人工智慧於乒乓球動態偵測及球速計算之研究與應用

黃偉祥、李易、陳先正、李錫捷*
國立金門大學資訊工程學系
*cjlee@nqu.edu.tw

論文摘要

本研究使用以 Roboflow 上的 MS COCO 為基礎模型，搭配自行錄製的訓練影片，以每秒 30 幀的速度對乒乓球進行物件偵測。該研究主要以解析度為 1280*720 的連續影片作為訓練集，並採由上而下的視角進行拍攝。

結合以 Python 開發的演算法來對每顆球的速度及落點進行分析，期望打造出能實時檢測乒乓球的位置與速度來進行得分計算的監測系統，以此幫助運動員在訓練時可以將注意力都用在訓練上並在訓練結束後可以得到分析結果。本研究希望能藉此提升乒乓球在國際賽場上的地位，幫助該運動未來的發展以及應用，同時促進 AI 在這一領域的突破，以此帶來更多貼近生活、並為生活帶來便利之研究。

關鍵詞：桌球、物體識別、Roboflow

前言

人工智慧是一種模擬人類智慧的技術，通過電腦系統模擬人類的思維過程和行為，從而執行各種任務，如學習、理解、推理和解決問題。它的目標是使機器能夠像人類一樣感知、理解和應對環境，並自主地進行決策和行動。機器視覺則是人工智慧的一個分支，它使機器能夠理解和解釋視覺信息，就像人類的視覺系統一樣。通過使用各種感測器和相機，機器視覺系統可以捕捉、處理和分析圖像或視頻數據，並從中提取有用的信息。這些系統通常使用深度學習和計算機視覺技術來檢測和識別圖像中的對象、人物、場景和動作，從而實現自主導航、監控、品質控制、醫學影像分析等應用。機器視覺的目標是讓機器能夠以類似於人類的方式理解和利用視覺信息，從而更好地與世界互動和協作。

本研究便是採用機器視覺的方式，先通過攝影機紀錄乒乓球的對打過程，再一幀一幀的對其進行標記，然後將資料集傳給Roboflow模型（MS COCO）進行訓練。Roboflow是一個強大的工具，可以幫助開發者和研究人員輕鬆管理、標記和訓練機器學習模型所需的圖像數據。它提供了一個直觀的界面，讓用戶可以快速上傳、轉換和標記圖像數據，同時還能自動處理數據增強、分類和標記等任務，節省了大量的時間和精力。Roboflow支持多種機器學習框架和平台，讓用戶可以輕鬆地將標記好的數據集集成到他們的項目中，加速模型的開發和部署過程。

在這次的研究中，我們參考了過往以單視角、低質量攝像頭為主的乒乓球分析[1]，以此加強我們在訓練資料上的蒐集過程，並透過基於機器視覺的乒乓球比賽戰術大數據分析及應用研究結構[2]，來協助進行實時的球體物件偵測。在該研究過程裡，我們主要使用乒乓球發球機作為我們的訓練對手。乒乓球發球機是一種專為乒乓球運動員設計的訓練工具。它通常由一個可調節的發球機構、發球速度控制器和發球頻率調節器組成。這種機器能夠模擬各種不同的球路和速度，讓運動員有機會在訓練中練習應對不同類型的球，提高他們的反應速度和技術水平。它也是一種很好的自我訓練工具，運動員可以根據自己的需要調整發球模式，隨時進行訓練，提升技術。為了能更了解發球機的運動過程，我們參考了與其相關的發現和研究[3]，以此更好的調整相關參數，包括速度、角動量、落點等等。

而在訓練資料的蒐集過程中，我們也需對其進行資料標記。資料標記是將原始資料加工處理的過程，透過標記人員根據特定規則或標準，將資料標上標籤、標記或分類，以便機器學習模型能夠理解和處理這些資料。這包括文字、圖像、音訊等各種形式的資料。資料標記的目的是讓機器能夠準確地理解和處理資料，從而實現各種應用，如自然語言處理、影像辨識、智能搜索等。本研究便是使用於 Roboflow 平台上內建的資料標記功能，對數百張的圖片進行相關圈選和處理。並在處理完畢後利用其內建的訓練系統對圖像進行分析，以此誕生我們的模型。

在訓練過程裡，我們參考了 YOLOv5 在類似項目上的球體偵測表現[4]，以此推斷我們的期望值。並效仿其他針對乒乓球的物體檢測器研究[5]，對我們的模型進行實際訓練。為了解決資料量不夠充足的問題，我們對每一幀進行分割和旋轉，並與驗證集進行搭配達到反覆訓練、確認之目的，以此降低最終的 loss 值。該作法除了可以減輕與測試集的差異，還可降低過擬合的可能性，從而符合現實世界的期望。

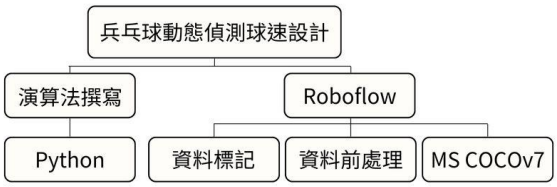
待訓練完畢，本研究使用 Python 來對偵測結果進行分析，並參考基於空間域資訊的乒乓球落點檢測演算法研究[6]，來計算球體在三維空間中的落點及速度，從而實現得分的計算。而為了能時刻追蹤乒乓球的軌跡來源，我們也參考了基於時態特徵的軌跡檢測研究[7]，以此避免多個球體出現時的路徑偵測錯誤，造成選手在得分過程中的誤判。

該研究希望能以此方法來達到實時的速度與分數檢測效果，從而降低相關的運動工作成本，達到令選手專注於訓練之目的。

研究方法

1、系統設計

本研究採用Roboflow平台來進行實作，然後將我們在Roboflow上訓練好的模型導入到Python來撰寫的演算法並進行進一步的分析。其系統架構圖如圖一所示。



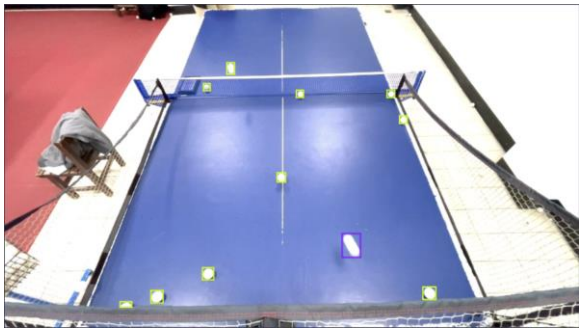
圖一、系統架構圖

2、Roboflow

本研究使用了約1000張的圖像來進行乒乓球物體識別的訓練，透過Roboflow平台上提供的影像前處理功能來增加資料量，最後透過平台上提供的 MS COCOv7 公開模型來進行訓練。MS COCO是基於YOLOv8這個強大的物體識別模型去進行訓練出來的模型。

A. 資料標記

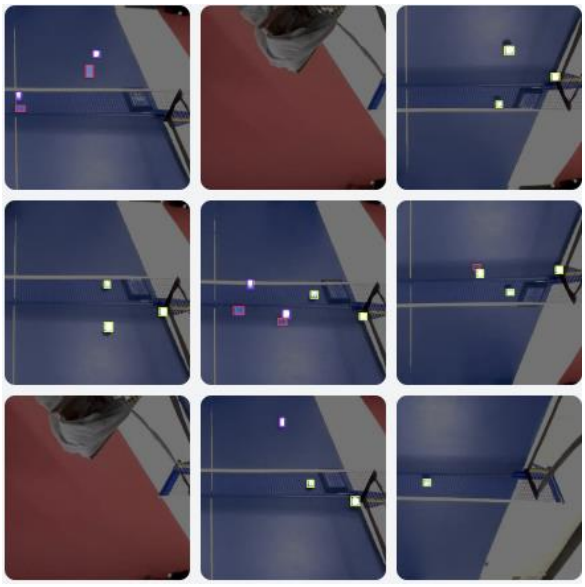
在資料標記的部分我們使用Roboflow平台上提供的Auto Label的功能去做第一步的標記，然後人工去進行下一步的審核與修正。Auto Label是Roboflow套用autodistill提供的自動標記功能。我們分別標記了靜止的球（綠色框）與正在快速移動的球（麵條狀，紫色框），範例如圖二所示。



圖二、標記範例

B. 資料前處理

首先我們將圖像重新調整為320*320以防訓練的速度太慢，之後對圖像進行Tile將圖像切割成2(row)*4(column)這是為了提升小型物體識別的準確率。之後對訓練資料做水平與垂直反轉，為了增加資料量，最後會將重複的圖像刪除。經過前處理後的圖像如圖三所示。



圖三、經過前處理後

C. 資料集介紹

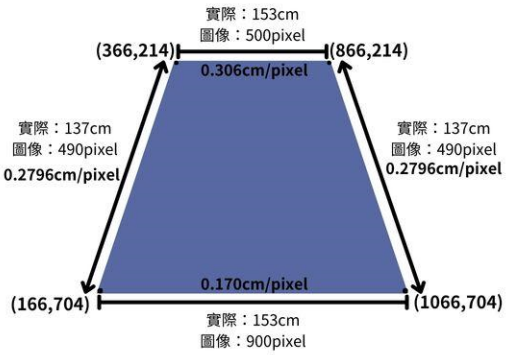
我們利用手機從上方拍攝影片，將其轉換為圖像，fps為30。我們的資料量從大約1000張的圖像經過前處理後的到15184張圖像，其中84%為訓練資料共12784張，11%為驗證資料共1600張，5%為測試資料共800張。

3、演算法

為了能抓到球的移動軌跡，我們對每顆球進行 id 標記，使其都能抓到最初的來源，從而更好的銜接之後的得分判定。在抓取過程中，我們將影片的每一幀獨自拆開，並利用 Roboflow 的 API 取得每一畫面下的球的位置、數量、座標等等，然後記錄在我們的陣列中。紀錄完畢後，我們將每一幀與前一幀進行比較，根據相對位置取得其目前球的速度、方向，並儲存在原始的資料結構裡。等到所有球的速度跟方向都抓到後，我們再對其進行大小比較，取得目前球的最小速度值；再對所有從來源球到目前球的配對進行比較，同樣以最小速度值的配對作為正確可能，從而排除球體來源路徑重複的情況，並更好的將結果印在螢幕上。該演算法的時間複雜度為O(n^3)，因為要先用迴圈抓到目前一幀下的球，再列出該幀下所有球的可能路徑，然後再對所有可能路徑進行比較、去除重疊的軌跡。未來還會再用更好的方式對其進行優化，但目前已經能出正確抓出球速及軌跡。

而為了計算出球速，我們透過球桌上四點的座標如圖十三所示，首先計算出對於Y軸每一個pixel等於0.2796cm (137cm/490pixel)。接下來計算X軸，但是我們拍攝的角度導致桌子在影像上呈現一個菱形，所以我們在計算X軸時需要對於每一個X軸在Y軸上的變化量以減少誤差。首先我們計

算出X軸上最短的邊（最大值）與最長的邊（最小值）的對應值分別為0.306cm/pixel 與0.17cm/pixel。相減後得到他們的相差值為0.136cm/pixel，接著用這個值除以Y軸的距離490（704-214）可得Y軸每改變1那麼X軸的對應就改變0.0002775cm/pixel。最後將X與Y軸的變化計算出實際移動的距離(cm)後除以每一幀的時間變化(0.03333s)可得速度(cm/s)，再經過單位轉換可得到常用的速度單位(km/h)。



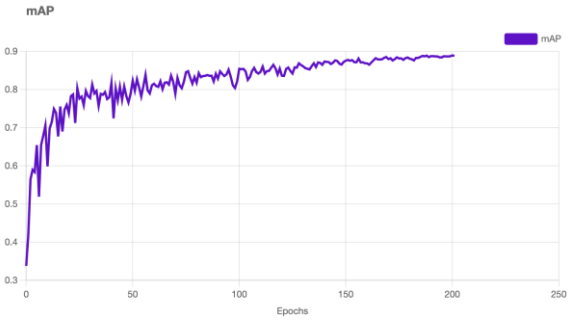
圖十三、圖像裡球桌的比例

舉例，像圖十二中的frame4 到fram5，球id為5從[646.50,256.00] 到 [682.50,345.00]為例。對Y軸來說球移動了89pixel，換算後實際移動的距離為24.8844cm。對X軸來說球移動了36pixel，X軸對應值為0.306-0.0002775*86.5=0.2819cm/pixel ，所以對X軸來說實際移動了10.15cm。透過直角公式可得實際移動距離為26.8748cm。我們每兩幀之間的時間差為0.03333s，速度為806.325cm/s，經過單位轉換可得29.0277km/h。

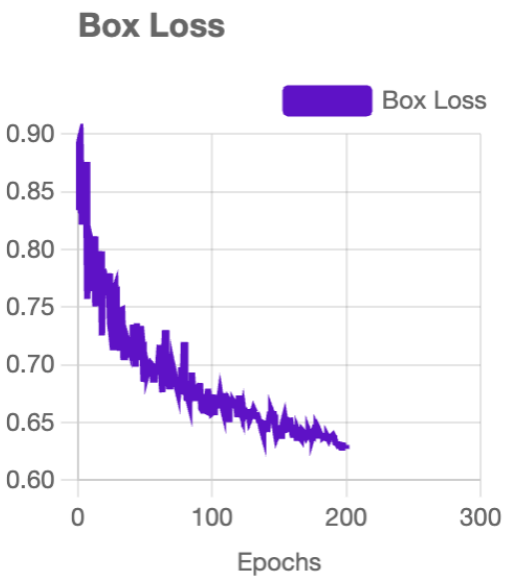
研究結果

1.模型訓練結果

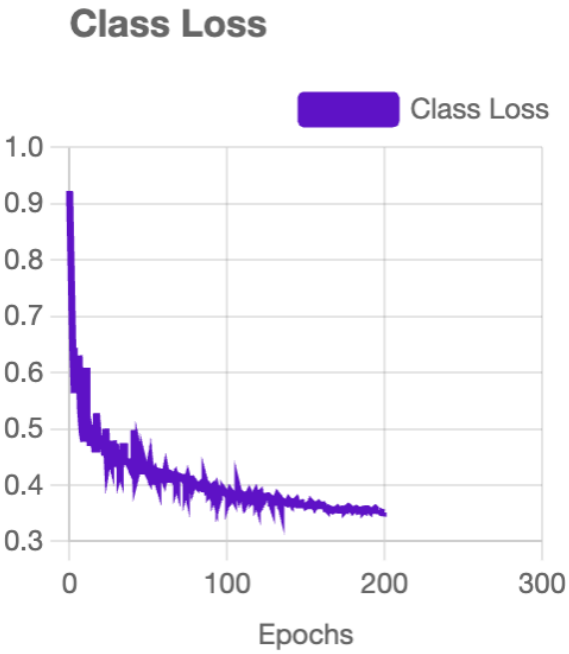
模型訓練結果的各指標與準確率如下列圖片所示。



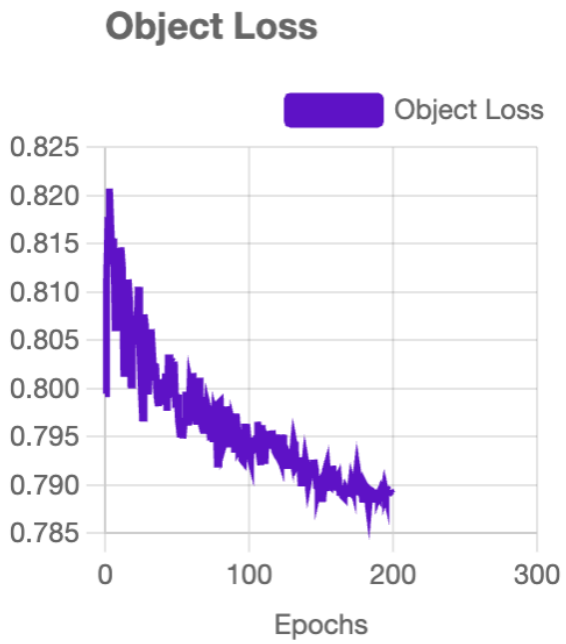
圖四、mAP



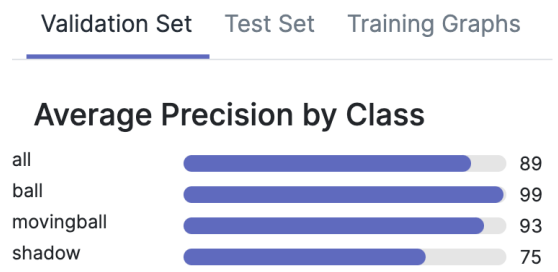
圖五、Box Loss



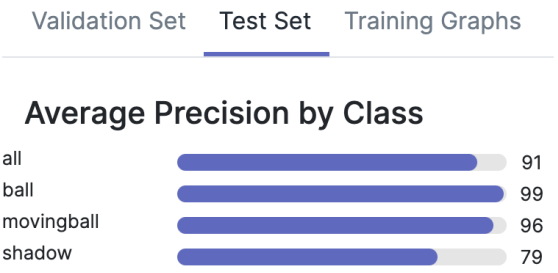
圖六、Class Loss



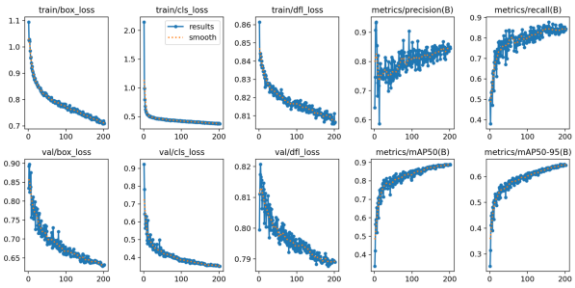
圖七、Object Loss



圖八、每個類別的精確度（驗證集）



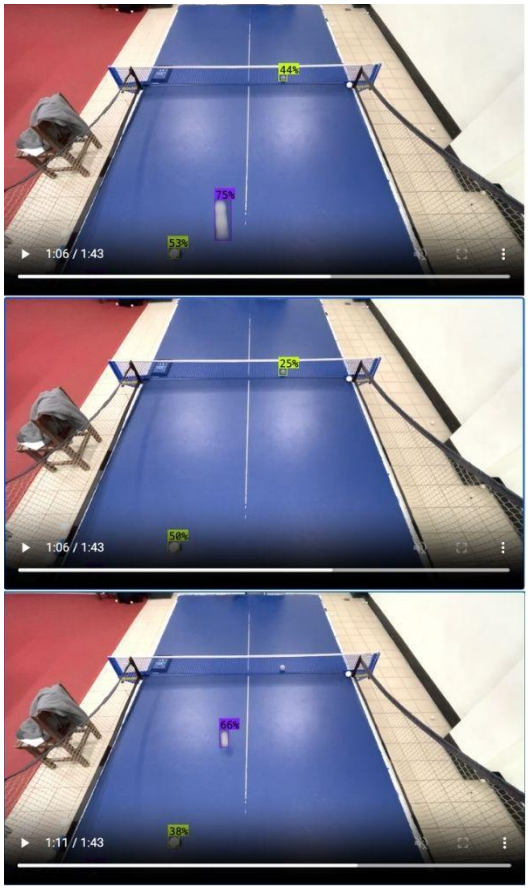
圖九、每個類別的精確度（測試集）



圖十、其他損失函數與指標

2. 模型表現

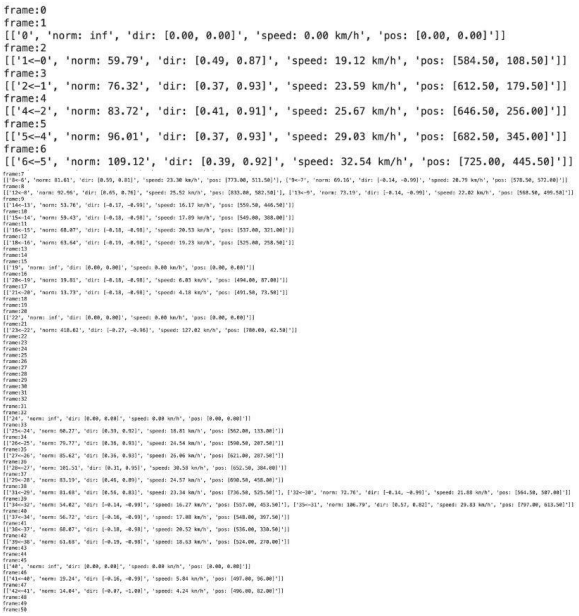
模型訓練好後，我們可以透過Roboflow的網頁將想要測試的影片去進行測試，它會實時地更新預測結果與預測信心，範例如圖十一所示。



圖十一、模型表現測試

3. 演算法結果

我們使用python語言來撰寫我們判斷移動中的球與計算速度的演算法，輸出結果會顯示抓到的正在移動中的球的一些重要數值，如：球的ID與來源ID、移動方向、速度。輸出結果範例如圖十二所示。



圖十二、演算法輸出結果範例

結論

透過本研究，我們成功使用物體識別做到乒乓球的移動捕捉，並搭配演算法計算出乒乓球的移動速度、移動方向以及移動路徑。本研究所蒐集的影片資料高達數十部，並藉由切割方式取得數千張圖樣，透過 Roboflow 的訓練可達到實時的檢測，並最大可能的避免漏抓的情況。本研究成功建構出一套系統，能讓使用者以投喂影像的方式了解到球的運動形式，藉此在不依賴肉眼捕捉的情況下做出合理的得分判斷。該研究希望能以此協助運動員的訓練，幫助打造良好的乒乓球練習環境，並拓展至諸如奧運、世運等國際賽事上，以此對該項運動帶來發展。

參考文獻

[1] S. Triamlumlerd, M. Pracha, P. Kongsuwan and P. Angsuchotmetee (2017) "A table tennis performance analyzer via a single-view low-quality camera"

[2] J. -R. Chang, C. -J. Wang, Z. -K. Wei, C. -J. Lu and H. -Y. Lin (2023) "A Research Structure of Big Data Analysis and Application for Table Tennis Match Tactics Based on Computer Vision"

[3] M. Esmael, H. Ahmad, K. AlKouh, M. Almohammad, T. Jamal and O. Accouche (2023) "Spin-Ninja Table Tennis Training Robot"

[4] Bhuvana J, T.T. Mirnalinee, B. Bharathi, Jayasooryan S, Lokesh N N (2021) "YOLOv5for Stroke Detection and Classification in Table Tennis"

[5] Kiran (2018) "Training an object detector to track a table tennis ball"

[6] Tao Ning, Changcheng Wang, Meng Fu & Xiaodong Duan (2023) "A study on table tennis landing point detection algorithm based on spatial domain information"

[7] Wenjie Li, Xiangpeng Liu, Kang An, Chengjin Qin and Yuhua Cheng (2023) "Table Tennis Track Detection Based on Temporal Feature Multiplexing Network"