2021, June 17). Fast R-CNN 學習筆記. Medium.
<https://didi3310781.medium.com/fast-r-cnn-%E5%AD%B8%E7%BF%92%E7%AD%86%E8%A8%98-dceb2e496b9b>
- 黃湯米. (2018, September 4). 深度學習-物件偵測:You Only Look Once (YOLO). Medium. <https://chih-sheng-huang821.medium.com/%E6%B7%B1%E5%BA%A6%E5%AD%B8%E7%BF%92-%E7%89%A9%E4%BB%B6%E5%81%B5%E6%B8%AC-you-only-look-once-yolo-4fb9cf49453c>
- 蔡承翰. (2021, August 19). YOLO-V5 使用教學. HackMD.
<https://hackmd.io/@4XoSjtMaS46Zzn7DwmEIEQ/SkC0ceHIF>
- 莫里斯. (2021, October 10). YOLO V5—解釋與揭秘（目標偵測）. Medium.
<https://medium.com/image-processing-and-ml-note/yolo-v5-explained-and-demystified-22f041ad316e>
- Austin70915. (2023, September 19). 【Day 4】Pytorch & TorchText 的正確開啟方式. IT 邦幫忙. <https://ithelp.ithome.com.tw/articles/10322104?sc=rss.iron>