

國立台北大學資訊工程學系專題報告

Mobile Application Development Based on Deep Learning for Real-time Sports Comparison and Guidance

專題組員:邱信瑋、溫佩旻、萬逸涵、劉嘉蕎

專題編號:PRJ-NTPUCSIE-110-006

執行期間:110 年 9 月 至 111 年 6 月

1. 摘要

根據統計，全球健身應用軟體市場在未來十年將以 21.6%的複合年增長率成長，顯示全球對於在家健身的需求十分龐大。本次專題製作一個開放性的即時動作比對及指導應用軟體，供使用者上傳感興趣的任意教學影片，和手機前置鏡頭拍攝的錄影檔，接著採用關節點預測模型 BlazePose 來預測關節點。為了做到即時的動作比對，我們必須先減少生成立體骨架模型的計算量，將輕量化的模型部屬在所撰寫的 Android 軟體中，使用此模型快速產生鏡頭前人物的立體骨架。最後，再即時計算使用者運動與練習影片的動作差距，對每個身體全身及每個不同部位的姿勢評分，同時給予動作指導。

2. 簡介

如今，互聯網上有著數以萬計的健身教學或跳舞影片，即使教練教學經驗豐富，知道初學者常犯的錯誤有哪些，仍無法針對每個人的情況即時指導。由於缺乏學習反饋，這些網路影片仍無法有效取代昂貴的健身或舞蹈課程。儘管市面上有許多健身即時比對系統，卻往往限縮使用者的選擇，只能夠訓練固定的課程影片，不能讓使用者針對任何想要的影片練習，而且常需搭配購買額外硬體設

備，要價不菲。如果無法選擇感興趣的影片練習，又需要花很多錢購買，或是需要配戴傳感器在身上，就會讓使用者缺乏健身或練舞的動力，很快便放棄自主運動或練舞。

除此之外，市售的健身比對系統仍然缺乏具體的運動指導，僅僅對使用者的動作給予模糊的評分，缺乏直觀且完整的反饋，與健身房教練、舞蹈老師能提供的建議相差甚遠。因此本次專題想製作一個開放性的即時動作比對及指導系統，只要使用行動裝置的前置鏡頭，就可以協助只用者練習任何想要的健身或舞蹈影片。

倘若能結合人工智慧 (artificial intelligence, AI)，將手機錄像即時偵測出使用者的關節點座標，產生立體的骨骼位置，並通過特定演算法比對使用者與教練的動作，再以直觀的文字及圖示即時回饋使用者姿勢的正確度或是給予指導，便能解決上述市售產品的痛點，另外又可以練習任意教學影片，讓使用體驗更好。

3. 專題進行方式

我們設計及開發一個供使用者選擇一個影片作為學習範例並跟隨它進行練習的 APP，圖 1 為本次專題的架構圖。使用者(user)上傳一個自己想要學習的教練(tutor)影片後，APP 端會將範例影片傳至 server 端透過 AI 模型生

成關節點資料並保存在 server 端，如圖 2。並在開始學習選定的影片時，將選定影片的關節點資訊傳回手機 client 端，影片開始後，client 端會透過邊緣運算呈現出使用者的各關節點，如圖 3 所示。且對各關節點和 tutor 比較並即時進行詳細的運動評估(包含指導和評分)，讓用戶可以像在教練的陪伴下一樣進行運動，並在學習影片結束後，再給出整體運動評分。

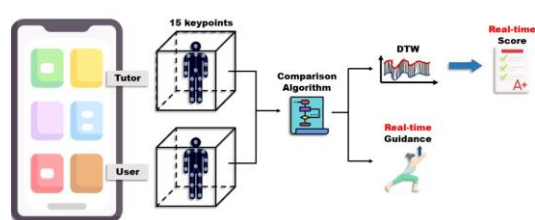


圖 1 軟體架構圖

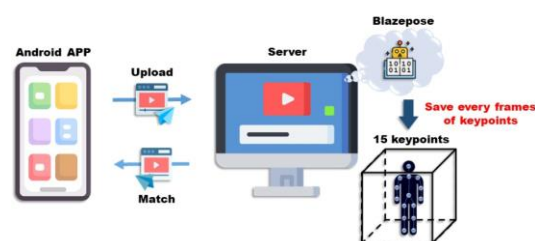


圖 2 Tutor 影片的計算

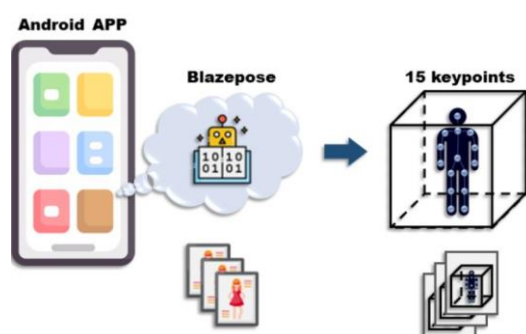


圖 3 User 影片的計算

3.1 AI 架構

我們使用 BlazePose 從 RGB 影像中抓取人體的 3D 關節點位置。模型採用編碼-解碼 (encoder-decoder) 架構預測每個關節點的熱圖(heat map)與偏移

圖(offset map)，再交給另一個編碼器(encoder)預測關節點座標。

BlazePose 的訓練數據集包括 6 萬張有單個或幾個人在場景中擺出常見姿勢的影像，和 2.5 萬張有單個人在場景中做健身運動的影像。也因此 BlazePose 在運動姿勢偵測的表現會比其他常見模型更加優秀。

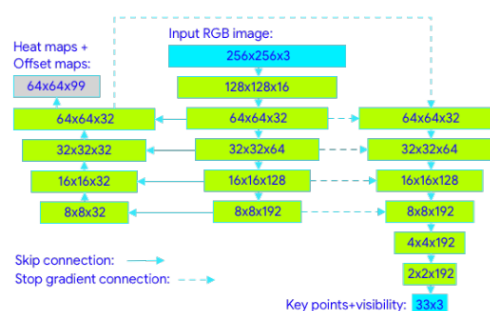


圖 4 BlazePose 架構

如圖 4 所示，左塔與中塔輸出關節點熱圖與偏移圖，交給右塔預測關節點座標。模型最終輸出為 33 個關節點的 x, y, z 座標，然後之中多數是五官和手腳位置，不適合進行運動比對，所以我們僅保留當中的 13 個四肢關節點，並額外生成 2 個軀幹關節點，以便計算關節夾角，如圖 5 所示。

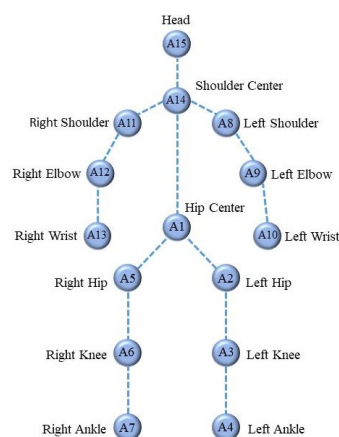


圖 5 關節點標示

為了達到即時預測的速度，BlazePose 針對連續影格只會抓取第一個影格的關節點位置，之後會以追蹤座標的方式產生後續影格的關節點，直到人物消失才會重新抓取。

本次專題在 server 端使用經過 XNNPACK 優化的 TensorFlow Lite 模型在 CPU 上預測關節點，在邊緣裝置則交由 TensorFlow Lite 套件自動以 CPU-GPU 並行的方式預測。Server 端處理器為 Intel® Xeon® W-2125 @ 4.00 GHz。

3.2 動作比對演算法

本次專題，我們將透過 AI 模型生成 15 個關節點，其位置如圖 5 所示。並且利用此 15 個關節點，計算關節點間的角度，如圖 6。

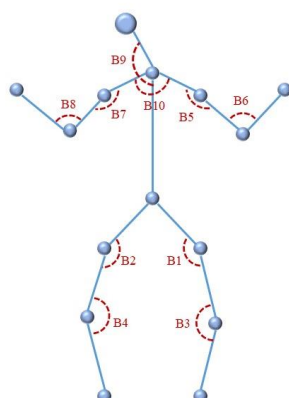


圖 6 關節點間角度標示

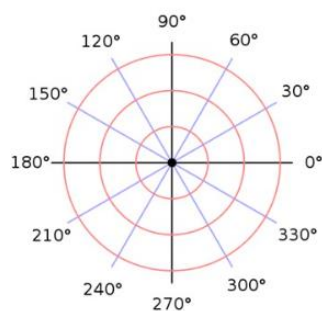


圖 7 極座標說明圖

並以 hip center 和 shoulder center 的中心為原點，分別計算各自的極座標，如圖 7。

另外，user 和 tutor 當前的 frame 分別和各自的前一個 frame 進行比較，若 x 軸方向的座標值增加，且 y 和 z 軸方向座標值未增加，則當前的 frame 會記錄(+x, 0, 0)，圖 8 為運動傾向說明圖。

Movement Tendencies		
-x	0	x
-y	0	y
-z	0	z

圖 8 運動傾向說明圖

評分部分分成兩個部分，即時評分以及最後評分。即時評分是透過 tutor 和 user 儲存每個 frame 中，15 個關節點各自的極座標、運動傾向和 10 個關節點間夾角計算數值，並且每 10 個 frame 進行一次動態時間規整 (dynamic time warping, DTW) 計算，兩速度不同的序列使用 DTW 進行比對的概念如圖 9 所示。即時評分的公式如下：15 個關節點各自計算極座標和運動傾向的 DTW 分數(共 30 個分數)*2% + 10 關節點間夾角的 DTW 分數(共 10 個分數)*4% = 100%。

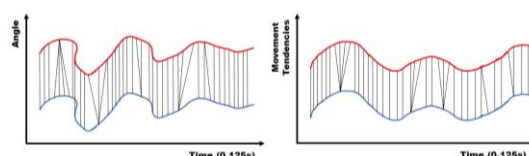


圖 9 使用 DTW 進行序列比對

最後評分部分，會將每次即時評分平均後*50% + 所有 frame 中 15 個關節點各自計算極座標和運動傾向的 DTW 分數(共 30 個分數)*2% + 10 關節點間夾角的 DTW 分數(共 10 個分數)*2% = 100%。

指導部分，本系統會透過 15 個關節點的級座標和 10 個關節點間的角度進行指導判斷，若 user 角度和 tutor 角度的差距大於一定的閾值(threshold)，我們就會及時給予 user 動作的指導。指導的方式有分為兩種。

第一種是 user 某個關節點間的角度和 tutor 不一樣時，會根據輕重程度，將其周圍兩個骨幹畫成綠色、黃色或紅色，綠色表示 user 和 tutor 幾乎一樣，黃色表示存在些許差異，紅色表示嚴重不一樣。第二種是 user 某個關節點的極座標和 tutor 不一樣時，系統會在該關節點上生成一個指示箭頭，讓 user 可以根據箭頭方向將動作調整成和 tutor 相近的動作。

3.3 APP 介面

為了避免使用者因為對 APP 的操作流程不熟悉而感到困擾，我們製作了一個簡易的 APP 流程圖(圖 10)供參考，希望能讓使用者有更好的使用體驗。



圖 10 APP 流程圖

初次使用之使用者，至註冊畫面進行註冊，如圖 11。往後使用僅需至登入畫面輸入帳號密碼即可使用，如圖 12。



圖 11 註冊畫面



圖 12 登入註冊畫面

將欲比對之教練影片上傳及命名，等待數分鐘後，將會在影片清單內找到上傳的影片，此時代表影片已經傳送至 server 端，並透過 BlazePose 提取關節點。

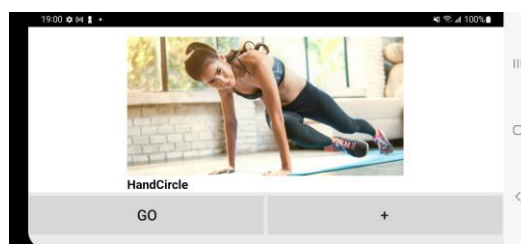


圖 13 上傳影片畫面

點選清單內欲比對的影片，並點選 GO，將會進入動作比對畫面，此時將已產生關節點之教練影片傳回使用者手機。動作比對畫面上方名稱為欲選擇的影片名稱，左側 tutor 畫面為欲選擇的教練影片，中間柱狀進度條為透過演算法計算的即時分數，右方 user 畫面中顯示使用者即時的關節點，透過輕量化 BlazePose 模型產生。骨幹的

顏色分為紅黃綠，分別表示動作準確的程度，紅色為不準確，而綠色為準確。此外，藍色箭頭提供動作指導，使用者可依照箭頭方向調整動作。動作比對及指導的畫面如圖 14 所示。

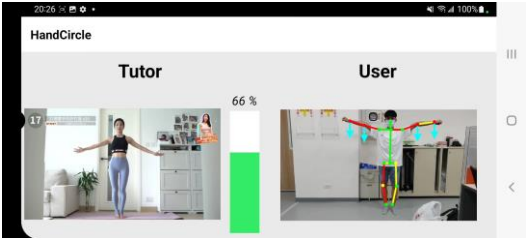


圖 14 動作比對 match 畫面

動作比對結束後，透過演算法計算出總分，顯示成績畫面如圖 15 所示。

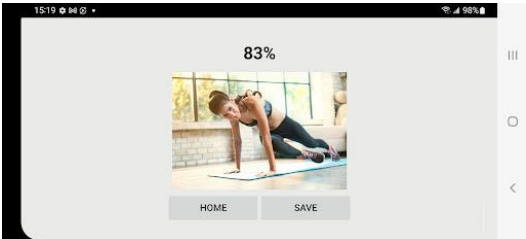


圖 15 動作比對結束顯示成績畫面

4. 主要成果與評估

此次研究成果包含 AI 模型、比對演算法及 APP，當使用者欲學習新的運動影片時須先上傳影片至 server 端為 tutor 影片預測關節點，之後可以選擇上傳過的任意影片作學習。在運動過程中 user 會透過我們輕量化過後的 model 預測關節點，並同時透過比對演算法給予及時分數和當下的指導。

AI 模型的部份我們不使用目前較常被用來預測關節點的模型 OpenPose，而是使用的 BlazePose 這個模型的原因如表 1 所示。BlazePose 在速度上快於 OpenPose 許多。再加上我們使用該模型的輕量化，即使在準確度上略低一點，但對於我們即時比對的特性來說，允許忽略些微差異換取

近 10 倍的速度。

我們提出了可以即時給予運動者回饋的系統，但在關節點預測上還存在些許誤差，在運動指導上也存在一些個體化差異的問題，因此希望在未來能夠透過改善模型架構來提升系統整體的準確性，讓 APP 能更精準地給予使用者運動反饋。

表 1. BlazePose vs. OpenPose

Model	FPS	AR Dataset, PCK@0.2	Yoga Dataset, PCK@0.2
OpenPose (body only)	0.4 ¹	87.8	83.4
BlazePose Full	10 ²	84.1	84.5
BlazePose Lite	31 ²	79.6	77.6

5. 結語與展望

此次專題我們成功開發出一個簡易且方便的 APP，讓使用者可以在運動時得到及時姿勢的正確度及指導的回饋。這樣能即時給予運動者反饋及指導的 APP 目前尚未有人提出，我們是第一個提出此方法的團隊。然而此應用程式還無法自動辨別非運動時的動作，例如:休息、說話，這導致使用者在上傳影片前需先將影片進行切割，否則在比對上會有一些誤差，因此我們期望在未來能做到讓 APP 做到自動區分運動及非運動狀態，增加使用的方便性，同時也希望能做到在上傳影片時與 YouTube 做連接，讓使用者不需再手動上傳影片，增加便利性。最後，本研究還無法自動忽略人體差異，需自行調整手機擺放位置，若與影片中 tutor 的拍攝方向差異過大會造成誤差，若體型和 tutor 差異太大也需要自行調整與鏡頭間的位置來減少差異性，因此期望在未來能改善此功能，讓 App 自動忽略人體及拍攝方向的差異性。

6. 銘謝

非常感謝指導教授提供大量資源與豐富的研究經驗分享，以及實驗室學長姐的協助，讓我們在進行專題的路上更加順利，也謝謝系上學姐來當我們 demo 影片的主角。最後也感謝組員們的合作才能使我們的專題能如期完成。

7. 參考文獻

- [1] 錢玉紘 (2021)。「手機鏡頭就能「骨骼標定」，靠 AI 揪錯誤姿勢！健身 App Uniigym 還有哪些亮點？」2021 年 11 月 10 日。取自 <https://www.bnext.com.tw/article/66084/workout-ai-app-deeplearning>
- [2] M.-C. Hu, C.-W. Chen, W.-H. Cheng, C.-H. Chang, J.-H. Lai, and J.-L. Wu, "Real-Time Human Movement Retrieval and Assessment With Kinect Sensor," *IEEE Trans. Cybern.*, vol. 45, no. 4, pp. 742–753, Apr. 2015.
- [3] Y. Liao, A. Vakanski, and M. Xian, "A Deep Learning Framework for Assessing Physical Rehabilitation Exercises," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 2, pp. 468–477, Feb. 2020.
- [4] C.-H. Huang, C.-F. Lin, C.-A. Chen, C.-H. Hwang, and D.-C. Huang, "Real-time rehabilitation exercise performance evaluation system using deep learning and thermal image," in *Proc. I2MTC*, Dubrovnik, Croatia, 2020, pp. 1–6.
- [5] A. Kamel, B. Sheng, P. Li, J. Kim, and D. D. Feng, "Efficient Body

Motion Quantification and Similarity Evaluation Using 3-D Joints Skeleton Coordinates," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 51, no. 5, pp. 2774–2788, May 2021.

- [6] Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F. and Grundmann, M., "BlazePose: On-device Real-time Body Pose tracking," in *Proc. CVPR*, Seattle, WA, USA, 2020.