

基於影像處理與深度學習的棒球進壘點追蹤系統之研究

林毅冠¹、林奕璿¹、陳明僑^{1*}、陳秀惠^{2、3}

¹ 國立臺東大學資訊工程學系、² 國立臺東大學體育學系、³ 國立臺東大學運動與競技學程

摘要

近年來，人工智慧與深度學習的相關應用蓬勃發展，再加上棒球是台灣的國球，最近的 U23 棒球比賽剛在台灣圓滿結束，基於能夠協助教練與選手，進而提升棒球水準，本研究將相關資訊技術應用於解決棒球相關問題。隨著科技進步，運動員的競爭也越趨激烈，人們可以透過科技來幫助運動員訓練的更到位，而本研究開發一套進壘點追蹤系統，能夠在賽事影片中，藉由偵測人體部位位置，自動生成好球帶並捕捉投手投出的棒球，進而產生球路軌跡，且在捕手接到球後即時顯示進壘點，以利相關人員進行賽事分析及選手能自我調整，來達到減少人力、時間成本的目的。本系統主要分成兩部份：(1)運用 Mediapipe Pose 偵測人體關鍵點，並根據好球帶定義捕捉打擊者肩、臀部、膝部以及進行本壘板偵測來生成好球帶，(2)以 YOLOv5 深度學習模型偵測投手投出的棒球，並且將棒球的座標記錄下來，完整呈現出投手的投球軌跡，從實驗結果數據顯示，本壘板的平均準確率是 76.35%，而棒球軌跡偵測的平均準確率是 62.36%。選手於比賽結束後，可藉由本系統回顧比賽期間的表現狀況，教練亦可透過科技的輔助在賽後更容易指導選手，不受時間、地點等因素限制，讓教練和選手擁有更方便、簡潔的資訊來提升自身的能力。

關鍵字：深度學習、影像處理、物件偵測、姿態偵測。

林毅冠，國立臺東大學資訊工程學系。E-mail: 10811138@gm.nttu.edu.tw

林奕璿，國立臺東大學資訊工程學系。E-mail: 10811139@gm.nttu.edu.tw

陳明僑(通訊作者)，國立臺東大學資訊工程學系助理教授。E-mail: joechen@nttu.edu.tw

陳秀惠，國立臺東大學體育學系教授兼運動與競技學程主任。E-mail: chensh1@nttu.edu.tw

The Study of Baseball location Tracking System Based on Image Processing and Deep Learning

Yi-Kuan Lin, Yi-Syuan Lin, Ming-Chiao Chen, Chen, Hsiu-hui

Abstract

With the rapid development of Artificial Intelligence (AI) and Deep Learning, which can help us to detect the baseball and generate strike zone. This study proposes a pitch location tracking system. The proposed system generates a strike zone automatically by detecting the body keypoints and captures the baseball to show the trajectory in the video, and displays the pitch location immediately after the catcher catching. This system includes two parts:(1) Using Mediapipe Pose model to detect body keypoints, and capture the batter's shoulders, hips and knees, and also detect home plate, and then generate a strike zone, (2) Using YOLOv5 deep learning model to detect the baseball thrown by the pitcher, and display the trajectory of baseball. The experimental results show the accuracy ratio of home plate detection is 76.35%, and the accuracy ratio of baseball trajectory detection is 62.36%. This study makes coaches and players reviewing their performance after the game and improves their capabilities efficiently.

Keyword: Deep Learning, Image Processing, Object Detection, Pose Detection

Yi-Kuan Lin, Undergraduate Student, Department of Computer Science and Information Engineering,
National Taitung University. E-mail: 10811137@gm.nttu.edu.tw

Yi-Syuan Lin, Undergraduate Student, Department of Computer Science and Information Engineering,
National Taitung University. E-mail: 10811139@gm.nttu.edu.tw

Ming-Chiao Chen (Corresponding Author), Assistant Professor, Department of Computer Science and
Information Engineering, National Taitung University. E-mail: joechen@nttu.edu.tw

Chen, Hsiu-hui, Chair, Department of Physical Education and Bachelor Program of Athletic Performance,
National Taitung University. E-mail: chensh1@nttu.edu.tw

壹、簡介

隨著運動科學的蓬勃發展，棒球比賽的數據收集與分析越來越受到重視，如何善用數據以提升自我訓練效率、幫助教練技戰術的下達、情報蒐集等，已成為一大研究課題。棒球是一項團隊運動，其中更以投打對決為一大看點，投打對決中主要包含投手、打者等元素，而棒球比賽紀錄相較於其他球類運動而言，相對的更加繁瑣與複雜，若能透過影像處理與深度學習等技術之協助，將能夠達到經濟又有效率的紀錄效果。

棒球比賽中所蘊含的資訊量相當繁複，一場比賽中投手的總投球數就高達上百顆。球場的不同及拍攝角度、球衣顏色、光線明亮、天氣狀況等諸多變因也會影響到球體的偵測；綜合上述，本研究運用 YOLOv5 深度學習模型偵測投手投出的棒球，並透過 Mediapipe Pose 人體關鍵點偵測與本壘板偵測模型創建好球帶，最終達到追蹤棒球進壘點之目的。

貳、文獻探討

一、深度學習

物件偵測是電腦視覺中的主要任務之一，在 2015 年，由 Joseph R.和 Santosh D. 等人(Joseph et al., 2015)提出 YOLO(You Only Look Once)，將物件偵測重新定義為回歸問題，並在單個神經網路中執行，因此讓物件偵測這塊領域在最近十年間拓展至更耀眼的成就，到目前為止，結合絕佳的電腦視覺研究，YOLO 已經更新到第五版。

Benjumea A.和 Teeti I.等人(Benjumea et al., 2021)提出了名為 YOLO-Z 的物件偵測模型，此模型是以 YOLOv5 為基底，且可偵測 YOLOv5 模型上無法偵測到的物件，目的是為了讓自駕車能夠在路上偵測更小的障礙物，讓車子能夠提早在遠處就能計算後續的路線，達到最終的目的，而依據實驗結果顯示此模型和 YOLOv5 相比，mAP 提升了高達 6.9%，且速度也比原本快了 3ms。

Nepal U.和 Eslamiat H.等人(Nepal & Eslamiat, 2022)提出了在城市中無人機的狀況檢測及排除方法，並利用 YOLOv3(Joseph & Ali, 2018)、YOLOv4(Alexey B. et al., 2020)及 YOLOv5(Glenn Jocher, 2020)模型去偵測城市中的障礙物，判斷哪個地方是合適的降落處，且同時比較這三種版本對障礙物的解讀速度，很明顯的 YOLOv5 在各方面完全的有優勢，最終能將無人機引導到偵測出的安全位置降落，得以讓維修人員檢查無人機的狀況。

Bazarevsky V.等人(Bazarevsky et al., 2020)提出了 Mediapipe Pose，其為姿態偵測的輕量級卷積神經網路架構，在推理過程中，該網路生成 33 個人體關鍵點。主要貢獻包含一個新穎的姿態偵測解決方案和一個使用熱圖和關鍵點座標回歸的輕量級人體姿態偵測神經網路。

二、影像處理

Guéziec A. (Guéziec, 2002)透過 ESPN 運動頻道所釋出的棒球比賽影片中，追蹤投手的投球過程作分析，其中以速度和影像處理精確度為主軸，提高影像處理精確度並克服以下條件：物體的位置變化(如球、球員的移動)及色彩的變化(如陽光、球場照明設備帶來的光影變化及球衣、場地背景的不同)等，為此使用計算機生成的圖形來創建帶陰影的半透明框，該框為觀看者概述了打擊區域的邊界，並與計算飛行軌跡的算式融合，用以追蹤每次的投球過程。

Chakraborty B.等人(Chakraborty & Meher, 2012)自籃球影片中，基於分析籃球投球軌跡，以分析球員的投球過程。為此提出了一種追蹤球的方案，此方案以分割運動對象(球、球員等)、形狀及物體圓度以偵測球的位置，並使光照條件與周圍環境的影響降低，以此條件總和，判斷籃球自出手後的運行軌跡，推算球的出手角度與速度(此兩者為影響投球準確度的重要因素)，進而進行數據統計與分析。

Li C. C.等人(Li et al., 2010)鑑於每次投球在速度、軌跡、加速度和形狀方面都有其自身的特徵，提出了一種棒球投球類型的識別方案，藉此分析球運行軌跡的時空特性。主要方式為使用隱藏式馬可夫模型(HMM, Hidden Markov Model)對投球類型的時間行為進行模型建置，並為了提取軌跡的加速度特徵，採用貝葉斯(Bayesian)決策對特徵進行分類，以進一步提高識別精度。

Shum H.等人(Shum & Komura, 2004)提出一種提取和計算錄像剪輯中投手投球的3D 軌跡的方法，以單一視角的影片片段，使用基於動態規劃的全局搜尋方法以查找球的軌跡。在投手投球過程中，利用所有幀的信息來找出球，且還可以從打者的角度於虛擬環境中建置3D 棒球移動軌跡，投手可以通過觀察自己的軌跡提高自己的技術，而打者可以查看各種投手的投球軌跡以輔助訓練。

Chen H. T.等人(Chen et al., 2012)提出了基於隱藏式馬可夫模型(HMM, Hidden Markov Model) 的擊球事件探索系統。在棒球影片中，檢測棒球比賽場地中的空間，例如場地線、四個壘包、土壤、草地等，並通過直方圖來定義主要顏色。運用 RGB 和 HSI (Hue Saturation Intensity) 顏色空間中的直方圖。HSI 顏色空間中的色調值足以定義主要顏色，原因有 2 點：(1)觀察到色調值即使受到光線的干擾，但在同一場比賽中仍相對穩定，以及(2)草和土壤的顏色在色調直方圖呈現突出的峰值，因此能良好的進行辨別。

Chen H. T.等人(Chen et al., 2010)提出利用比賽的棒球影像偵測好球帶，透過對於好球帶的定義，需找出本壘板，以及打者的肩部、臀部及膝蓋等特徵的相對位置。本壘板部分運用色彩分佈(Color Distribution)中找出本壘板的顏色與符合本壘板形狀的物體去找出其所在位置，好球帶的寬度便可得知。再者建立打擊者區域的像素差異圖(Moving Pixel Image, MPI)，找出打擊者區域的極右像素差異圖(Rightmost Moving Pixel Image, RMPI)及極左像素差異圖(Leftmost Moving Pixel Image, LMPI)，將 RMPI 與 LMPI 中未連通的像素依照縱軸從上至下連結，成為一個打擊者右半邊的輪廓，並對

輪廓做平滑化處理，可以求出最小負曲率(Negative Minimum Curvature, NMC)對應到膝蓋，及最大正曲率(Positive Maximum Curvature, PMC)對應到臀部，而透過輪廓計算出在橫軸顏色差異度大的地方，(因為球衣和臉造成顏色差異度變大)，及可判斷為肩部位置，而找出本壘板與打者的肩膀、臀部、膝蓋相對位置後便可找出好球帶。

Wong K. C. P.等人(Wong & Dooley, 2010)提出一個在桌球運動錄像中，自動檢測與追蹤球的系統，進而提高比賽判決的精準性。過程中使用一套算法，可於影片中自適應地利用時間與空間的信息，以高速物體的運動為特徵，為避免相似物體形狀相似、物體模糊與其他物體阻擋等造成判斷失準，使用卡爾曼濾波器(Kalman)和增量貝斯演算法(Bayesian)來追蹤球在真實比賽場景中的情況。

Myint H.等人(Hnin et al., 2015)提出了一種有效的檢測和追蹤策略，其使用低成本的立體視覺形式呈現，利用兩個單視圖拍攝的立體影片中的網球，其為運用分割、檢測、追蹤的方法來尋找之。此外，為了解決球體可能被遮住的問題，還提出了一種具有校正機制的立體視覺系統，分成四個主要模型：(1)結合應用顏色閾值和移動偵測的分割模型、(2)用於軌跡偵測的二階運動模型、(3)球體特徵檢測模型和(4)自我校正模型，來尋找乒乓球在移動過程中可能被其他物件擋住或顏色相近困擾。

Chu W.等人(Chu et al., 2006)基於提供一些有助於提高投手能力的線索，提出了一種基於卡爾曼濾波器的方法，可以根據單視角投球序列追蹤球的軌跡。此方法用於從電視頻道，捕獲或從網路上下載的投球序列。為了驗證檢測到的球的軌跡，過程中的應用如：著名的 K-zone 以追蹤俯仰視角的球的軌跡，以偵測球的位置、速度、旋轉軸及旋轉等增加球運行軌跡的精確度，最終以基於棒球物理模型研究出了球軌跡的特徵，提出了球的軌跡提取和球位置檢測的有效性。

Yen W. C.等人(Yen et al., 2019)提出了棒球運動員行為識別系統。訊號來自可穿戴傳感器和 Kinect 相機。當身體部位被遮擋或在相機中有運動模糊效果時，傳感器的傳感信號可以補償識別能力。此外，通過分析從異構傳感器中獲得的特徵，結果可以顯示在具有增強功能的智能玻璃上現實展示。在這個原型中，玩家的行為可以由教練監督，以協助建議過程場上和場外的棒球比賽環境。

參、研究方法

如圖 1 系統架構圖所示，本系統分為 4 個部分：影像前處理、棒球偵測、好球帶生成以及進壘點判斷，系統功能與實作方法的詳細說明如下：

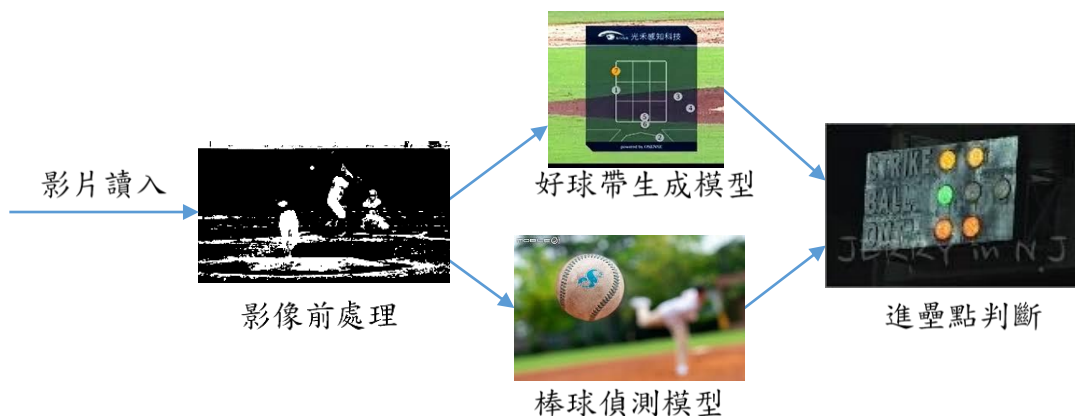


圖 1 系統架構圖

3.1 影像前處理

由於投手投球畫面的攝影機都是固定不動的，不會受到攝影機移動的影響，因此適合利用前後影像差異來擷取畫面中所有移動中的物件。

3.1.1 灰階化

公式 1 是將彩色影像轉換成灰階影像。灰階值(Gray)為公式計算得出的值，為了減少資料的複雜度，以利往後步驟的處理。

$$Gray = 0.299 * R + 0.587 * G + 0.114 * B \quad (\text{公式 1})$$

R、G、B 分別代表圖中紅色、綠色、藍色的值。

3.1.2 連續影像差異

透過每兩個連續灰階影像像素值帶入公式 2，像素值相差較明顯的設為 255，而較不明顯的設為 0，擷取兩張影像中的移動物件。 $I_{FDI_n}(x,y)$ 代表連續影像差異的像素值，n 代表第 n 張連續影像，T 為影像差異的閾值， $I_{GRAY_n}(x,y)$ 代表當前灰階影像之像素值， $I_{GRAY_{n-1}}(x,y)$ 代表前一張灰階影像之像素值。

$$I_{FDI_n}(x,y) = \begin{cases} 255, & \text{if } I_{GRAY_n}(x,y) - I_{GRAY_{n-1}}(x,y) > T \\ 0, & \text{otherwise} \end{cases} \quad (\text{公式 2})$$

3.2 好球帶生成模型

本模型主要分為兩個部分，藉由人體姿態偵測捕捉關鍵點座標取得其高度，接著運用本壘板偵測找出本壘板並取得其寬度，最後根據高度與寬度生成一個好球帶。

3.2.1 人體姿態偵測

採用 Mediapipe Pose 中的 BlazePose 來進行人體姿態偵測，其具有總計 33 個人體關鍵點(如圖 2)。在機器學習推理過程中採用檢測器-追蹤器管道(如圖 3)，管道由一個輕量級的身體姿態檢測器和一個姿態追蹤器網路組成。檢測器首先定位當前幀上的人的存在及感興趣區域(Region of Interest, ROI)，追蹤器隨後使用 ROI 裁剪幀作為輸入，來預測 ROI 內的關鍵點和分割掩碼(Segmentation Mask)。在影片中，檢測器只在第一幀上或是當追蹤器無法再識別出前一幀中的人體姿勢時運行。對於第一幀後的後續幀，管道將從前一幀的關鍵點導出 ROI。

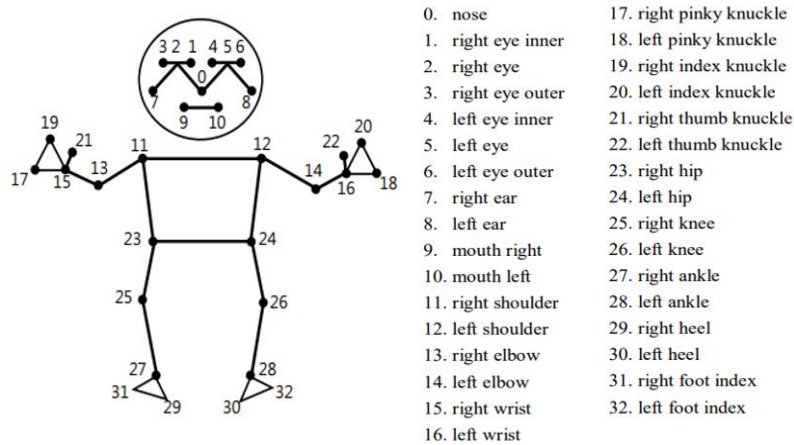


圖 2 人體關鍵點示意圖

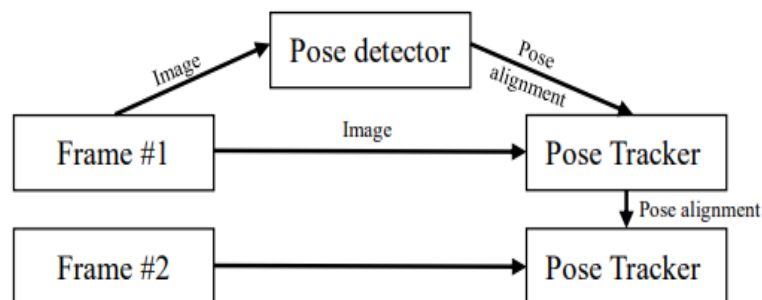


圖 3 檢測器-追蹤器管道

神經網路架構採用整合熱圖(Heat Maps)、偏移(Offset)和回歸(Regression)的方式(如圖 4 所示)。利用熱圖和偏移只在機器學習訓練階段，並在運行推理之前從模型中刪除相應的輸出層。在訓練過程中，首先採用熱圖和偏移損失訓練網路的左塔與中心

(圖 4 a 部分與 b 部分)，然後移除熱圖輸出並訓練回歸編碼塔(圖 4 c 部分)，從而有效的運用熱圖來監督輕量級嵌入。

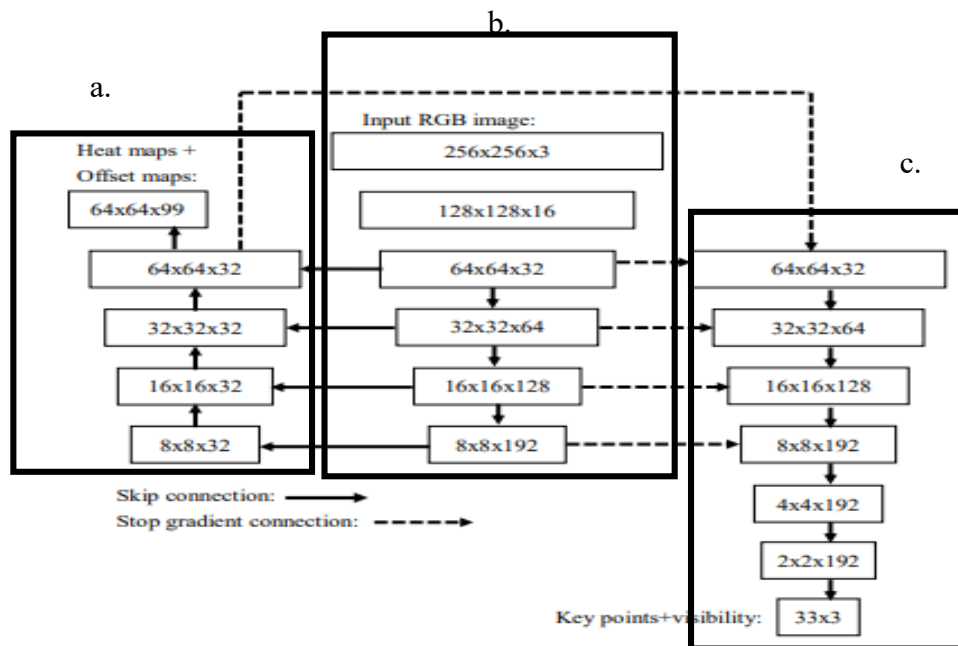


圖 4 神經網路架構

我們將採用 Mediapipe Pose 捕捉出打擊者的肩部、腰部、膝部，再依好球帶定義(圖 5)取肩部到腰部之間的中線為好球帶上緣，膝蓋為好球帶下緣。



圖 5 好球帶定義

3.2.2 本壘板偵測

如圖 6 所示，在棒球影片中，因拍攝角度關係，本壘板以接近白色的短水平線段顯示(如圖 6 a)，主要分布在幀中心區域，本研究中採用一個本壘板偵測法，首先去除掉低強度的像素，然後只保留幀內中心四分之一區塊的像素(如圖 6 b 所示)。再根據棒球運動規則中定義的規格，棒球寬度是 2.9 英吋，本壘板寬度是 17

英吋，本壘的幀內寬度由棒球的幀內直徑按比例推估:2.9/17(棒球規格/本壘板規格)，最後將寬度最接近本壘板估計幀內寬度的短水平物體視為本壘板。(如圖 6 c)



圖 6 a.原始影像 b.在幀中心的高強度像素 c.檢測本壘板

3.3 棒球偵測模型

將影片輸入 YOLOv5 模型執行，並且將每幀偵測到的棒球留在畫面上，形成一個棒球軌跡，最後再分析棒球進壘點。如圖 7 架構圖所示，採用 YOLOv5 模型來偵測棒球，其主要分為四大部分:輸入端、Backbone、Neck、以及 Prediction。架構圖內相關名詞說明如下：

- Focus 結構為影像經切片及卷積後的過程
- CBL 由卷積層(Convolutional Layer)、BatchNorm 和 Leaky_relu 組成
- CSP1 由 CBL、Res unit、卷積層以及 Concat 組成，其中 Res unit 為殘差(Residual)結構，其過程經由 CBL 後再將張量相加
- CSP2 由卷積層、CBL 以及 Concat 組成
- SPP 為池化後再經 Concat 組成

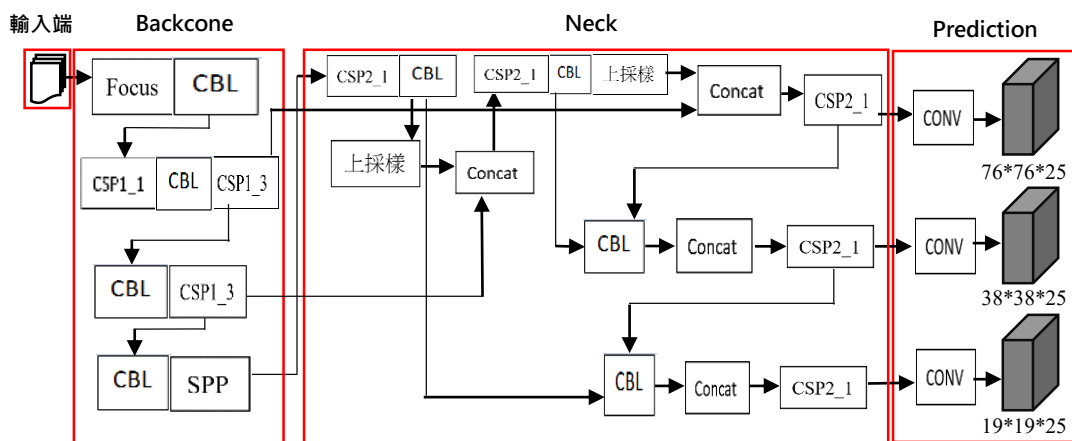


圖 7 YOLOv5 偵測模型架構圖

3.3.1 輸入端

在執行 YOLOv5 時，由於每個圖片或影片的長寬不一致，因此會將原始影像縮放到統一長寬(416x416)後，加上計算好的黑框寬度，再執行偵測的動作。以一個長寬為 800x600 的影像為例，以 416 除以此影像長與寬，得到 0.52 和 0.69 兩個數值，再以最小值(0.52)乘以原始影像長與寬，各得到 416x312 的影像，將此長與寬相減(416-312)後得到 104，再將 104 除以 32(因為 YOLOv5 的網路經過 5 次採樣， 2^5 是 32)取餘數得到 8，將此結果平分至影像寬度(上下各增加 4 的黑框寬度)，得到 416x320 的影像，則是此次進行 YOLOv5 模型偵測的影像長與寬。

3.3.2 Backbone

此部分是在 YOLOv5 新增的，主要是使圖片在採樣的過程中，減少了一定的模型計算量，為後續的圖片保留了更完整的特徵採取，並且不會有任何資訊丟失，依舊仍有非常高的精確度。此過程是將三維陣列的圖像經過切片後，再經過多次的卷積核操作，得到最終的三維特徵圖。

3.3.3 Neck

此部份分為 FPN(Feature Pyramid Networks)和 PAN(Pixel Aggregation Network)兩大結構，且從 YOLOv3 就存在，到了 YOLOv5 也被優化許多。從下圖 8 架構圖來看，FPN 先採用上採樣的方式，將高層的特徵資訊傳遞，得到預測的特徵圖。而 PAN 則是從 2018 年的圖像分割(Image Segmentation)衍伸而來，主要是提高特徵提取的能力，採用的是下採樣的方式。

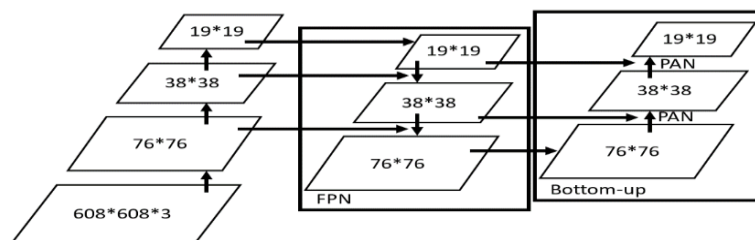


圖 8 FPN 與 PAN 架構圖

3.3.4 Prediction

YOLOv5 中，採用此方法作為 Bounding box 目標偵測的損失函數，和過往用的 DIOU_Loss 公式一樣，不過這次將預測框和目標框的長寬比都列入公式 3 裡。CIOU_Loss 中的重要元素包含： IOU 代表真實框(I)與預測框(P)重疊面積的百分比、 p 代表預測框距離中心點長度， c 代表標籤框距離中心點長度，及 v 代表

$$\frac{4}{\pi^2} \left(\arctan \frac{w_I}{h_I} - \arctan \frac{w_P}{h_P} \right)^2。$$

$$CIOU_Loss = 1 - CIOU = 1 - \left(IOU - \frac{p^2}{c^2} - \frac{v^2}{(1-IOU)+v} \right) \quad (\text{公式 3})$$

肆、結果與討論

一、實驗數據集

本研究的實驗數據採用本校運動競技學士學位學程的比賽影片，總共五場比賽，分別是台東 VS 中信(第一場)、台東 VS 中信(第二場)、台東 VS 文化(第一場)、台東 VS 嘉大(第一場)、台東 VS 嘉大(第二場)，每一場比賽的用球數分別為 200 球、127 球、186 球、191 球、70 球，如表 1 所示。

二、好球帶生成

偵測結果如表 1 所示，為本壘板偵測之準確率。從 3.2 節好球帶生成模型，捕捉人體姿態關鍵點中的肩部、腰部、膝部，經由計算取得好球帶的高度，本壘板寬度則由影像處理偵測方法取得，即為好球帶寬度，以所計算出的寬高為基準，最終自動生成出好球帶。

如表 1 中所示，其中場次 1.台東 VS 中信(第一場)、場次 3.台東 VS 文化(第一場)、以及場次 5.台東 VS 嘉大(第二場)本壘板偵測皆取得不錯的結果，準確率都達到 92% 以上，而反觀，場次 2.台東 VS 中信(第二場)以及場次 4.台東 VS 嘉大(第一場)的準確率不甚理想，由於棒球場地的環境因素複雜多變，分析其是因球場光線、影片噪音、本壘被紅土弄髒等原因所導致，綜觀這五場比賽下來，本壘板偵測準確率平均為 76.35%。

表 1. 本壘板偵測準確率

場次	比賽隊伍	比賽用球數	準確率
1	台東 VS 中信(第一場)	200	94%
2	台東 VS 中信(第二場)	127	33.07%
3	台東 VS 文化(第一場)	186	94.08%
4	台東 VS 嘉大(第一場)	191	63.35%
5	台東 VS 嘉大(第二場)	70	92.85%

三、棒球偵測與追蹤

此部分乃採用 YOLOv5 偵測模型，來辨識表 1 中第一場次及第三場次的比賽，系統會先將影片裁成一幀幀的影像進行偵測，偵測後結果如圖 10 所示，偵測框上會顯示所偵測物件的名稱，名稱後包含此物件的配對權重。



圖 10 YOLOv5 偵測之結果

由於 YOLOv5 可偵測到的物件高達 30 多種，因此我們將不是棒球的其它物件過濾掉，如下圖 11 所示，過濾後僅保留被標記的棒球，使整個畫面顯示較為清晰。



圖 11 經過濾後僅保留棒球之結果

在偵測棒球的過程中，YOLOv5 原先設定的偵測範圍為整張影像，因拍攝棒球影像的鏡頭是固定的，為避免與減少非棒球物件之干擾，本研究將偵測範圍的高度縮小為畫面的 2/5 到 4/5，以及寬度縮小為 1/8 到 7/8，如圖 12 所示。



圖 12 YOLOv5 偵測區域縮小示意圖

如表 2 所示為棒球偵測準確率之結果，兩場比賽平均偵測準確率各為 58.6%及

68.42%，右打者打擊時平均偵測準確率各為 80.87%及 74.92%，左打者打擊時平均偵測準確率為 20.7%以及 36.48%，其中左打者影片兩者結果都較不佳，分析其原因係因棒球為白色的物件，影片中的呈現的棒球顏色會和白天的背景、球員的球衣等因顏色相近，而導致模型無法偵測到球的正確位置；而且球在飛行過程中，如果是左打者打擊時的影片，較容易和棒球重疊，而導致其偵測失敗，如圖 13，由於第二場中左打者影片數量較少，故第二場比賽平均偵測結果較第一場平均為佳。然而左打者的影片中亦有不錯的結果，最高偵測準確率可達 80%，甚至會比右打者影片結果還要好，所以棒球在飛行過程中，和其他物件是否重疊，是影響影片偵測準確率的關鍵。

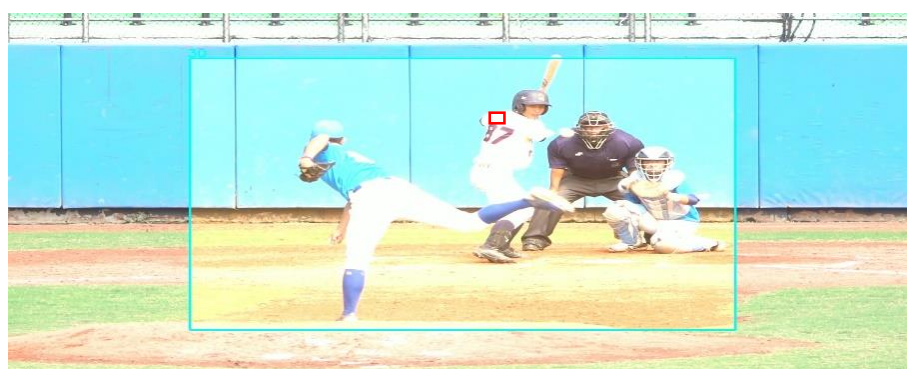


圖 13 YOLOv5 無法偵測和球衣重疊的球

表 2. 棒球偵測準確率

棒球比賽 平均準確率	台東 VS 中信(第一場)	台東 VS 中信(第二場)	二場比賽結果
整場比賽	58.6%	68.42%	62.36%
右打者打擊影片	80.87%	74.92%	78.2%
左打者打擊影片	20.7%	36.48%	24.19%

四、進壘點追蹤

結合以上好球帶生成與棒球偵測與追蹤，可實作出最終進壘點之結果，如圖 14 所示，在影片中呈現好球帶及棒球軌跡並追蹤其進壘點。



圖 14 軌跡與正確進壘點位置

如圖 15 所示為系統偵測失敗的例子，系統判定之進壘點與正確進壘點不符，分析原因是由於棒球顏色與打者球衣顏色相近而造成重疊，導致影響最終進壘點判定的結果。



圖 15 進壘點位置偵測失敗

伍、結論

一、好球帶生成自動化

本研究採用 Mediapipe Pose 人體姿態偵測模型捕捉打者人體關鍵點，並利用影像處理技術偵測影片中的本壘板，其中，在五場比賽中三場比賽的本壘板偵測準確率都高達 92%，而本壘板偵測準確率平均為 76.35%。最終將上述所得到的資訊，於影片中自動生成出好球帶，以利相關紀錄人員能藉著本研究的系統進行紀錄，取代傳統肉眼方式判別的紀錄方法。

二、以深度學習模型偵測棒球

本研究此部分採用 YOLOv5 物件偵測模型來偵測棒球，實驗數據主要分成兩個部分，分別為右打者打擊影片以及左打者打擊影片，右打者打擊影片準確率為 78.20%，左打者打擊影片準確率為 24.19%。從此次實驗來看，背景亮度、拍攝角度、影像的解析度等條件有很大的相關性，且和偵測目標顏色相近的物件也和準確率息息相關，最終將一幀幀移動中的棒球利用標記形成棒球軌跡，讓選手以及相關人員透過此研究來增強個人的棒球能力。

陸、引用文獻

Alexey, B., Chien-Yao, W., & Hong-Yuan M. L., (2020). "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv:2004.10934v1*.

Benjumea, A., Teeti, I., Cuzzolin, F., & Bradley A. (2021, 22 Dec). "YOLO-Z: Improving small object detection in YOLOv5 for autonomous vehicles," *arXiv: 2112.11798v1 [cs.CV]*.

Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., & Grundmann, M. (2020, 17 Jun). "Blaze Pose: On Device Real-time Body Pose Tracking," *arXiv: 2006.10204v1*

[cs.CV].

- Chakraborty B. & Meher S. (2012). "A Trajectory-Based Ball Detection and Tracking System with Applications to Shot-type Identification in Volleyball Videos," *Proceedings of IEEE International Conference on Signal Processing and Communication*.
- Chen, H. T., Chou, C. L., Tsai, W. C., Lee, S. Y., & Lin, B. S. P. (2012). "HMM-based Ball Hitting Event Exploration System for Broadcast Baseball Video," *Journal of Visual Communication and Image Representation, Vol. 23, pp. 767 - 781*.
- Chen, H. T., Tsai, W. J., & Lee, S. Y. (2010). "Contour-based Strike Zone Shaping and Visualization in Broadcast Baseball Video: Providing Reference for Pitch Location Positioning and Strike/Ball Judgment", *Multimed Tools Application, pp. 239 - 255*.
- Chu, W., Wang, C., & Wu, J. (2006). "Extraction of Baseball Trajectory and Physics-based Validation for Single-view Baseball Video Sequences," *Proceedings of 2006 IEEE International Conference on Multimedia and Expo, pp.1813-1816*.
- Glenn, J., (2020 May). "YOLOv5", <https://github.com/ultralytics/yolov5>.
- Guéziec A. (2002). "Tracking Pitches for Broadcast Television," *Proceedings of IEEE International Conference on Computer, Vol. 35, No. 3, pp.38-43*.
- Hnin, M., Patrick, W., Laurence, D., & Adrian, H. (2015 May). "Tracking a Table Tennis Ball for Umpiring Purposes", *Proceedings of Conference on Machine Vision Applications (MVA2015) on Fourteenth IAPR International, pp. 18 - 22, Tokyo, Japan*.
- Joseph, R., & Ali, F. (2018). "YOLOv3: An Incremental Improvement," *arXiv: 1804.02767*.
- Li, C. C., Lin, C. W., & Yu, J. Y. (2010, July). "Statistical Pitch Type Recognition in Broadcast Baseball Videos," *Journal of Computers, pp. 15 - 23, Vol.21, No.2*.
- Nepal, U., & Eslamiat, H. (2022, 8 Jan). "Comparing YOLOv3, YOLOv4 and YOLOv5 for Autonomous Landing Spot Detection in Faulty UAVs," *Sensors 2022, 22, 464*.
- Shum, H. & Komura, T. (2004). "A Spatiotemporal Approach to Extract the 3D Trajectory of the Baseball from a Single View Video Sequence," *Proceedings of 2004 IEEE International Conference on Multimedia and Expo (ICME 2004), pp. 1583 - 1586 Vol.3*.
- Wong, K. C. P., & Dooley, L. S. (2010). "High-Motion Table Tennis Ball Tracking for Umpiring Applications," *Proceedings of IEEE International Conference on Signal Processing, pp. 2460 -2463*.
- Yen, W. C., Fang, C. C., Sun, S. W., Hua, K. L., & Shih, H. C. (2019). "Baseball Player Behavior Recognition System using Multimodal Features with an Augmented Reality Display on a Smart Glass," *Proceedings of IEEE International Conference on Multimedia & Expo, pp. 607 - 607*.