

國立東華大學資訊工程系
National Dong Hwa University
109 學年度大學部專題研究報告
109 CSIE Project Report

基於深度學習之桌球事件偵測計分系統



指導教授 Advisor：張意政 教授

專題參與人員 Team Member：蔡景丞
簡廷州
吳侑達

中 華 民 國 110 年 5 月 21 日

國立東華大學資訊工程學系

專題報告原創性聲明

National Dong Hwa University Department of Computer Science and Information Engineering Statement of Originality

本人鄭重聲明：

所呈交的專題報告是在指導老師指導下進行的研究工作及取得的研究成果。除文中已經註明引用的內容外，本報告不包含任何其他個人或集體已經發表或撰寫過的研究成果。對本文的研究做出重要貢獻的個人與集體，均已在文中以明確方式標明。若有違上述聲明，願依校規處分及承擔法律責任。

I hereby affirm that the submitted project report is the result of research under the supervision of my advisor. Except where due references are made, the report contains no material previously published or written by another person or group. All significant facilitators to the project have been mentioned explicitly. Should any part of the statement were breached, I am subject to the punishment enforced by the University and any legal responsibility incurred.

學號	學生姓名	親筆簽名
410714203	蔡景丞	蔡景丞
410714212	簡廷州	簡廷州
410721371	吳侑達	吳侑達

日期 Date : 5 / 21

專題摘要

本專題以日常的桌球比賽影片為基準，透過深度學習的方式，來實現球體追蹤與球體事件偵測，進而實現桌球自動計分系統。(使用者可以使用相對低的成本,體驗到高階套裝系統的部分功能)使用者可以已相對較低的硬體成本，體驗到高階套裝系統的功能。克服了以往需要專業攝影機才能準確地進行球體追蹤，讓使用者能盡情享受桌球的樂趣。

關鍵詞：卷積式神經網路、軌跡追蹤、桌球、電腦視覺

目錄

1. 前言.....	6
2. 研究動機與研究問題.....	6
3. 文獻回顧與探討.....	7
4. 研究方法及步驟.....	8
4.1 硬體與軟體選擇.....	8
4.2 系統簡介與架構.....	8
4.3 球體追蹤網路.....	9
4.3.1 球體追蹤架構.....	9
4.3.2 神經網路輸出與熱度圖.....	10
4.4 事件偵測網路.....	11
4.4.1 事件偵測架構.....	11
4.4.2 機率標記與輸出.....	12
4.5 計分演算法.....	12
5. 結果與討論.....	13
5.1 UI 介面.....	13
5.2 系統訓練與準確度.....	14
5.2.1 球體追蹤訓練與準確率.....	14
5.2.2 事件偵測訓練與準確率.....	16
5.3 總結.....	19
6. 參考文獻.....	19

圖片目錄

圖一、系統架構圖.....	9
圖二、偵測桌球熱度圖.....	11
圖三、計分演算法流程圖.....	13
圖四、使用者介面.....	14
圖五、球體追蹤神經網路之損失值(loss)趨勢圖.....	15
圖六、定位誤差(positioning error)分布圖.....	16
圖七、原始架構神經網路之損失值(loss)趨勢圖.....	17
圖八、原始架構神經網路之損失值(loss)趨勢圖.....	17
圖九、改良架構神經網路之損失值(loss)趨勢圖.....	18
圖十、改良架構神經網路之準確值(accuracy)趨勢圖.....	18

表格目錄

表一、球體追蹤神經網路架構.....	10
表二、事件偵測神經網路架構.....	12
表三、訓練球體追蹤神經網路的關鍵參數.....	15
表四、訓練事件偵測神經網路的關鍵參數.....	16

1.前言

現今存在著各式各樣的球類運動，可分為使用球拍、球門、擊球等類型。我們著重研究球類運動中的一項競技：桌球。近年來，球類運動在賽事分析與賽後檢討中，利用錄下的比賽過程做為戰術分析以及球的落點紀錄是一種常用的方式，但因為桌球的體積小且移動快速，在錄影過程中會容易產生球體影像模糊的問題，使得選手們不易觀察影片中的資訊，造成分析上的困難。雖使用高階的攝影器材能克服球體影像模糊的問題，但高階的攝影器材的成本往往不是一般民眾所能負荷的。

在本次專題研究中，我們結合文獻[1]所提出的 TrackNet 系統與文獻[2]中的 TTNet 系統，以日常的桌球比賽影片為題材進行研究做出一自動計分的系統。利用我們提出的架構來進行球體的定位以及事件偵測，且改善以往需要透過專業攝影機拍攝才能準確分析的困難。本系統只需要以入門手機錄製影片，即可達到準確判斷球體位置與動作。此專題除了可以克服影像模糊及殘影等問題，並且還可以預測被物件遮蔽的桌球位置。利用 TrackNet 追蹤球體位置，並結合 TTNet 的事件偵測技術，在理想情況下改善原本的 TTNet 系統，完成本研究計劃的核心內容，且讓使用者透過較低階的硬體設備，進行桌球球體追蹤並達成桌球事件偵測，做出自動計分的系統。

2.研究動機與研究問題

近年來運動偵測的研究大多藉由影片中球體的運動軌跡來分析比賽的賽況，而在攝影技術的進步下，這些研究的相關應用更加受到重視，例如：分析選手的球路特徵、判斷選手的擊球反應、和觸網球的處理等等。然而，現今多數桌球運動偵測系統遇到的難題有：（1）專業的高速攝影器材過於昂貴，或是需要搭配特殊設計的相機，才可運行的大型系統，都具有相當大架設成本，對一般民眾負擔較大。（2）因為球體體積小，所以球體在畫面中所占的面積較小，因而容易與畫面中相似顏色的物體或背景混淆，產生錯誤偵測(false positive)。（3）由於桌球球桌的範圍較小，球體在相當有限的範圍內移動，且速度相對較快，會造成影像模糊，可能會有錯誤偵測(false negative)發生。

本專題將使用較低階的硬體設備需求的自動桌球事件偵測計分系統，將透過組合現有 TrackNet 與 TTNet 系統解決以上描述遇到的難題。為了解決高速攝影機的成本問題，我們將使用一般入門級手機或初階相機即可拍出的 60fps 影片進行桌球事件偵測。利用 TrackNet 來進行球體定位，藉由一次輸入多張連續的幀，不只可以從中學習球的影像特徵，也可以學習球飛行的軌跡特性，以期達到特徵加成的效果，進而降低錯誤偵測的發生機率。最後根據神經網路計算出的熱度圖來判斷畫面中球體的位置，再利用 TTNet 來進行事件偵測，使用

TurboJPEG 套件來提升資料讀取速度，降低整體系統輸出的負擔。因為桌球的核心事件觸發前後經過時間相當短，就單獨球體彈跳事件而言，發生的間隔僅有 0.25 秒，但一般低速攝影機所提供的影格率僅 60fps。為了符合一般低速攝影機的影格率，我們將輸入連續五幀的影格到事件偵測系統中，並輸出指定事件的發生機率。利用發生機率來判斷當前事件為何，以達到事件偵測的目的。

本專題實現一個基於低速攝影機達成之事件偵測系統，我們將解決現有桌球事件偵測系統的硬體負擔問題，並且為一般民眾提供一個結合球體追蹤與事件偵測之完整桌球系統。

3. 文獻回顧與探討

透過軟硬體來輔助球類運動的技術在近年來有相當大幅的成長，現今已有非常多的球體追蹤技術，但是與事件偵測技術結合的應用卻相對較少。對於不同球類競技的球體追蹤，文獻[3]鷹眼技術證明球體追蹤技術目前已突破圓形球體的限制，可以應用於非圓形球體例如羽毛球，在球體追蹤上已採用各式各樣的技術達到目的。例如文獻[4]中，系統透過去除靜止背景並偵測動態物件以提供直徑大小於 5~60 像素的桌球球體追蹤的功能。在文獻[5]中，系統透過給定的理想環境與條件下創建一個三維模擬空間，使用 ROI(Region of interest)限制區塊並計算光源與球體陰影以取得桌球球體位置。

在我們主要所使用的深度學習技術中，對於球體追蹤的問題與解決已有了相當程度的進展。基於卷積神經網路的技術下，許多已完成的網路架構都具有相似的結構。諸如使用在足球球體追蹤的文獻[6][7]，與運用在桌球球體追蹤之文獻[8]，透過深度學習技術所完成的系統，正確率皆具有 95% 以上，與文獻[4][5]的 80% 正確率相比有著更高的準確度，但上述系統卻都缺乏計分與主要事件的判斷功能，並未結合事件偵測成為一套裝系統。

而事件偵測部分，現已有少數完善系統可以運行，但卻需要搭配特殊要求才能使用。在文獻[9]中雖可準確判定網球球體彈跳時的位置，但卻需要依靠熱感應器才能判定球體彈跳時與場地的摩擦位置。在文獻[10]需透過 OpenCV 進行大量運算，且須提供特殊角度與解析度的攝影機才能成功運行。在文獻[11]中提供了一套完整的計分系統，系統先提供了網球球體追蹤的技術後，再透過取得的球體位置達到事件偵測，系統依據球體的標準黃色來當作分類標準，透過橢圓法線計算得出球體座標，再結合運動員位置後即可進行事件偵測。

在本專題主要參考的技術之一，文獻[1]中的 TrackNet 系統便提供了最基本的球體追蹤技術；TrackNet 以網路公開之網球比賽為主，使用卷積神經網路設計了網球球體位置的追蹤系統。在事件偵測部分，則參考文獻[2]的 TTNNet 系統。

TTNet 為使用者提供了全套的賽事計分系統，但其代價即為高速攝影機與套裝設備的成本。

4. 研究方法及步驟

研究方法與步驟將會按照專題內容分成硬體與軟體介紹、系統的簡介與架構以及相對應實現系統的方法介紹。

4.1 硬體與軟體選擇

硬體：

- Intel(R) Core(TM) i7-9700 CPU @ 3.00GHz
- Nvidia GTX1080Ti
- 1080p 60fps 低速攝影機

環境：

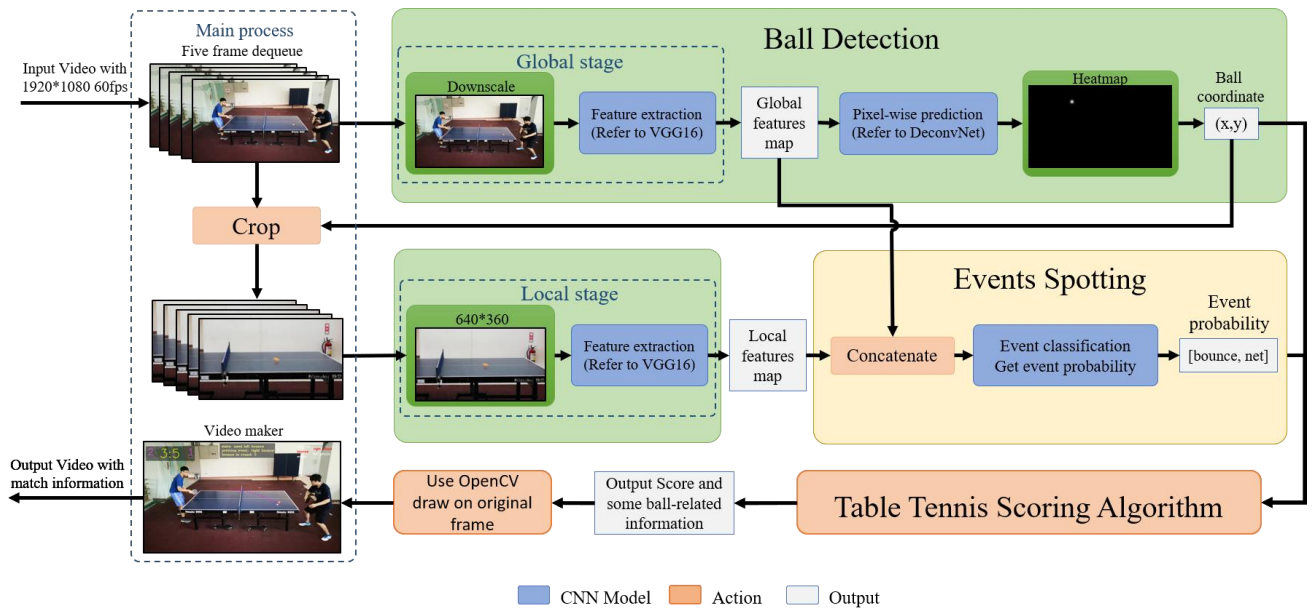
- Ubuntu 18.04
- Tensorflow 2.3.1、OpenCV 4.4.0
- Python3

4.2 系統簡介與架構

本專題的自動桌球事件偵測計分系統，讓使用者無論是平常與球友的對打亦或是正式的比賽時皆可使用，只要使用一般低速攝影機或手機，對比賽進行全程的錄影後輸入至本系統，便可協助進行比賽復盤檢討。本系統由兩個核心部分組成，分別是球體追蹤網路及事件偵測網路。系統會先經過球體追蹤網路來標記出桌球球體的確切位置，接著經過事件偵測網路來判斷桌球球體事件，再根據上述資訊，透過自行設計的計分演算法來判定該場比賽的得分狀況，進而達到自動計分的效果。

當影片輸入至本系統時，首先會將連續五幀畫面的解析度降至固定大小，接著輸入至球體追蹤的神經網路，再透過此神經網路生成用於偵測最後一幀球體的確切座標熱度圖。得到球體的座標後，以球體座標為中心，將原圖裁減成固定大小，再輸入至球體追蹤前 13 層的神經網路。將原圖未裁減與裁減過的畫面分別經過球體追蹤前 13 層神經網路，分別稱為 Global stage 與 Local stage，然後取其特徵圖進行結合，作為事件偵測的輸入。在訓練事件偵測網路時，由於特徵圖的提取在球體追蹤時便已完成，故在訓練時可節省許多時間。我們以包含核心事件的前五幀畫面進行事件的訓練，目的是為了避免球體追蹤與事件偵測之間的延遲，意即完成球體追蹤後等待事件偵測的資料不斷堆疊，事件偵

測的設計將會盡量避免過量的計算，以加快神經網路的速度。結合上述球體追蹤與事件偵測兩個核心神經網路，本系統架構圖如圖一所示。



圖一、系統架構圖

4.3 球體追蹤網路

本專題基於現有的 TrackNet[1]系統進行改善，提出一球體追蹤神經網路。利用卷積與反卷積神經網路組成，將數個連續幀畫面一起輸入球體追蹤網路，以生成桌球球體位置的熱度圖，來達到球底追蹤的目的。

4.3.1 球體追蹤架構

TrackNet 的球體追蹤架構參考文獻[12] VGG16 的前 13 層的設計，用來對物體作特徵擷取；14-24 層為反卷積神經網路，用來做像素級別的預測，利用 Upsampling 來恢復因為 Pooling 所引起的 Downsampling 效果，整體架構如表一所示。原先在 TrackNet 網路中是使用連續三幀畫面做為輸入，為了配合我們的事件偵測網路，我們將其增加至連續五幀畫面，這樣可以使神經網路更好的學習球體飛行的軌跡。又因為 TrackNet 神經網路原先是用於追蹤網球球體中心座標，其球體直徑大小在畫面上與桌球球體直徑大小不盡相同。經過我們的測量發現桌球球體在畫面中直徑大小平均為 12 pixels。我們預計將 12 pixels 作為球體追蹤神經網路的 Ground Truth 依據以及判定桌球球體座標時的重要參數。

表一、球體追蹤神經網路架構

Input size	Operator	In-channels/ Out-channels	Padding	Stride	Activation
640 x 360	Conv 3x3	15/64	2	1	ReLU+BN
640 x 360	Conv 3x3	64/64	2	1	ReLU+BN
640 x 360	MaxPool 2x2	64/64	-	2	-
320 x 180	Conv 3x3	64/128	2	1	ReLU+BN
320 x 180	Conv 3x3	128/128	2	1	ReLU+BN
320 x 180	MaxPool 2x2	128/128	-	2	-
160 x 90	Conv 3x3	128/256	2	1	ReLU+BN
160 x 90	Conv 3x3	256/256	2	1	ReLU+BN
160 x 90	Conv 3x3	256/256	2	1	ReLU+BN
160 x 90	MaxPool 2x2	256/256	-	2	-
80 x 45	Conv 3x3	256/512	2	1	ReLU+BN
80 x 45	Conv 3x3	512/512	2	1	ReLU+BN
80 x 45	Conv 3x3	512/512	2	1	ReLU+BN
80 x 45	UpSample 2x2	512/512	-	-	-
160 x 90	Conv 3x3	512/256	2	1	ReLU+BN
160 x 90	Conv 3x3	256/256	2	1	ReLU+BN
160 x 90	Conv 3x3	256/256	2	1	ReLU+BN
160 x 90	UpSample 2x2	256/256	-	-	-
320 x 180	Conv 3x3	256/128	2	1	ReLU+BN
320 x 180	Conv 3x3	128/128	2	1	ReLU+BN
320 x 180	UpSample 2x2	128/128	-	-	-
640 x 360	Conv 3x3	128/64	2	1	ReLU+BN
640 x 360	Conv 3x3	64/64	2	1	ReLU+BN
640 x 360	Conv 3x3	64/256	2	1	ReLU+BN
230400	Softmax	256/256	-	-	-

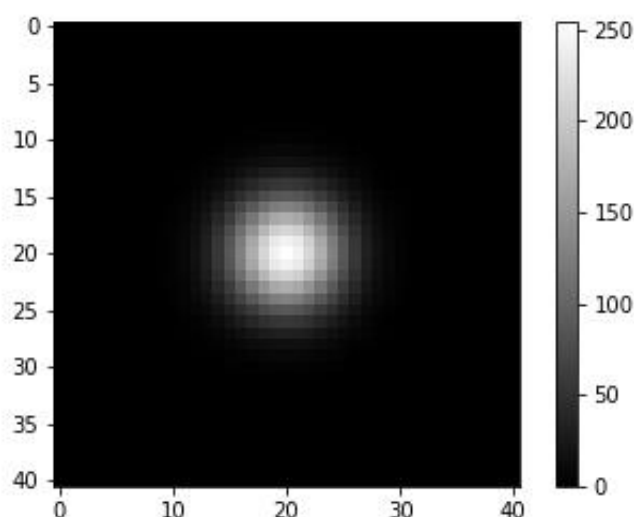
4.3.2 神經網路輸出與熱度圖

本專題中的球體追蹤神經網路在經過最後一層卷積後，會生成與輸入大小相同，但深度為 256 的圖像。其 256 層深度是為了對應灰階值 0~255，值越大代表該像素為此深度所對應灰階值的可能越大。經過 Softmax 層後，會將 256 層深度中的值正規化，進而做出球體位置的熱度圖。

令 $P(i, j, k)$ 表示灰階 k 的於位置 (i, j) 處的機率值，深度指數為 $0 \leq k \leq 255$ ，像素值為 $(0, 0) \leq (i, j) \leq (639, 359)$ ，而熱度圖在位置 (i, j) 的值即是具有最高機率灰階值，如下式一。

$$h(i, j) = \arg P(i, j, k) \quad (1)$$

透過式一得到熱度圖，如圖二所示，此圖即為桌球在畫面中的機率分布，接著，可以如以下步驟取得桌球球體位置。首先，熱度圖通過所設定的閾值轉換為黑白二位元圖。如果熱度圖上的值大於閾值，則輸出設置為 255，否則輸出設置為 0。此計畫中，閾值設定為 127。然後將會使用 OpenCV 中名為 Hough Gradient Method 的圓形查找演算法，去查找熱度圖上的圓形。若只有一個圓形被找到，則輸出圓的中心。反之，則認為畫面中沒有桌球球體的存在。



圖二、偵測桌球熱度圖

4.4 事件偵測網路

在一場桌球比賽中影響比賽勝負的核心事件分別為彈跳(Ball Bounce)以及飛越網子(Over Net)，我們透過事件偵測系統找出得分的主要事件。根據球體追蹤網路中 Global stage 與 Local stage 所獲得的特徵圖，將此特徵圖延用於事件偵測網路中。利用事件偵測神經網路判斷出連續畫面中核心事件，以此界定是否得分的標準。

4.4.1 事件偵測架構

參考文獻[2]中的原始事件偵測架構，本系統將會同時使用 Global stage 與 Local stage 所輸出的特徵圖，以提升事件偵測網路的準確度。原始神經網路架構是依照 120fps 高速攝影機의 影像為輸入進行設計，而本系統將會受限於 60fps 低速攝影機的使用。若本系統直接沿用該神經網路架構，將會因為低速攝影的限制，進而導致事件偵測的準確率下降。

我們參考文獻[12]中的架構設計，改良準確率不佳的原始事件偵測架構。藉以提升，在 60fps 攝影設備的條件下之事件偵測準確度。改良後的完整神經網路架構如表二所示。

表二、事件偵測神經網路架構

Input size	Operator	In-channels/ Out-channels	Padding	Stride	Activation
80 x 45	Conv 3x3	1024/512	2	1	ReLU+BN
80 x 45	Conv 3x3	512/512	2	1	ReLU+BN
80 x 45	Conv 3x3	512/512	2	1	ReLU+BN
80 x 45	Conv 3x3	512/64	2	1	ReLU+BN
80 x 45	Conv 3x3	64/64	2	1	ReLU+BN
80 x 45	Conv 3x3	64/64	2	1	ReLU+BN
80 x 45	Flatten	64/-	-	-	-
230400	FC	-	-	-	ReLU
2048	FC	-	-	-	ReLU
512	FC	-	-	-	-
2	Sigmoid	-	-	-	-

4.4.2 機率標記與輸出

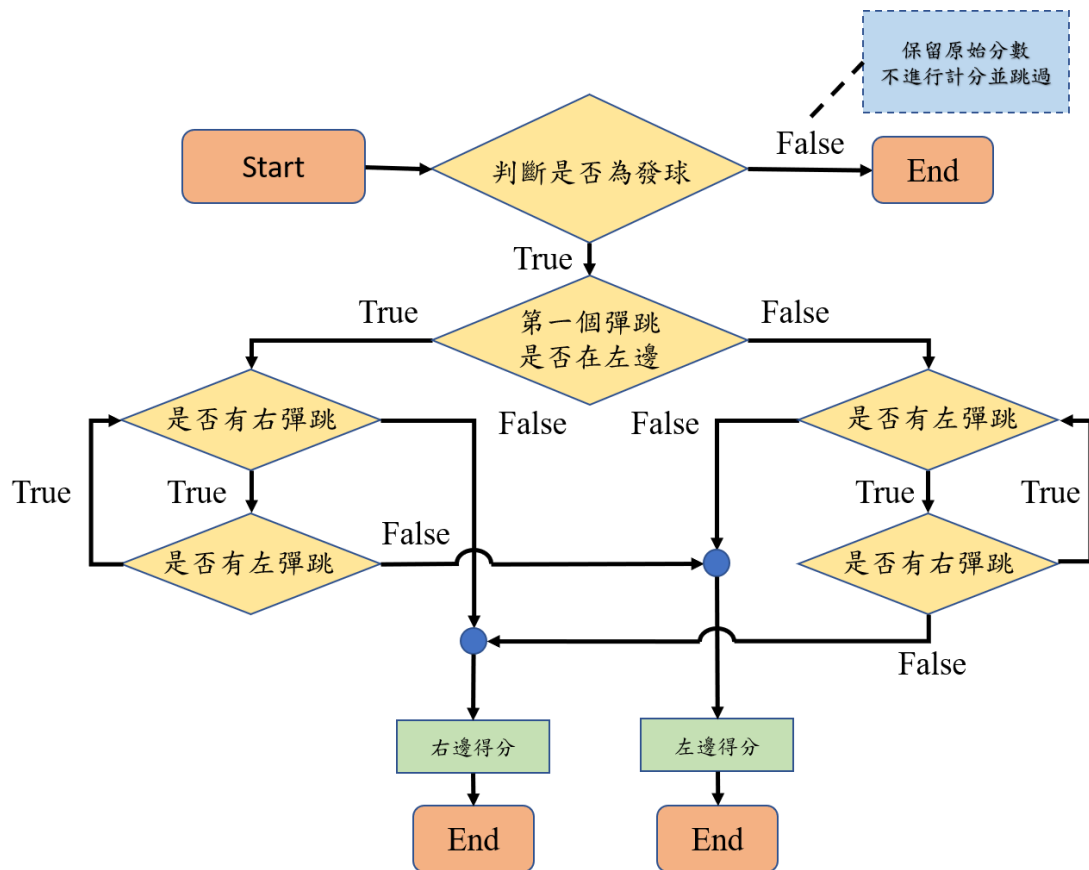
在事件偵測的輸入中，因沿用由球體追蹤網路所獲得的特徵圖，所以可以減少大量的卷積計算。在後半則使用全連接層輸出兩個主要事件(彈跳及越網)之事件機率。由於事件機率的輸出將會是一組機率數值，我們將會把最後全連接層的 Activation function 改為 Sigmoid 函數而非 ReLU，以限制最後的輸出結果為 0 至 1 的機率。

事件偵測的神經網路會以核心事件幀為最後一幀，與球體追蹤網路相同的共五幀畫面的機率進行標記與訓練。由於影格的輸入將會以五幀序列為輸入，在已完成事件標記的各組輸入中，我們將定義核心事件幀的機率為 1，並增加除核心事件幀五幀範圍外的部分五幀序列，並將其標記為空事件。其中空事件的機率為 0，代表並無發生彈跳或越網之事件。

4.5 計分演算法

由圖一可知，在系統進行完球體追蹤與事件偵測網路後，所輸出的球體座標以及事件機率將會共同輸入至計分演算法。計分演算法流程圖如圖三所示。

最一開始，會判斷當前幀是否為發球。若不符合發球條件則代表該幀不需進行計分；若符合發球條件，則判斷由哪方發球。得知發球成立以及發球方後，即可對接下來數幀的事件機率進行判斷，直到不符合保持比賽進行的指定事件發生，即可判定該次比賽由哪方玩家獲勝。



圖三、計分演算法流程圖

5. 結果與討論

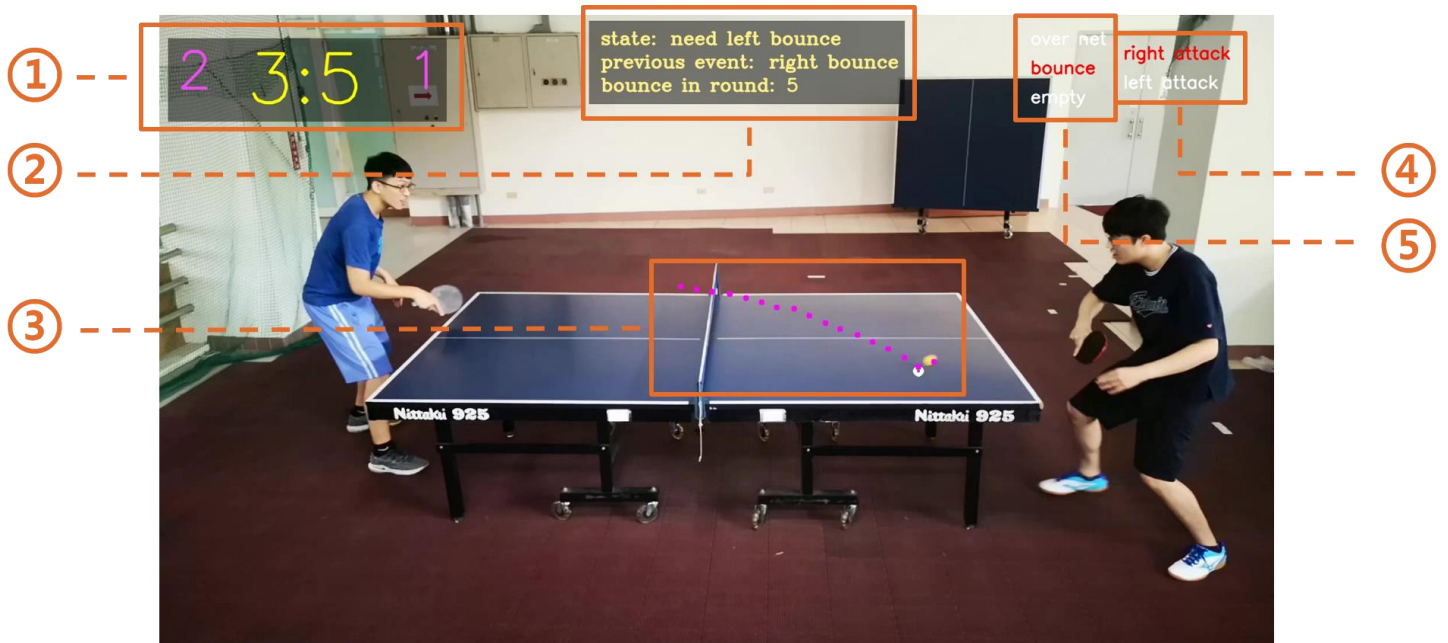
我們將結果與討論分成兩個部分，分別為 UI 介面功能介紹及模型偵測結果。其中偵測結果，將分析球體追蹤與事件偵測神經網路的準確率和執行速度。

5.1 UI 介面

本專題研究透過上述所提出的兩個核心網路，實現一自動桌球計分系統。當使用者想要對比賽復盤時，將桌球比賽影片輸入至系統後，本系統會先透過球體追蹤網路定位出球體位置。接著再透過事件偵測網路得到球體的事件機率，綜合以上資訊後，最後輸入至計分演算法，來判斷當前的得分情況，達成自動計分的功能。

當影片計算完成後，使用者可以直接觀看影片。畫面中除可以看到當前的比賽畫面外，我們也提供球體位置與事件相關資訊，如下圖四所示。將顯示的資訊分為五個部分(1) 當前比分，紫色為大局比分，黃色為當局比分，並且分數會隨著影片的進行，透過計分演算分來改變。(2) 球體事件相關資訊，第一行為下一次要發生甚麼狀態才能滿足比賽繼續的條件。第二行為前一次的狀態，而最後一行為某一方發球後球總共發生了幾次彈跳。(3) 我們將透過醒目色彩為使

用者提供球體位置，且將會滯留數幀畫面，使能夠更加明確地查看球體軌跡，並且也會標記出球體彈跳的位置。(4) 標示出此次發球為哪一方所進行的 (5) 目前球體事件為何。可分為彈跳事件、越網事件以及空事件



圖四、使用者介面

5.2 系統訓練與準確度

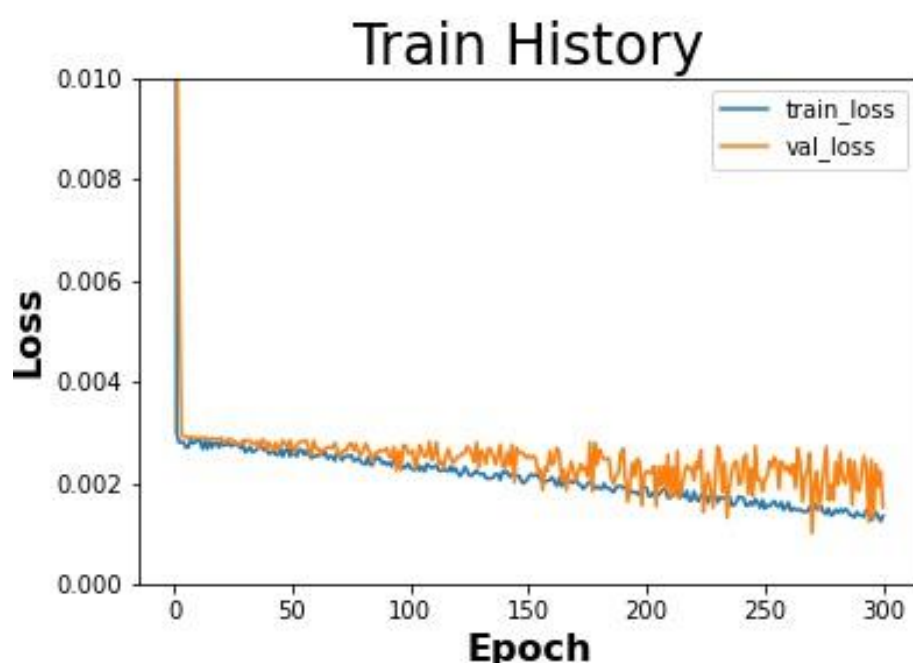
此章節將介紹球體追蹤網路與事件偵測網路的個別訓練與準確度。其中球體追蹤網路將採用定位誤差(positioning error)分布方法，而事件偵測網路則採用驗證集準確度趨勢表示。

5.2.1 球體追蹤訓練與準確率

本專題球體追蹤神經網路中，數據集中共有 25895 張照片，其中訓練集由數據集中的 20000(77%)張照片組成，驗證集由數據集中的 2000(8%)張照片組成，剩下的 3895(15%)張用於測試集中。數據集取至在多容館桌球室拍攝，總共有四種不同的背景，以及每個背景有數個些微不同的角度。訓練時的相關參數參考了文獻[1]，如表三所示，與訓練過程的 Loss 圖如圖五。

表三、訓練球體追蹤神經網路的關鍵參數

Parameters	Setting
Learning rate	1.0
Batch size	2
Steps per epochs	200
epochs	300

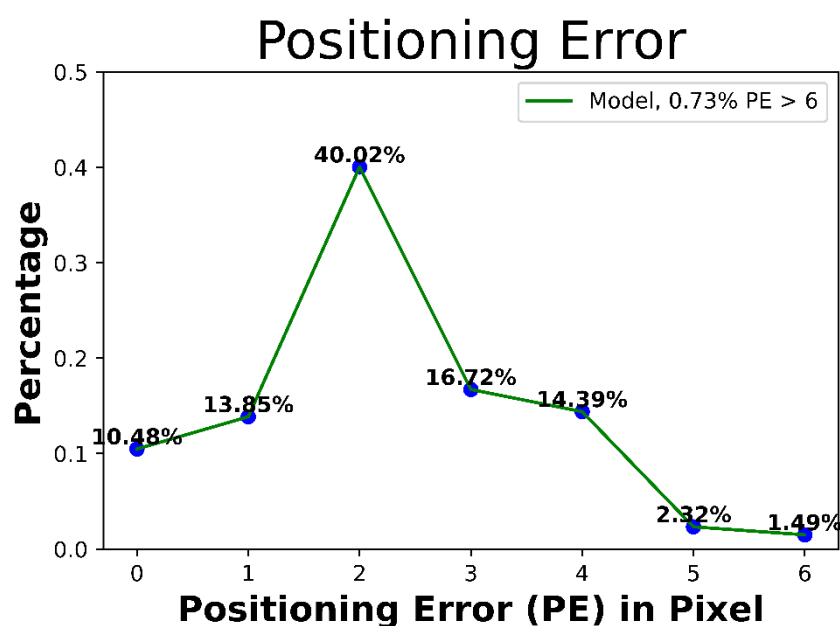


圖五、球體追蹤神經網路之損失值(loss)趨勢圖

在訓練階段，訓練球體追蹤網路總計耗時 50 小時左右，從圖四中可看到 loss 在最剛開始就很低，並且有持續緩慢下降，驗證集也有緩慢下降的趨勢，驗證球體追蹤網路的訓練是成功的。

在說明測試集的準確率前，我們先定義定位誤差(Positioning Error)，以下使用 PE 作為簡稱。PE 是由歐幾里得距離來計算預測出的座標與標記的座標之距離，測量的 PE 分布圖如圖四所示。在 $x=0$ 處是精確偵測到的百分比， $x=1$ 處表示 $0 < PE \leq 1$ 的累積機率， $x=2$ 處表示 $1 < PE \leq 2$ 的累積機率，以此類推。因為影片中桌球的平均半徑大概為 6 個像素。因此我們以 6 個像素作為判斷基準，已決定球是否被正確的偵測與定位，所以將 PE 大於 6 個像素歸類為偵測錯誤(False positive)，球體追蹤網路有 0.73% 的球是 PE 超過 6 個像素或是未被偵測到，換句

話說，有 99.27% 的球是 PE 不超過 6 個像素。經過計算得到，球體平均定位誤差為 2.2 個像素，此誤差在我們標記影片中的球時，並不會影響觀看體驗。



圖六、定位誤差(positioning error)分布圖

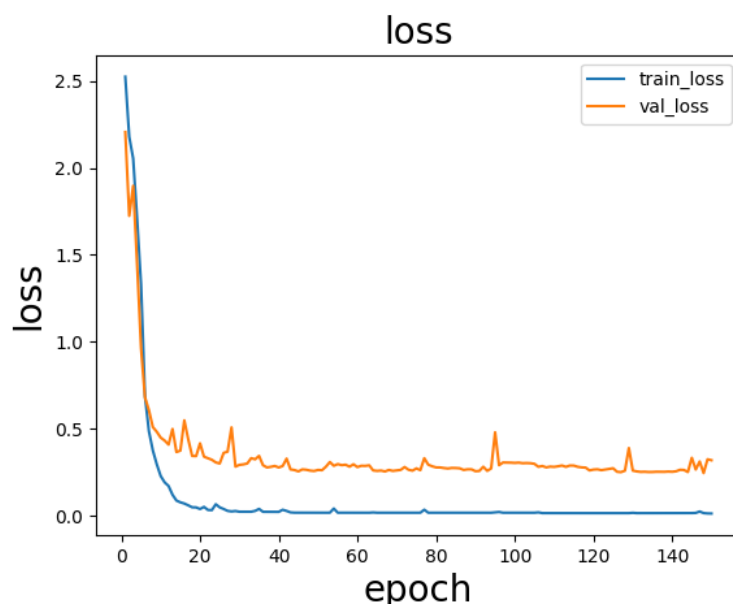
5.2.2 事件偵測訓練與準確率

本專題事件偵測神經網路中，數據集，共有 3812 筆標記。其中事件標記共 1895 筆，空事件標記共 1917 筆。其中訓練集由數據集中的 3416(90%) 筆標記組成；驗證集由數據集中的 384(10%) 筆標記組成。數據集來源與球體追蹤網路相同，總共有四種不同的背景，以及每個背景有數個些微不同的角度。訓練參數依照神經網路架構表二進行調整，完整參數設置如表四所示。

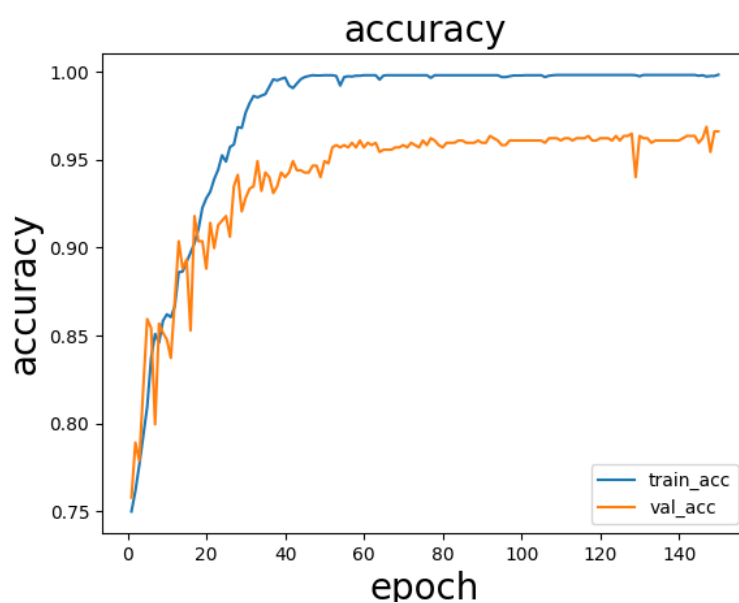
表四、訓練事件偵測神經網路的關鍵參數

Parameters	Setting
Learning rate	1e-6
Batch size	4
Steps per epochs	854
epochs	150

承 4.4.1 所述，文獻[2]中的原始事件偵測系統架構並不適合使用以 60fps 為基礎的輸入影像。我們的系統受限於 60fps 低速攝影機的攝影速度，若我們直接使用原始事件辨識網路，則原始事件偵測網路的訓練成果，將無法達到預期的效果。原始架構訓練過程的 Loss 圖如圖七所示，而原始架構的準確度趨勢圖如圖八所示。

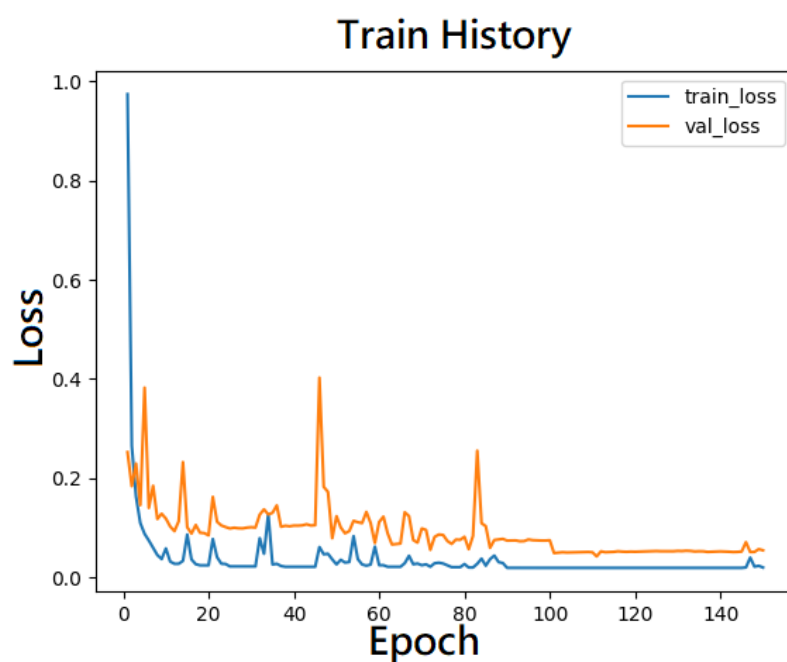


圖七、原始架構神經網路之損失值(loss)趨勢圖

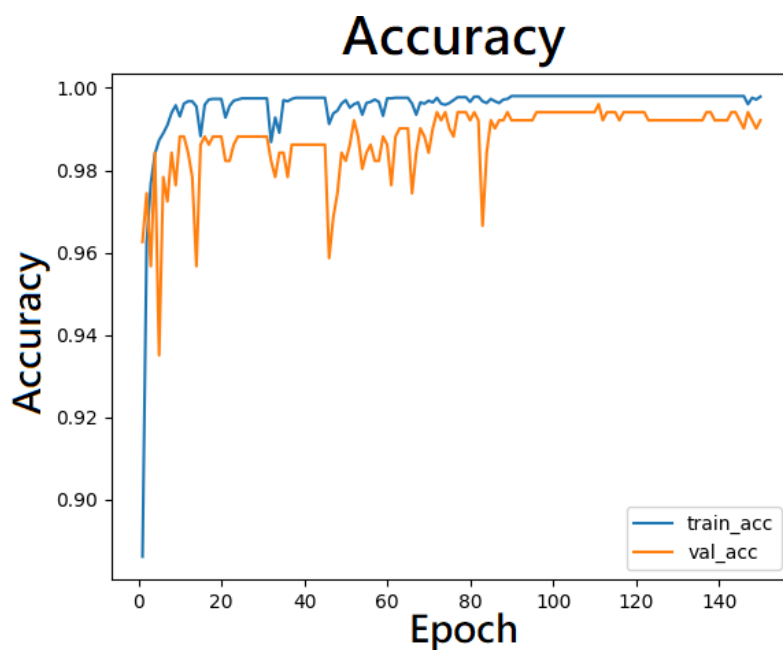


圖八、原始架構神經網路之準確率(accuracy)趨勢圖

在數據集相同以及關鍵訓練參數的設置策略相同的條件下，我們改善事件偵測網路架構後進行訓練。改良後的事件偵測神經網路如表二所示。改良後架構的損失值趨勢圖如圖九所示，改良後架構的準確值趨勢圖如圖十所示。



圖九、改良架構神經網路之損失值(loss)趨勢圖



圖十、改良架構神經網路之準確率(accuracy)趨勢圖

根據圖表所示，以 60fps 低速攝影機為目標，原始事件偵測網路的驗證集損失值最低僅能達到約 0.1，而原始架構的準確值則是約 0.95。在改良事件偵測網路架構後，改良後之神經網路的損失值最低可以低於 0.05，準確值則可到達 0.992 以上。

5.3 總結

在球體追蹤網路上，我們參考了文獻[1]的神經網路架構，將原始網路架構輸入成功的增加至連續五幀的圖片。由圖六的誤差圖可看出，球體追蹤網路擁有著極高的辨識準確率；在事件偵測的網路架構上，透過改良事件偵測的原始網路架構，我們突破原始架構在低影格數的限制下，事件偵測準確率的瓶頸。從準確率的趨勢圖對比可知，我們所改良的神經網路架構，可以成功適用於一般低速攝影機的事件偵測系統，並有效的解決低影格數下，事件偵測準確率不足的問題。

最後，我們結合高準確率的球體追蹤網路，以及改良後的事件偵測，完成本系統的核心目的：基於深度學習之全自動桌球事件偵測計分系統。與現有的系統相比，我們為使用者提供了一套，透過簡單機器即可達成，但卻又不失精準準確率的全自動桌球計分系統。作為一套使用簡單機器，卻又精緻且完整的計分系統，以此為使用者提供了另一個全新的計分選擇。

6. 參考文獻

- [1] Y. Huang, I. Liao, C. Chen, T. Ik, W. Peng, "TrackNet: A Deep Learning Network for Tracking High-speed and Tiny Objects in Sports Applications," *16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2019, pp. 1-8.
- [2] R. Voeikov, N. Falaleev and R. Baikulov, "TTNet: Real-time Temporal and Spatial Video Analysis of Table Tennis," *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2020, pp. 3866-3874.
- [3] Bal, B. S. and G. Dureja. "Hawk Eye: A Logical Innovative Technology Use in Sports for Effective Decision Making", *Sport Science Review*, vol. XXI, No. 1-2, 2012, pp. 107-119.
- [4] Y. Zhang, W. Wei, D. Yu, C. Zhong, "A Tracking and Predicting Scheme for Ping pong Robot," *Frontiers of Information Technologu and Electronic Engineering*, vol. 12, 2011, pp. 110-115.
- [5] J. Mao, D. Mould, S. Subramanian, "Background Subtraction for Real-time Tracking of a Tennis Ball," *Computer Vision Theory and Applications*, 2007, pp. 427-434.
- [6] M. Teimouri, M.H. Delavaran, M. Rezaei , "A Real-Time Ball Detection Approach Using Convolutional Neural Networks," *RoboCup 2019: Robot World Cup XXIII*, 2019, pp. 323-336.
- [7] J. Komorowski, G. Kurzejamski, G. Sarwas, "DeepBall: Deep Neural-Network Ball Detector," *14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, vol. 5, 2019, pp. 297-304.

- [8] Y. Zhao, J. Wu, Y. Zhu, H. Yu and R. Xiong, "A Learning Framework Towards Real-time Detection and Localization of a Ball for Robotic Table Tennis System," *IEEE International Conference on Real-time Computing and Robotics (RCAR)*, Okinawa, 2017, pp. 97-102.
- [9] S. Lewis, "Experimental Rig & Procedure," *New Techniques for Tennis Ball Motion Tracking*, 2016, pp. 39-41.
- [10] C.J. Soler, "Implementation," *Table Tennis Ball Tracking and Bounce Calculation Using OpenCV*, 2017, pp. 5-17.
- [11] F. Yan, W. Christmas, J. Kittler, "A Tennis Ball Tracking Algorithm for Automatic Annotation of Tennis Match", *British Machine Vision Conference*, vol. 2, 2005, pp. 619–628.
- [12] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", *International Conference on Learning Representations*, 2015, pp. 1-14.