

國立雲林科技大學

資訊管理研究所

機器學習專案作業三

指導教授： 許中川 博士

學生：M11223038 陳品佑

M11223032 張祥恩

中華民國 113 年 5 月

April 2024

## 目錄

摘要.....	1
Abstract.....	2
一、緒論.....	3
1.1 研究動機.....	3
1.2 研究目的.....	3
二、研究方法.....	3
三、研究實驗.....	4
3.1 資料集簡介.....	4
3.2 前置處理.....	4
3.3 實驗設計.....	5
3.3.1 績效評估指標.....	5
3.4 實驗結果.....	6
3.4.1 YOLOv5 Model .....	6
3.4.2 Faster R-CNN Model.....	8
3.4.3 光學文字辨識.....	9
四、結論.....	10
五、參考文獻.....	11

## 表目錄

表 1	YOLOv5 Model Training Parameters.....	5
表 2	Faster R-CNN Training Parameters .....	5
表 3	YOLOv5 訓練/測試集之績效表 .....	6
表 4	R-CNN 訓練/測試之資料表 .....	8
表 5	OCR 平均辨識準確度表 .....	9

## 圖目錄

圖 1 貨櫃資料集示意圖.....	4
圖 2 實驗流程圖.....	5
圖 4 訓練集 Precision 績效圖.....	7
圖 5 訓練集 F1-score 績效圖.....	7
圖 3 訓練集績效圖.....	7
圖 6 Recall Curve 圖.....	8
圖 7 貨櫃號碼區塊成功辨識示意圖.....	8
圖 8 影片資料集辨識成功示意圖.....	9
圖 9 成功辨識之過程圖.....	10

## 摘要

---

全球化的背景下，國際貿易和物流業的重要性日益增強，貨櫃運輸作為其中的關鍵一環，對物流管理的效率和安全性至關重要。然而，目前大多數貨櫃號碼的辨識仍依賴人工操作，效率低且易出錯，尤其在大型港口和物流中心，這樣的低效和高錯誤率可能帶來巨大的經濟損失和管理挑戰。因此，本研究旨在利用先進深度學習的物件偵測技術，實作一個高效率且準確度高的貨櫃號碼自動辨識系統。透過訓練 YOLOv5 和 Faster R-CNN 模型進行貨櫃號碼定位和檢測，並結合 OCR 技術進行文字辨識。

實驗結果顯示，兩種模型在物件偵測上均取得了超過 0.90 的 mAP 值，而在 OCR 辨識中均達到 82% 的準確率，參數優化和資料增強等方式或許還能再提升辨識的準確度。本研究實作之系統在物流管理中具有重要的應用價值和提升空間。

**關鍵字：**物件偵測、YOLO、R-CNN、文字辨識

---



## Abstract

---

In the context of globalization, the importance of international trade and logistics industries is increasingly prominent. Container transportation, as a crucial part of these industries, is essential for the efficiency and security of logistics management. However, the majority of container number recognition still relies on manual operations, which are inefficient and prone to errors. This is particularly problematic in large ports and logistics centers, where such inefficiency and high error rates can lead to significant economic losses and management challenges. Therefore, this study aims to implement a high-efficiency and highly accurate automated container number recognition system using advanced deep learning object detection techniques. By training YOLOv5 and Faster R-CNN models for container number localization and detection, and combining OCR technology for text recognition, we aim to achieve this goal.

The experimental results show that both models achieved mAP values of over 0.90 in object detection and an accuracy of 82% in OCR recognition. There is potential for further improvement in recognition accuracy through parameter optimization and data augmentation. The system implemented in this study has significant application value and room for improvement in logistics management.

**Keywords:** Object Detection, YOLO, R-CNN, Text Recognition

---

## 一、緒論

### 1.1 研究動機

在全球化的背景下，國際貿易和物流業已是愈來愈重要，而貨櫃運輸是其中關鍵的一環。而貨櫃號碼是物流管理中的重要標籤，能夠準確高效地辨識貨櫃號碼，對於提升物流運營效率和確保貨物安全具有重要意義。然而，目前多數貨櫃號碼的辨識仍然仰賴人工操作，不僅效率低下，還容易發生人為疏失，尤其在大型港口和物流中心，這樣低效率和錯誤率的方式可能帶來巨大經濟損失以及管理上的挑戰。因此，基於人工智慧的辨識技術，不僅有助於提高物流系統的自動化程度，還能顯著減少人力成本，降低錯誤率，提升整體運營效率。

### 1.2 研究目的

隨著深度學習技術的快速發展，利用先進的物件偵測和文字辨識技術來實現貨櫃號碼的自動化辨識，成為一個具備重要實際應用價值的研究方向。本研究旨在利用先進的深度學習技術，開發一個高效率且準確的貨櫃號碼自動辨識系統，以解決目前物流行業中人工辨識效率低下和錯誤率高的問題。主要研究目的包括以下方面：

1. **訓練並測試不同的物件偵測模型：**利用 YOLOv5、R-CNN 物件偵測模型對貨櫃號碼進行定位和檢測，並比較其在準確率和效率方面的表現。
2. **文字辨識技術的應用：**利用 OCR 技術對檢測到的貨櫃號碼區域進行文字識別，確保能夠從圖像中準確提取貨櫃號碼。
3. **性能評估：**透過衡量指標如 mAP、Recall、Precision 和 F1-score 對模型進行評估。
4. **應用於動態影像：**除了靜態影像之外，本研究將模型應用至貨櫃影像的號碼辨識上，實現即時影像的貨櫃號碼自動辨識。

## 二、研究方法

於本研究中，首先選用 PyTorch 深度學習框架進行模型訓練和測試，接著將專案作業之貨櫃號碼資料集下載(含訓練集、測試集、驗證集)，每個集合都有數百張標註有貨櫃號碼的圖像以及對應的 XML 標註檔案，將這些圖像標註檔案轉換為模型可讀取之格式。而本研究選擇 yolov5、Faster R-CNN 進行模型訓練，而後利用衡量指標做性能評估。而於物件偵測後，使用 Tesseract-OCR 對檢測到的貨櫃號碼區域進行文字識別，確保準確提取貨櫃號碼。

### 三、研究實驗

#### 3.1 資料集簡介

本研究所使用的資料集為專案作業提供之真實貨櫃影像資料集，資料集當中之影像包含光源足夠、清晰的照片，也有陰暗、不清晰的影像。檔案中包含訓練集、測試集、驗證集以及各別之 xml 標註檔。而在影片資料及部分共有 10 部影片來做實驗測試。下圖 1 為部分資料內容示意圖。

- 資料集名稱：貨櫃資料集
- 訓練資料筆數：2125 筆資料
- 測試資料筆數：755 筆資料
- 驗證資料筆數：536 筆資料
- 影片資料筆數：10 筆資料

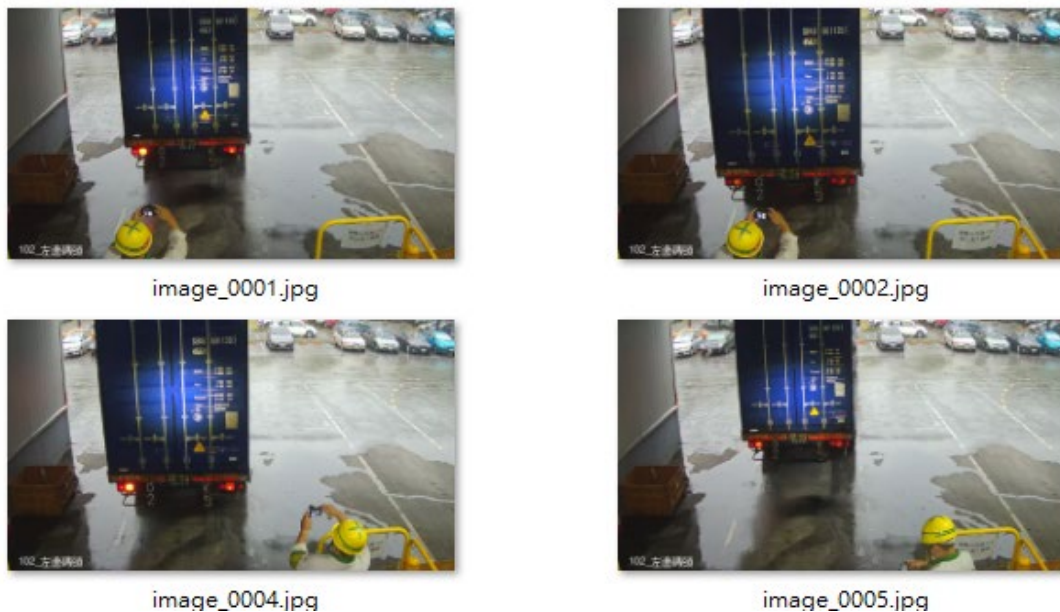


圖 1 貨櫃資料集示意圖

#### 3.2 前置處理

在本研究中，資料集之前置處理首先於資料集中載入影像以及對應的標註文件，而後修正標籤文件中的坐標值、資料增強，並將這些影像調整模型輸入之大小，本實驗採用(640\*640)。資料增強部分則包含對影像做水平翻轉、隨機剪裁和翻轉、調整對比度和亮度。如此即可完成模型之前置處理。



### 3.3 實驗設計

本章節將使用 yolov5 以及 Faster R-CNN 模型進行實驗，兩大模型各有其優勢，本研究主要使用這兩個模型做訓練以比較其績效表現。在圖像資料的實驗流程如下圖 2 所示。而在圖像資料預測上使用的參數訓練如下表 1, 2 所示。

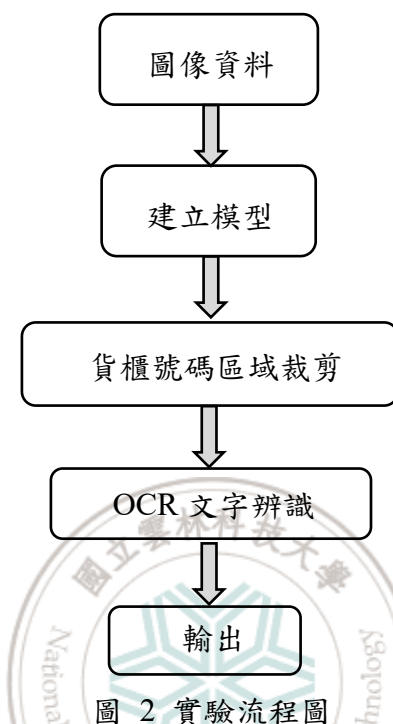


表 1 YOLOv5 Model Training Parameters

Model	epoch	Batch_size	Patience
YOLOv5	100	16	10

表 2 Faster R-CNN Training Parameters

Model	epoch	Learning Rate	Weight Decay	Momentum
Faster R-CNN	1	0.005	0.0005	0.9

#### 3.3.1 績效評估指標

本研究主要使用四種主要的績效評估指標來衡量模型的效果：平均精度均值（mAP）、召回率（Recall）、精確率（Precision）和 F1-score；計算 OCR 準確率則由 Accuracy 作為衡量指標。這些將在本次實驗做為評估模型表現的指標，確保其在貨櫃號碼辨識中的有效性和準確性。

$$Precision = \frac{TP}{(TP + FP)}$$

$$Recall = \frac{TP}{(TP + FN)}$$

$$F1\_Score = \frac{2}{(\frac{1}{Precision} + \frac{1}{Recall})}$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

### 3.4 實驗結果

在經過兩者演算法模型的訓練、測試結果後，在物件偵測上測試集得到的績效至少都有在 0.90 以上，而在圖片資料集之 OCR 辨識率上 YOLOv5 以及 R-CNN 皆為 0.82，由實驗結果可知，在參數優化改進、資料增強等等方面，準確率也許有進一步的提升空間。實驗結果如下所示。

#### 3.4.1 YOLOv5 Model

本實驗之 YOLOv5 模型在物件偵測上之績效表現優異，mAP、Precision、Recall、F1-score 都有著優秀的表現。而在 OCR 文字辨識中也有著不錯的表現，本研究在訓練資料的 mAP 上新增了  $mAP_{0.5:0.95}$ ，這是指在多個 IoU 閾值下的平均精度均值，從 0.5 到 0.95，以 0.05 為步長做計算共 10 個不同的 IoU 閾值。如此可以更全面地評估模型在不同檢測嚴格程度下的表現。如下表 3；訓練集之績效如下圖 3 所示。在本次實驗當中，針對訓練資料做 Precision Curve、F1 Curve、Recall Curve 做可視化圖形呈現，如下圖 3, 4, 5 所示。

表 3 YOLOv5 訓練/測試集之績效表

YOLOv5	Recall	Precision	F1-score	mAP	mAP_0.5:0.95
Training Data	100%	99.81%	100%	99.45%	86.01%
Testing Data	100%	100%	100%	92.38%	

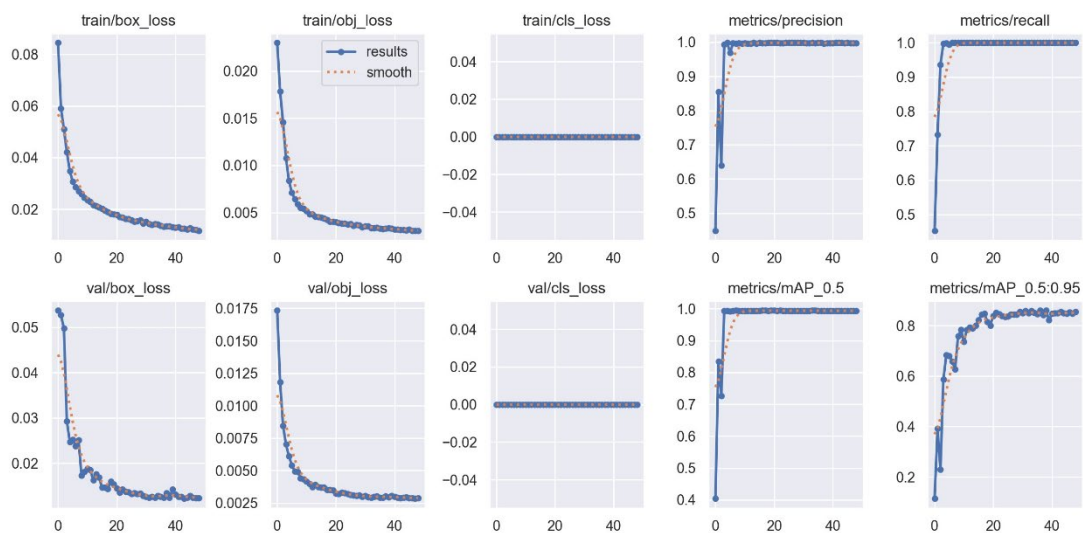


圖 5 訓練集績效圖

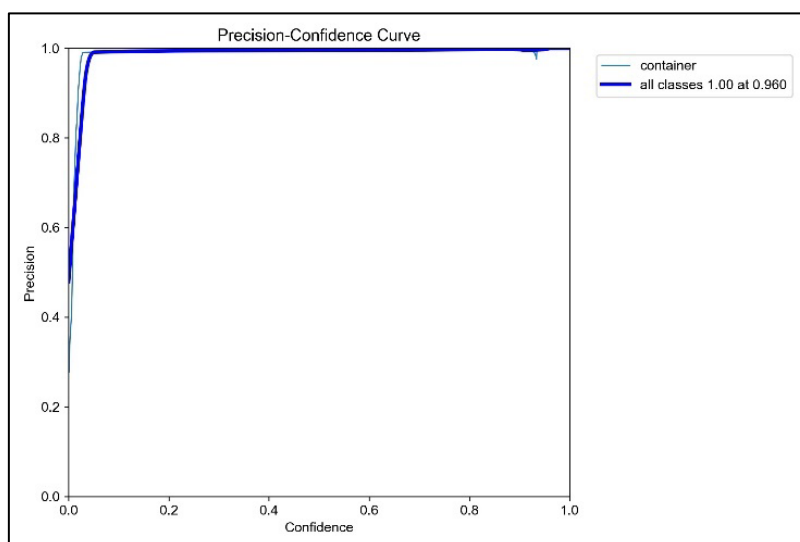


圖 3 訓練集 Precision 績效圖

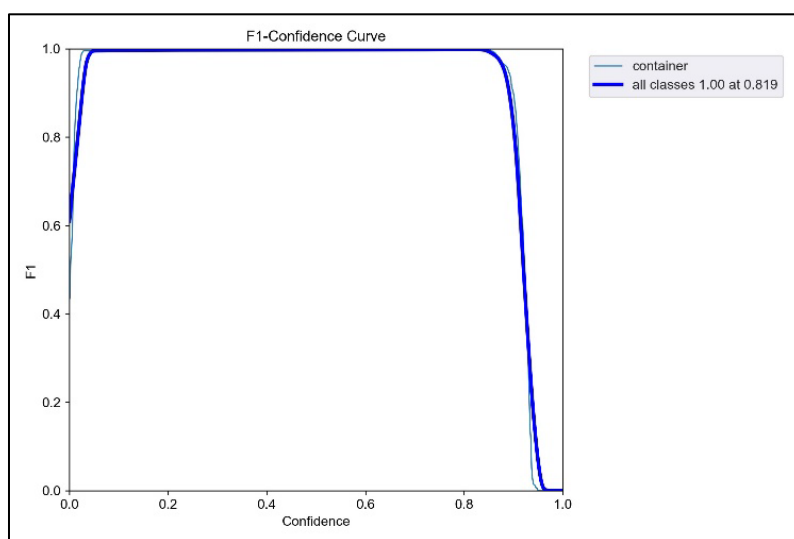


圖 4 訓練集 F1-score 績效圖

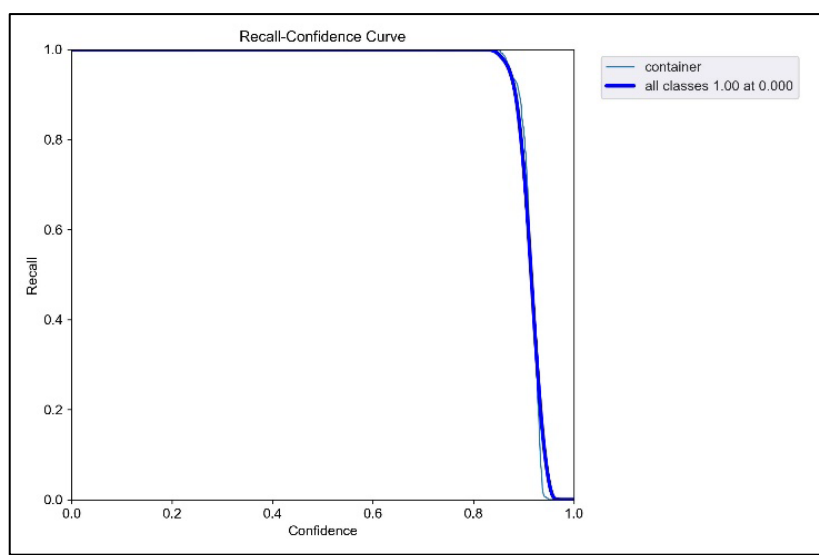


圖 6 Recall Curve 圖

### 3.4.2 Faster R-CNN Model

本實驗之 Faster R-CNN 模型在訓練、測試集在物件偵測區塊上的績效表現相對於 YOLOv5 平分秋色，無論在精確度(Precision)還是平均精確均值(mAP)來看，兩者表現相差不超過 3%。績效結果如下表 4 所示，成功辨識文字區域示意圖如下圖 6 所示。

表 4 R-CNN 訓練/測試之資料表

R-CNN	Recall	Precision	F1-score	mAP
Training Data	100%	96.75%	98.35%	97.30%
Testing Data	100%	94.53%	97.19%	100%

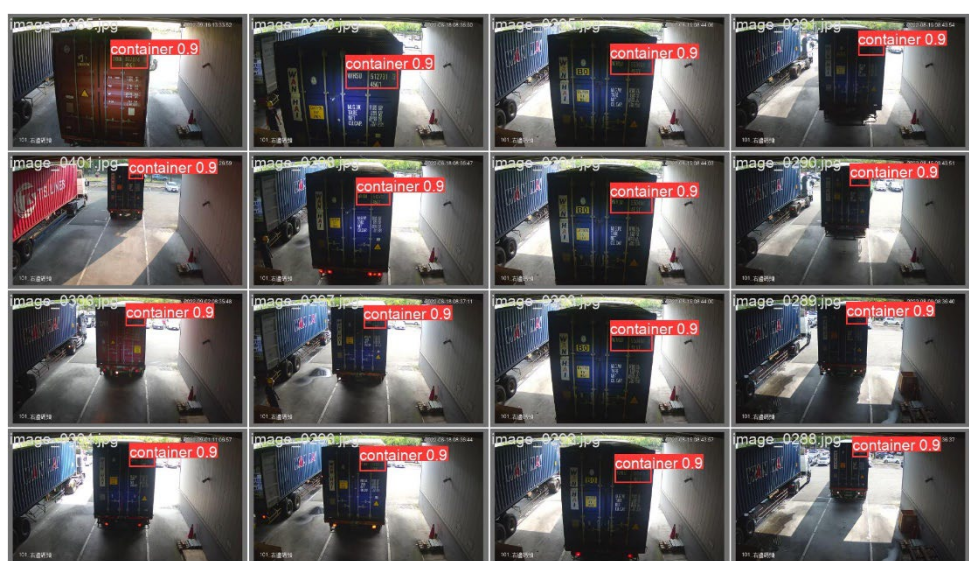


圖 7 貨櫃號碼區塊成功辨識示意圖

### 3.4.3 光學文字辨識

最後，本實驗將訓練好的 YOLOv5、R-CNN 兩個模型做光學文字辨識，並以影片資料集作為測試。在圖像資料集的辨識中，會將圖像做前置處理，例如增加對比度、轉灰階圖像、二值化處理，並選用 Tesseract-OCR 作本次文字辨識使用，以此比較前 11 位數字是否正確。ORC 辨識準確率如下表 5 所示。辨識情形則由下圖 7, 8 所示。

表 5 OCR 平均辨識準確度表

	YOLOv5	R-CNN
圖像 OCR Accuracy	82%	82%
影片 OCR Accuracy	90%	90%

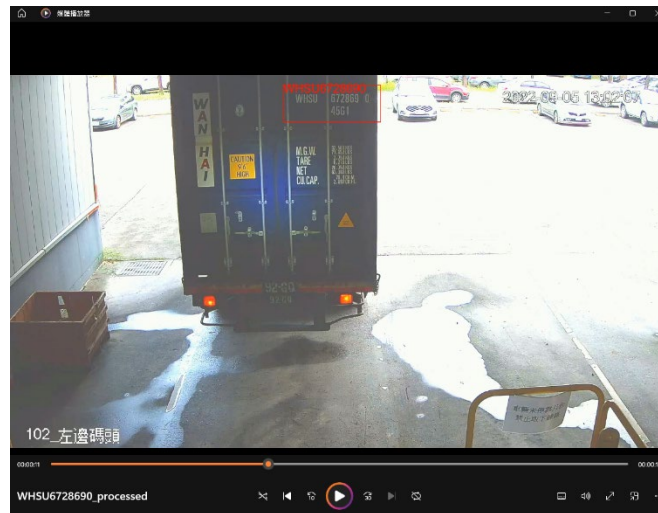


圖 8 影片資料集辨識成功示意圖

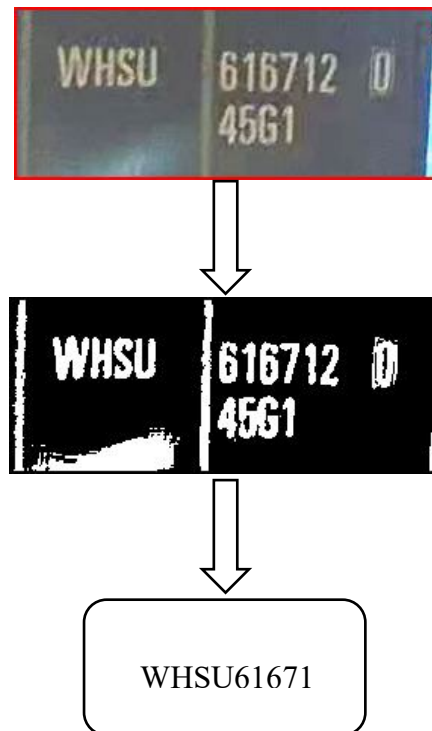


圖 9 成功辨識之過程圖

#### 四、結論

本次作業為實作一個基於深度學習的貨櫃號碼自動辨識系統，旨在提升物流管理的效率和準確性。在實驗中，本研究選擇了 YOLOv5 和 Faster R-CNN 兩種物件偵測模型進行訓練和測試，並結合 Tesseract-OCR 文字辨識技術對檢測到的貨櫃號碼區域進行文字辨識。實驗結果顯示，YOLOv5 在物件偵測方面表現優異，測試集上的 mAP 達到 92.38%，而 Faster R-CNN 則在精確度和召回率方面稍有落後，但仍保持較高的準確性。此外，兩種模型在 OCR 辨識中的準確率均為 82%，如此也表示出在圖像之處理上和參數調整方面還有進一步提升的空間。

本研究過程中發現資料增強技術對提升模型的泛化能力有顯著效果，透過旋轉、平移等技術擴展訓練數據，有助於模型在不同環境下保持穩定的辨識績效。同時，在模型訓練中使用 Early Stopping 策略可以有效防止 Overfitting，確保模型在新數據上的表現。

未來的研究可以進一步優化模型的參數設置，嘗試更多不同的深度學習框架和模型結構，並且擴展資料集的規模和多樣性，以提升系統的整體績效。此外，探索在光源不足之環境和模糊圖像中的辨識能力，也是提高系統實用性的重要方向。綜上所述，本研究成功實作了深度學習方法在貨櫃號碼上的自動辨識，也從本次專案作業中更瞭解了模型的應用。



## 五、參考文獻

- 楊宇任(2023)。自動識別集裝箱序列碼。 <https://hdl.handle.net/11296/2f8dcs>
- Hao-Chieh Kuo(2021)。yolov5 基本設置與物件偵測(上)。 <https://kuonumber.medium.com/yolo5%E5%9F%BA%E6%9C%AC%E8%A8%AD%E7%BD%AE%E8%88%87%E7%89%A9%E4%BB%B6%E5%81%B5%E6%B8%AC-%E4%B8%8A-98f4d8d748cc>。
- Hao-Chieh Kuo(2021)。yolov5 tips(下)。 <https://kuonumber.medium.com/yolo5-tips-%E4%B8%8B-a4559d275501>
- Xiaotudui(2019)。【PyTorch 教程】P1. PyTorch 環境的配置及安裝。 <https://medium.com/@zhou.lawliet/pytorch%E6%95%99%E7%A8%B-p1-pytorch%E7%92%B0%E5%A2%83%E7%9A%84%E9%85%8D%E7%BD%AE%E5%8F%8A%E5%AE%89%E8%A3%9D-c40d6a395531>
- Jocher, G., Stoken, A., Borovec, J., Chaurasia, A., Changyu, L., Hogan, A., ... & Laughlin, M. (2020). YOLOv5. Available at <https://github.com/ultralytics/yolov5>
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 580-587). <https://doi.org/10.1109/CVPR.2014.81>
- Girshick, R. (2015). Fast R-CNN. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448). <https://doi.org/10.1109/ICCV.2015.169>
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems (pp. 91-99). <https://doi.org/10.1109/TPAMI.2016.2577031>