

基於 YOLOv8 機器視覺應用於港區船舶監測之研究

溫嘉佑*¹ 李子宜* 郭紹宇**

*國立高雄科技大學 造船及海洋工程系
**エスオーエル株式会社

關鍵詞：影像辨識、深度學習、YOLOv8、資料增強、速度檢測、Python

摘要

大型商港每日船舶進出眾多以及大船進港時需要的拖船作業，隨著智慧港口成為全球商港的新趨勢，對於提高港口作業的效率和安全性有著迫切需求，其中若能精確辨識船舶進出和進港速度是對於提升港口安全和運營的關鍵。本研究採用最新一代的 YOLO 通用物件視覺辨識技術 YOLOv8 (You only look once version 8) 演算法進行深度學習權重訓練，模型網路資料集與測試地點皆為進出高雄港第二港口之船舶，於實驗室個人電腦架設下訓練權重。在樣本總數相同的情況下，藉由將資料集資料增強，比較驗證資料集的比例 20%、30%、40% 以及船舶速度準確度，選取最佳的模型。本研究開發辨識船舶及量測船速之系統，此裝置可採用港區攝影機畫面，對船舶進出港時進行數量統計和測速，以及在拖船作業時進行監測避免小型快艇違規高速進入，藉此降低船舶碰撞事件發生，以提高港口安全和運營效率。

前言

港口安全和運營效率的提升是海運業的核心關鍵，作為臺灣主要的重要出口管道之一，商港承擔著巨大的貨物吞吐量，每日都有眾多船舶進出。

為了進一步提高港口的安全性和效率，本研究開發了一套系統，能夠利用港區的攝影機畫面來辨識船舶並量測速度。這套系統的應用不僅限於對進出港船隻的數量和速度進行監測統計，還包括在拖船與引水人作業期間的安全監控，防止小型快艇違規高速進入港口區域，從而有效降低船隻碰撞的風險，以提高港口安全和運營效率。

目前有許多針對影像辨識開發之演算法，主要是以卷積神經網路 (Convolutional Neural Network, CNN) 作為主幹，透過匯入大量人工事前標記好之影像資料產生二進制編碼權重，形成一個影像辨識的演算法。

You Only Look Once (YOLO) 是一種屬於 One-stage learning 的影像辨識演算法，由美國華盛頓大學學者 Joseph Redmon, Ali Farhadi 於西元 2016 年提出一種開源 (Open Source) 深度學習演算法之影像辨識技術 [1]，歷經多代的演進已從 YOLOv1 發展到 YOLOv9。

本研究採用 2023 年推出的 YOLO 通用物件辨識技 YOLOv8m (You Only Look Once version 8 medium model) 演算法進行深度學習權重訓練，在樣本總數相同的情況下，藉由將資料集資料增強，比較驗證資料集的比例 20%、30%、40% 以及船舶速度準確度，選取準確度和泛化能力最佳的模型應用於本系統。

YOLOv8 介紹

YOLOv8 [2] 是由 Ultralytics 的公司所開發，是基於 2023 年 1 月 10 號開源的 YOLOv 5 的更新版本。

YOLOv8 並不是一個單獨的模型，其中包含四個網路模型 (YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l、YOLOv8x)，這五個模型的網路類型是一樣的，網路模型的深度和寬度由深度因子和寬度因子控制，使得 YOLOv8 的五種網路模型具有不同的深度和寬度，官方在 COCO Val 2017 數據集測試各模型的總平均準確度 (mAP, mean-AP)，隨著模型愈複雜其 mAP 越高，如圖 1 所示。

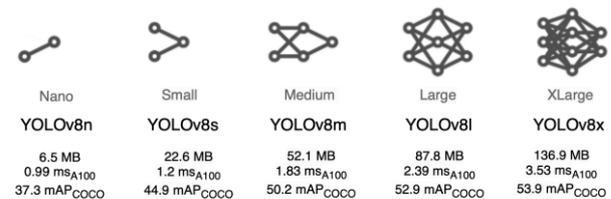


圖1 YOLOv8 各模型網路比較 [2]

研究方法

本研究的辨識流程圖 (圖 2) 中各重要區域分別說明如下：

本研究使用的 YOLOv8 演算法屬於監督式學習 (Supervised Learning)，因此需準備訓練模型的訓練集與驗證集的影像，在樣本總數相同的情況下，藉

¹溫嘉佑 (F111186107@nkust.edu.tw)

由調整驗證資料集的占比，三個權重驗證資料集占比分別為 20%、30%、40%。

(1) Dataset

本研究資料庫照片經過資料增強後共 6394 張，拍攝地點為進出高雄港第二港口之船舶，標記目標共有三種名稱分別為 Ship、Tug boat、Speed boat，標記格式為 YOLO(.txt)。

(2) Training

本研究使用 YOLOv8m 的網路架構，YOLOv8m 為五種模型中網路適中，既有一定的準確率且訓練花費的時間也不長，考量時間成本以及物件準確率，故作為研究中主要模型。

(3) Detection

將要檢測的影片或即時直播畫面導入本系統中，觀察即時檢測畫面。

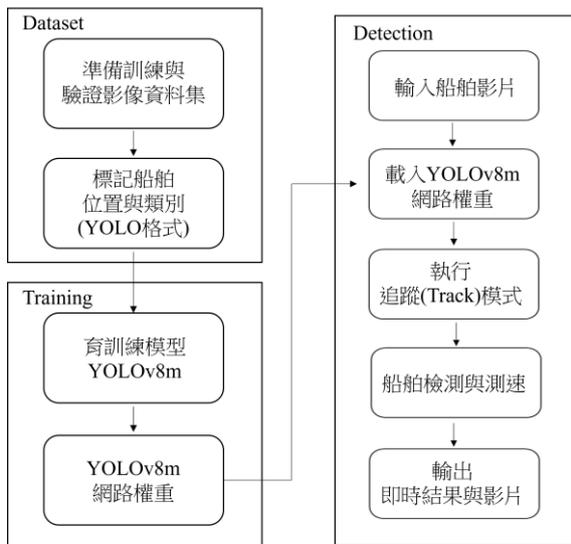


圖2 YOLOv8 辨識流程圖

訓練資料庫建置

1. 資料收集與標記

本研究資料庫照片與測試影像皆為進出高雄港第二港口之船舶。標記目標共有三種名稱分別為 Ship、Tug boat、Speed boat，標記名稱 Ship 泛指大型船舶如：貨船、散裝船、油輪、遠洋漁船等大噸位船舶，標記名稱 Tug boat 泛指拖船，標記名稱 Speed boat 泛指航速快的船舶。

標記工具採用 LabelImg，為美商微軟公司旗下軟體原始碼代管服務平台 Github 上的一項開源專案，以 Python 編寫支援 YOLO(.txt)、ImageNet(.xml) 標記格式。

2. 資料增強(Data Augmentation)

在照片樣本數較少的情況下進行訓練容易發生 Overfitting(過度學習)，為了避免此情況發生，採用影像增強技術使樣本數增加，且也能大幅提升模型的學習效果和泛化(Generalization)能力，在應對非資料庫數據時能有不錯的表現。

2.1 影像增強工具-Albumentations

Albumentations[3]是一個快速靈活的圖像增強工具，專為機器學習和深度學習中的圖像前處理任務而設計，它提供了豐富的增強技術，如旋轉、裁剪、模糊等，且增強後的照片也無需再做標記，大幅的節省人工標記的時間。

為了增加目標偵測物件的多樣性，本文所使用 3 種方法為旋轉 180 度、運動模糊、隨機排列 RGB 通道，圖 4 為圖 3 使用圖像增強後之圖片。



圖3 原始船舶圖片



圖4 增強後船舶圖片

2.2 透視變形(Perspective Transformation)

透視效果是一種視覺現象，特別是在攝影和拍照中常見，當目標物體位於相機或視覺感知系統較遠的距離時，會產生明顯的深度感和變形，遠處的物體看起來比實際小，而且隨著距離的增加，這種大小差異會變得更加明顯。

在 YOLOv8 版本中，透視變換是可以通過調整其參數來改變對物體大小的估計。當透視值設定較高時，模型會更加關注遠處的物體，並根據透視效應來估計其大小。反之，若透視值較低，則模型將專注於近處的物體，以減少透視效應的影響，如圖 5 展示了透視變形的效果。

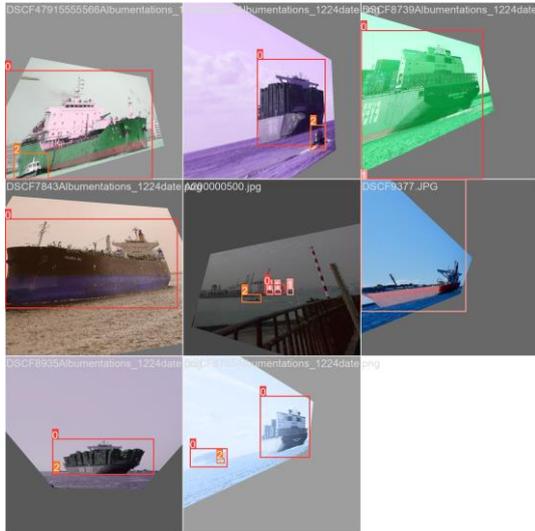


圖5 透視轉換示意圖(perspective=0.003)

船舶檢測與測速

1. 船舶檢測

本研究開發辨識船舶及量測船速之系統，是使用 YOLOv8 任務中追蹤 (Track) 模式進行影片分析，它不僅可以辨識影片中每幀物體的位置和類別，還可以将偵測到的物件命名一個唯一的 ID (identity)。

當船舶進出港時，系統辨識到目標物會顯示該目標物的辨識框(bounding box)在此上方有該目標物的訊息如：ID、種類、辨識信心度(如圖 6)，當目標物的辨識框(bounding box)先碰觸綠色線時判定為進港船隻，反之若先碰觸藍色線時判定為出港，視窗左上角為進港船隻訊息統計版，左下角為出港船隻訊息統計版(如圖 6)，會進行更新。



圖6 船舶進出港統計畫面

2. 船舶測速

本系統測量船舶船速，為當目標物的辨識框(bounding box)在碰觸藍線與綠線後，記錄行駛間之幀數除上當前影格速率(FPS)得出通過的時間和來計算船舶船速。為了將測量的距離從像素轉換成公尺(m)，假設一個恆定的像素(pixel)與公尺的比率，通過建立一致的距離測量標準來進行船速(knot)計算成為可參考性。



圖7 船舶進出港測速畫面

實驗結果

本研究在樣本總數相同的情況下，三個權重驗證資料集占比分別為 20%、30%、40%，訓練參數訓為：影像解度(imgsz) 640x640、訓練步數(epochs) 300、訓練的批量大小(batch)8、透視變形(perspective)0.003。

使用電腦硬體規格為 CPU i9-11900K，RAM 96GB，GPU RTX 3070TI 8G，作業系統為 Ubuntu 20.04 LTS，程式語言主要使用 Python 3.8.10 開發，搭配 Pytorch 1.11.0，CUDA 版本為 11.5.119，cuDNN 版本為 v8.3.0。

在機器人學習領域與統計分類問題中，在真實情況獲得 4 個一級指標，分別為真陽性(True Positive)、偽陽性(False Positive)、真陰性(True Negative)、偽陰性(False Negative)。

1. 評價指標

(1) 精確率 (Precision)

在所有分類為陽性樣本中，被正確判定與真實情況相符為陽性的占比。

$$Precision = TP / (TP + FP) \quad (3)$$

(2) 召回率 (Recall)

召回率則是在所有陽性樣本中，被正確判定為陽性比例之占比。

$$Recall = TP / (TP + FN) \quad (3)$$

(3) P-R Curve (Precision and Recall Curve)

前述兩個重要的二級指標，精確率和召回率繪製成的曲線圖，曲線圖下的面積就是代表平均準確率 (Mean Average Precision, AP)，mAP 則是所有類別的 AP 平均值，數值越趨近於 1，代表該模型有著更良好的表現。

驗證資料集在 20%、30%、40% 占比下訓練出的權重檔案，P-R Curve 圖中各 mAP 相差不大(表 1)，但以時間成本來說 40% 花費 9.432 小時時間較短，20% 花費 21.061 小時、30% 花費 18.265 小時。

在驗證資料集占比 20% P-R Curve 圖，由訓練資料集與驗證資料集獲得，Ship-mAP 0.994 有最高平均準確度。總平均準確度(All mean-AP,0.978)以下有 Speed boat-mAP 0.961 如圖 8 所示。

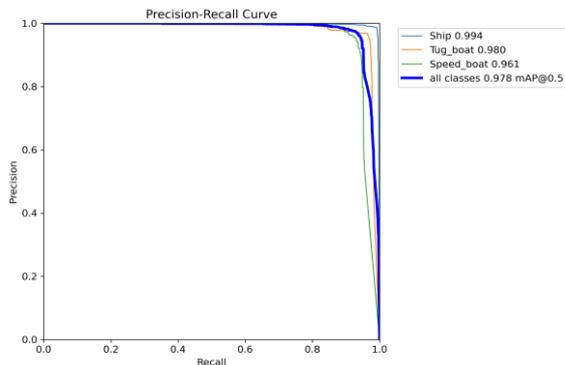


圖8 驗證資料集 20% P-R Curve 圖

在驗證資料集占比 30% P-R Curve 圖，由訓練資料集與驗證資料集獲得，Ship-mAP 0.993 有最高平均準確度。總平均準確度(All mean-AP,0.978)以下有 Tug boat-mAP 0.976、Speed boat-mAP 0.965 如圖 9 所示。

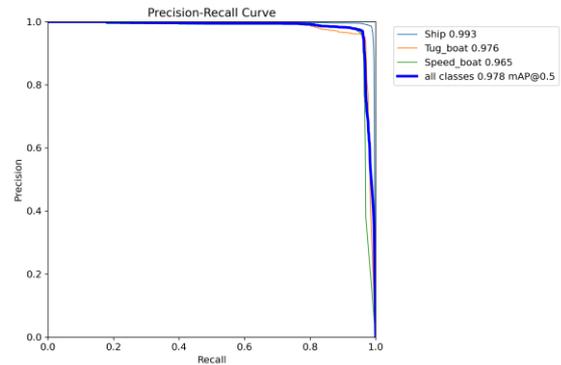


圖9 驗證資料集 30% P-R Curve 圖

在驗證資料集占比 40% P-R Curve 圖，由訓練資料集與驗證資料集獲得，Ship-mAP 0.994 有最高平均準確度。總平均準確度(All mean-AP,0.979)以下有 Tug boat-mAP 0.980、Speed boat-mAP 0.962 如圖 10 所示。

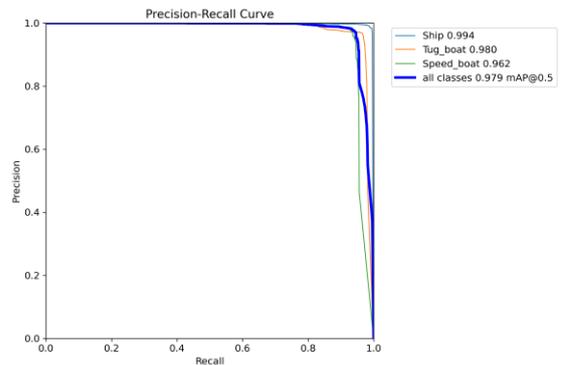


圖10 驗證資料集 40% P-R Curve 圖

表 1 各權重結果 P-R Curve 比較

驗證資料集	Ship	Tug boat	Speed boat	All classes
20%	0.994	0.980	0.961	0.978
30%	0.993	0.976	0.965	0.978
40%	0.944	0.980	0.962	0.979

2. 影片測試

測試影片地點為進出高雄港第二港口，在驗證資料 20%、30%、40% 權重下比較各模型的泛化能力和船舶速度準確度。

圖 11、12、13 皆以人工現場量測船速當成實驗參考值，圖 11 為船舶進港影片擷取畫面、圖 12、13 為船舶出港影片擷取畫面，將上述 3 部影片由各權重輸出辨識結果做比較，如表 2 所示各表中各驗證資料集占比權重都成功辨識到目標物和船速，其中在圖 11、13 驗證資料集 30%、40% 權重下測量船速較接近實驗參考值，圖 12 各權重測速值與實驗參

考值的百分誤差過大，在驗證資料集 20%較接近實驗參考值。

圖 14、15、16 為船舶進出港船舶作業影片擷取畫面，主要測試辨識能力，輸出船舶數量結果做比較，如表 3 所示各各驗證資料集占比權重都成功辨識到目標物，在圖 16 影片中驗證 20%權重沒精準辨識到出港 Tug boat 導致少統計到，其中驗證資料集 30%、40%權重下表現良好皆符合實驗參考值。



圖11 船舶進港 驗證 20% 檢測畫面



圖12 船舶出港 驗證 30% 檢測畫面



圖13 船舶出港 驗證 40% 檢測畫面

表 2 各權重辨識結果比較 (船速)

驗證資料集	圖 11	圖 12	圖 13
20%	12.13 Knot 誤差 9.948%	6.16 Knot 誤差 18.52%	7.98 Knot 誤差 3.39%
30%	14.74 Knot 誤差 9.428%	5.84 Knot 誤差 22.75%	8.06 Knot 誤差 2.42%
40%	14.35 Knot 誤差 6.533%	5.88 Knot 誤差 22.22%	8.05 Knot 誤差 2.54%
實驗參考值	13.47 Knot	7.56 Knot	8.26 Knot



圖14 船舶進港拖船作業 驗證 40% 檢測畫面

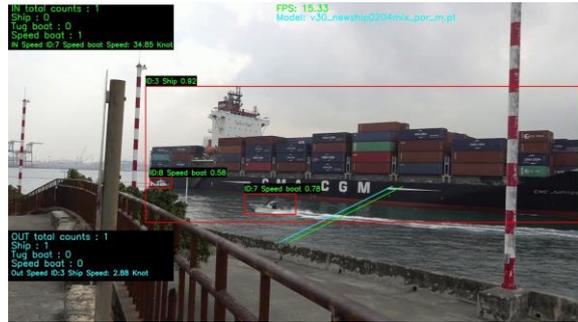


圖15 船舶出港引水人作業 驗證 30% 檢測畫面



圖16 船舶進港拖船作業 驗證 20% 檢測畫面

表 3 各權重辨識結果比較(統計數量)

驗證資料集	圖 14	圖 15	圖 16
20%	進港 Ship:1 Tug boat:1	進港 Speed boat:2 出港 Ship:1 Speed boat:1	進港 Speed boat:2 出港 Ship:1 Tug boat:1
30%	進港 Ship:1 Tug boat:1	進港 Speed boat:2 出港 Ship:1 Speed boat:1	進港 Speed boat:2 出港 Ship:1
40%	進港 Ship:1 Tug boat:1	進港 Speed boat:2 出港 Ship:1 Speed boat:1	進港 Speed boat:2 出港 Ship:1 Tug boat:1

實驗 參考值	進港	進港	進港
	Ship:1	Speed boat:2	Speed boat:2
	Tug boat:1	出港	出港
		Ship:1	Ship:1
		Speed boat:1	Tug boat:1

結論

本研究歸納出以下結論：

- (1) 即使在樣本數較少的情況下，仍可以使用資料增強讓樣本數增加後進行深度學習權重訓練，在本研究中驗證集 40% 下訓練出模型有較好的準確度和泛化能力以及所需訓練時間較短。
- (2) 本研究成功開發了一個利用計算機視覺進行船舶監測的系統。首先，使用 YOLOv8m 模型辨識影片中的船舶，然後通過分析邊界框訊息並跟踪船舶的移動，計算出其節速(Knot)。本系統展示了計算機視覺應用在港口管理和智慧港口上的潛在應用。但是，若要在真實場景中實現準確速度測量時，則需要進行的校正和微調。

參考文獻

1. Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," Proceedings of the IEEE conference on computer vision and pattern recognition, Las Vegas, NV, USA, (2016).
2. Glenn Jocher, YOLOv8 開源代碼 GitHub , <https://github.com/ultralytics/ultralytics> 。 (2023)
3. Jian ZhouHirt, MingHan Yan, Cong Luo, XiaoXue Xing, "Underwater Sonar Target Detection Based on YOLOv5," International Conference on Electronic Information Engineering and Computer Science (EIECS), pp729-732 (2021).
4. Alexander Buslaev , Vladimir I. Iglovikov, Eugene Khvedchenya , Alex Parinov , Mikhail Druzhinin 5and Alexandr A. Kalinin, "Albumentations: Fast and Flexible Image Augmentations," Information," 11-2(2020).
5. 郭紹宇，攜帶式即時辨試船艙標誌影像裝置之研究，國立高雄科技大學，碩士論文(2021)。
6. You-Wei Wu , Dong-Lin Li, "Deep Learning-Based Fish Classification with Data Augmentation," IEEE Computer Society (2023).
7. Renuka Devi, G., Athiyan, S., Ilan Thamizh, k. and Hema Varthini, V., " Speed and Distance Detection for Highway Traffic using Python," ICSCAN International Conference on System, Computation, Automation and Networking (2023).

Research on the Application of YOLOv8 Machine Vision for Ship Monitoring in the Port

J-Y Wen Z-Y Lee* S-Y Kuo***

*Department of Naval Architecture and Marine Engineering, National Kaohsiung University of Science and Technology

**SOL Corporation

Keywords: Image recognition, Deep learning, YOLOv8, Data Augmentation, Speed Detection, Python.

ABSTRACT

With the large number of ships entering and leaving major commercial ports daily and the tugboat operations required for large ships entering the port, the rise of smart ports as a new trend in global commercial harbors has created an urgent need to improve port operation efficiency and safety. Among these needs, the ability to accurately identify ship entries and exits and their speeds is key to enhancing port safety and operations.

This study employs the YOLO general object visual recognition technology, YOLOv8 (You Only Look Once version 8) algorithm, for deep learning weight training in 2023. The model network dataset and testing location are both focused on ships entering and exiting the second port of Kaohsiung Harbor, with weight training conducted on a personal computer in the laboratory. Under the condition of having the same total number of samples, the study enhances the dataset and compares the validation dataset ratios of 20%, 30%, and 40%, as well as the accuracy of ship speed, to select the model with the best accuracy and generalization ability.

The system developed in this study recognizes ships and measures their speeds using port area cameras. It can count ships entering and exiting the harbor, monitor tugboat operations, and prevent small speedboats from illegally entering at high speeds. This reduces the occurrence of ship collisions, thereby enhancing port safety and operational efficiency.