



運動影像技術於籃球罰球動作 特徵擷取與評估之應用

蘇明祥^{1,*} 許叡哲² 李佳芸³

¹ 東吳大學資料科學系副教授

² 東吳大學資料科學系碩士

³ 東吳大學資料科學系學生

摘要

本研究探討人體姿態估計技術於運動科學之應用，聚焦於籃球投籃技術的動作分析。為提升運動技術分析之精準度與科學化，本研究結合 OpenPose 與深度學習模型，即時擷取投籃動作中的人體關節座標，精確計算腕關節距離與各部位關節角度，以呈現投籃過程中的姿態變化與運動細節。所得動作資料可作為評估投籃穩定性與準確性的客觀依據。研究亦根據學術文獻與教練專業經驗，建立標準化的投籃姿勢評估準則，用以辨識動作特徵並評估技術優劣。進一步運用機器學習分類模型對姿態特徵進行訓練與測試，成功區分高品質與低品質的投籃動作。實驗結果顯示，本研究方法之分類準確率達 100%，驗證其於運動姿態分析中之高效性與可行性。

關鍵詞：人體姿態估計、深度學習、機器學習、籃球投籃、OpenPose

* 通訊作者：蘇明祥

電子郵件：huntfox.su@gmail.com

(收件日期：2025 年 9 月 17 日；修正日期：2025 年 10 月 28 日；接受日期：2025 年 10 月 28 日)



The Use of Motion Image Technology in Feature Extraction and Assessment of Basketball Free-Throw Techniques

Ming-Hsiang Su^{1,*}, Jui-Che Hsu², Chia-Yun Lee³

¹Associate Professor, Department of Data Science, Soochow University

²Master, Department of Data Science, Soochow University

³Student, Department of Data Science, Soochow University

Abstract

This study examines the application of human pose estimation technology in sports science, with a focus on analyzing the motion of basketball shooting techniques. To enhance the precision and scientific rigor of movement analysis, the research integrates OpenPose with deep learning models to capture real-time human joint coordinates during shooting motions. The system accurately calculates hip distances and joint angles of various body parts to illustrate postural variations and kinematic details throughout the shooting process. The collected motion data serve as an objective foundation for evaluating the stability and accuracy of shooting performance. Based on academic literature and expert coaching experience, a standardized evaluation framework for shooting posture was established to identify motion characteristics and assess technical proficiency. Furthermore, machine learning classification models were trained and tested on the extracted pose features, successfully distinguishing between high- and low-quality shooting performances. Experimental results indicate that the proposed method achieved a classification accuracy of 100%, demonstrating its effectiveness and feasibility in sports motion analysis.

Keywords: human pose estimation, deep learning, machine learning, basketball shooting, OpenPose

* Corresponding Author: Ming-Hsiang Su
E-mail: huntfox.su@gmail.com

壹、研究動機與研究問題

近年來，隨著國內職業籃球聯盟的建立與發展，大眾對籃球運動的熱情逐漸高漲。然而，我國球隊在國際賽事中的表現仍不理想。部分原因在於，相較於鄰近國家如日本與韓國，這些國家早已透過科學化訓練方法顯著提升運動員的競技表現，顯示我國在基層訓練與運動科學應用方面仍存在明顯差距。在籃球運動中，「罰球」雖屬基礎動作，卻是決定比賽勝負的關鍵得分手段之一。其命中率的高低往往直接影響球隊整體表現。因此，如何提升投籃動作的穩定性與罰球命中率，成為球員訓練的重要課題。根據過往學術研究（Okazaki, Rodacki, and Satern, 2015; Okubo and Hubbard, 2016; Tran and Silverberg, 2008）指出，投籃命中率的關鍵因素在於動作姿勢的穩定性，以及出手時對球體投射角度（Ball Angles）與力量的精確控制。然而，傳統訓練多依賴教練的經驗判斷與運動員的主觀感受，此類方法主觀性高且缺乏客觀數據支撐，限制了訓練的效率與精準度。隨著深度學習技術的快速發展，人體姿態估計（Human Pose Estimation, HPE）已逐漸成為運動科學中的關鍵工具，可應用於投籃動作的分析、矯正與個別化訓練（Ji, 2020; Oproescu et al., 2023; Yan, Jiang, and Liu, 2023）。其中，OpenPose（Cao, Simon, Wei, and Sheikh, 2017）作為一種 Bottom-Up 的骨架偵測技術，兼具即時性與高擴展性，具備應用於籃球投籃姿勢分析的潛力。因此，本研究旨在結合 OpenPose 與深度學習技術，建立一套能同時分析投籃姿勢與球體軌跡的即時評估系統。該系統可透過姿勢分類與動態數據辨識技術，提供投籃表現的量化分析結果，協助教練與球員依據客觀數據進行技術改善，進而促進我國籃球訓練的科學化與系統化發展。

貳、文獻回顧與探討

一、影像辨識（Image Recognition）

影像辨識是人工智慧與機器學習領域中的關鍵技術之一。其核心目標在於模擬人類視覺系統的運作，使電腦能理解、分析並自動辨識視覺資料，例如圖片與影片中的物體、人物或文字等資訊。此技術廣泛應用於各類場域，包括物體檢測（Seo and Park, 2019）、安全監控（Awais et al., 2019）、自動駕駛車輛（Muhammad, Ullah, Lloret, Del Ser, and de Albuquerque, 2021）、醫學影像分析（Ker, Wang, Rao, and Lim, 2018）以及人臉辨識（L. Li, Mu, Li, and Peng, 2020）等領域，顯示其在現代科技中的重要性與普及性。影像辨識技術的核心基礎在於深度學習模型。這些模型透過在大量標註數據集上的訓練，能自動學習影像的多層次特徵，並有效提升對不同物體與模式的辨識能力。卷積神經網絡（Convolutional Neural Network, CNN）（O'Shea and Nash, 2015）是影像辨識中最具代表性且應用最廣泛的深度學習模型。其架構靈感源自人類視覺皮層的層級結構，透過多層卷積（Convolution）與池化（Pooling）操作，逐步提取影像中的低階特徵（如邊緣與紋理），再進一步整合為高階語義特徵，進而達成對影像內容的理解與分類。Girshick, Donahue, Darrell, and Malik（2014）提出的區域卷積神經網絡（Region-Based Convolutional Neural Network, R-CNN）為影像辨識

技術帶來重大突破。該方法透過 Region Proposals 產生候選區域，並將每個區域輸入預訓練 CNN 進行特徵提取，再利用分類器進行物體識別。然而，R-CNN 在每個候選區域皆需獨立運行 CNN，導致計算成本與時間消耗過高。為改善此問題，後續的 Fast R-CNN (Girshick, 2015) 與 Faster R-CNN (Ren, He, Girshick, and Sun, 2017) 分別引入共享卷積特徵與區域建議網路 (Region Proposal Network, RPN)，顯著提升了處理速度與檢測效率，同時維持甚至提升了識別精度 (Bardou, Zhang, and Ahmad, 2018; Nakazawa and Kulkarni, 2018)。

二、YOLO (You Only Look Once)

YOLO 是一種將物件偵測視為單一迴歸問題 (Single Regression Problem) 的深度學習架構。與傳統的 R-CNN 系列方法不同，YOLO 模型僅需對整張影像執行一次 CNN 運算，即可同時預測影像中物體的類別與位置 (Bounding Box Coordinates)。這一創新設計大幅提升了檢測速度，並有效降低背景誤檢 (False Positives) 的發生率。YOLO 模型最早由 Redmon, Divvala, Girshick, and Farhadi (2016) 提出，隨後推出 YOLOv2 (Redmon and Farhadi, 2017) 與 YOLOv3 (Redmon and Farhadi, 2018) 版本，分別在檢測精度、速度以及多尺度檢測能力上獲得顯著提升。其後，Ultralytics 公司於 2020 年發布 YOLOv5 模型，雖非 Joseph Redmon 官方版本，卻以其高效的運算架構與優異的影像處理效能而廣泛應用於產業與研究領域 (Bachir and Memon, 2024; Jubayer et al., 2021; Lamane, Tabaa, and Klilou, 2022)。YOLOv5 能在單次影像處理中同時預測目標存在的機率與邊框座標，並整合數據增強 (Data Augmentation) 及訓練流程優化技術，使模型操作更為簡便且實用性更高。如圖 1 所示，YOLOv5 的架構可分為四個主要部分：Input、Backbone、Neck 與 Head。

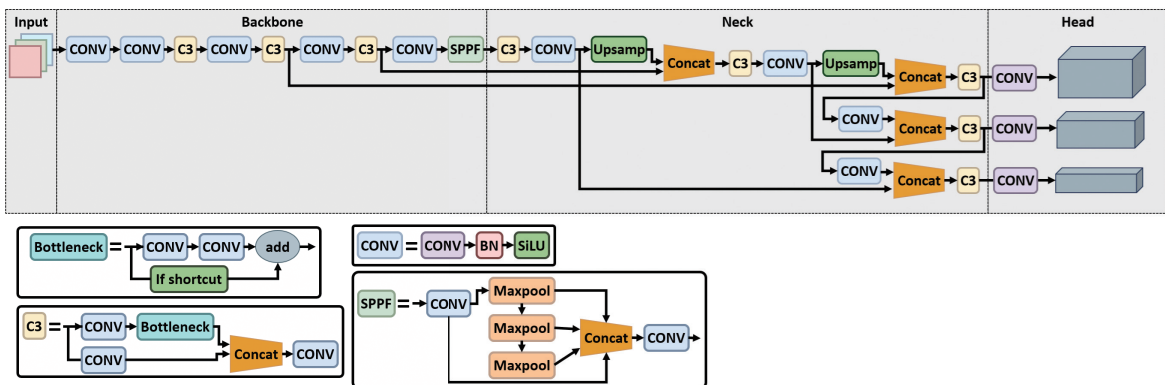


圖 1 YOLOv5 模型架構

資料來源：本研究自行整理。

(一) Input

輸入階段主要進行影像前處理，包括影像尺寸縮放與歸一化，使資料符合網路輸入需求。與先前版本不同的是，YOLOv5 內建錨框 (Anchor Box) 自動計算機制，能於每次訓練時自動根據資料集特性生成最適錨框，提升模型對不同場域資料的泛化能力。

(二) Backbone

Backbone 主要負責多層次特徵提取，透過深度卷積運算獲取高、中、低階特徵資訊。其結構包含 C3 模組與 SPPF (Spatial Pyramid Pooling-Fast) 模組。其中，CONV 模組 (Convolution Module) 輔助 C3 模組進行特徵學習，而 C3 模組結合殘差結構 (Residual Structure) 以加強梯度傳遞並減少訊息損失。SPPF 模組則為 SPP 之改良版本 (Karim, Majumdar, Darabi, and Chen, 2018; Zhang, Zhong, and Li, 2019)，透過多尺度池化操作 (Multi-Scale Pooling) 提升特徵表徵能力，並在保持高效的同時縮短計算時間。

(三) Neck

Neck 結合特徵金字塔網路 (Feature Pyramid Network, FPN) (Lin, Dollár, Girshick, He, Hariharan, and Belongie, 2017) 與金字塔注意力網路 (Pyramid Attention Network, PAN) (H. Li, Xiong, An, and Wang, 2018)，主要用於多層特徵融合。如圖 2 所示，FPN 的核心概念是透過上採樣 (Upsampling) 將 Backbone 產生的多層特徵圖 (F1、F2、F3) 轉換為新的金字塔特徵層 (P1、P2、P3)，以實現多尺度物件的辨識；而 PAN 則反向進行下採樣 (Downsampling)，將特徵圖還原至原始解析度，進一步強化特徵資訊流通與檢測精度，同時降低運算成本。

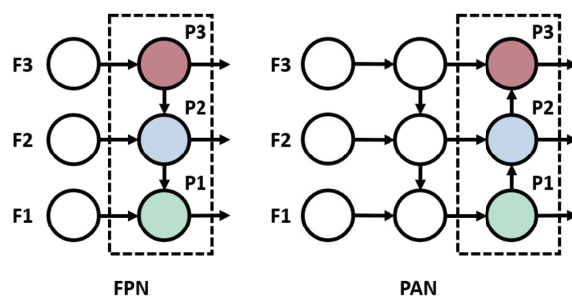


圖 2 FPN 和 PAN 網路結構

註：FPN：特徵金字塔網路 (Feature Pyramid Network)；PAN：金字塔注意力網路 (Pyramid Attention Network)。資料來源：本研究自行整理。

(四) Head

Head 模組負責執行邊界框迴歸運算，預測物體的中心點、寬度與高度，並同時輸出分類結果與信心分數 (Confidence Score)。YOLOv5 採用三個偵測分支，分別負責不同尺寸的目標偵測，從而提升模型對多尺度物體的辨識能力。在訓練階段，採用廣義交並比 (Generalized Intersection over Union, GIoU) (Rezatofighi, Tsoi, Gwak, Sadeghian, Reid, and Savarese, 2019) 作為邊界框損失函數，以更準確地衡量預測框與真實框之間的重疊度。

在後處理階段，YOLOv5 使用非極大值抑制 (Non-Maximum Suppression, NMS) (Neubeck and Van Gool, 2006) 濾除重疊候選框。具體作法為：先將低於信心閾值的候選框移除，再依據信心分數排序，保留最高者，並逐步刪除與其交並比 (Intersection over Union, IoU) (Rezatofighi et al., 2019) 超過閾值的候選框。最終，保留下的框將被標註為最終偵測結果。

三、HPE 模型

HPE 旨在從影像或影片中自動辨識並定位人體的關鍵關節 (Keypoints)，以重建人體姿勢結構。早期的姿態估計方法主要可分為兩類：基於模型的方法 (Model-Based Methods) (Andriluka, Roth, and Schiele, 2009) 與基於外觀的方法 (Appearance-Based Methods) (Johnson and Everingham, 2011)。前者透過建立人體幾何模型或骨架結構，結合人體運動學約束進行姿勢預測；後者則依賴局部影像特徵進行關節辨識與匹配。然而，這些傳統方法在低解析度影像、複雜背景或人體遮擋情況下表現不佳，且普遍缺乏對不同場景的適應性與泛化能力 (Andriluka et al., 2009)。

隨著深度學習技術的興起，2014 年出現的 DeepPose (Toshev and Szegedy, 2014) 成為第一個將 CNN 應用於 HPE 的模型。DeepPose 利用 CNN 強大的特徵抽取能力，有效提升 HPE 的準確性，為後續研究奠定了深度學習化的基礎 (Go and Aoki, 2016)。此後，眾多基於深度神經網絡的 HPE 方法陸續被提出 (Zhou, Mok, and Zhou, 2019)，顯示該領域已進入以端對端深度架構為核心的發展階段。目前，HPE 方法可依處理策略大致分為兩大類，分別是 Top-Down 方法與 Bottom-Up 方法 (Munea, Jembre, Weldegebriel, Chen, Huang, and Yang, 2020)。

(一) Top-Down 方法

Top-Down 方法的典型流程為：首先偵測影像中所有人體區域，接著對每一個檢測到的人體分別進行 HPE。代表性模型包括 AlphaPose (Zhou et al., 2019) 與卷積姿態機 (Convolutional Pose Machines, CPM) (Fang et al., 2022)。此類方法因能針對單一個體進行精細估計，通常在準確率上表現較佳。然而，其主要缺點為運算效率受影像中人體數量影響較大，當多人同時出現時，計算成本顯著增加，導致即時性下降。

(二) Bottom-Up 方法

Bottom-Up 方法則採取相反策略，直接從整張影像中偵測出所有人體關鍵點，再透過演算法將這些關鍵點分配至不同個體。此類方法的代表包括 DeepCut (Pishchulin et al., 2016; Wei, Ramakrishna, Kanade, and Sheikh, 2016) 與 OpenPose (Rajchl et al., 2017)。相較於 Top-Down 方法，Bottom-Up 對影像中人物數量的變化較不敏感，因此具有更佳的可擴展性與即時性能。然而，該方法在處理人體遮擋、交互或複雜姿勢時，容易因關節匹配錯誤而降低準確性。

整體而言，Top-Down 與 Bottom-Up 各具優缺點：前者具備高精度但低即時性，後者具備高效率但易受姿勢干擾。隨著硬體效能提升與深度模型優化，近年研究趨勢逐漸朝向結合兩者優勢的混合式架構發展，期望在保持高準確度的同時，兼顧即時運算需求。

四、OpenPose 原理與應用

OpenPose (Cao et al., 2017) 是一種基於 CNN 與監督式學習 (Supervised Learning) 的深度學習架構，為目前最具代表性的 Bottom-Up 即時多人姿態估計 (Multi-Person Pose Estimation) 方法之一。作為開源工具，OpenPose 能從靜態影像或動態影片中自動偵測人體多個關節節點與部位位置。其主要優勢在於靈活的應用程式介面 (Application Programming Interface, API) 設計，可整合多種輸入來源，如網路攝影機、行動裝置或嵌入式系統。此特性使得 OpenPose

不僅具備高擴展性，也能與其他深度學習框架（如 TensorFlow、PyTorch）結合，廣泛應用於運動分析、人機互動及姿態識別等領域。

如圖 3 所示，OpenPose 的整體架構可分為特徵提取（Feature Extraction）階段、關鍵點預測（Keypoint Prediction）階段及關聯場估計（Part Affinity Estimation）階段。其中，符號 F 代表輸入影像經由 VGG-19 Backbone Network 前 10 層所提取之特徵圖。這些特徵圖會傳入後續的多階段卷積網路，以進行 HPE 與優化。

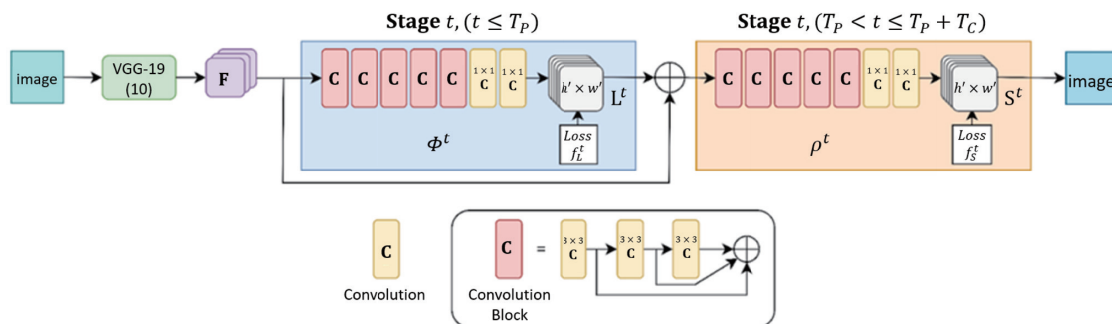


圖 3 OpenPose 架構圖

資料來源：本研究自行整理。

模型主要包含兩組 CNN：

（一）部件關聯場（Part Affinity Fields, PAF）

用於預測人體各部位之間的關聯向量場（Association Vector Fields），表示不同關節間的空間關係。

（二）關節點熱圖（Confidence Maps）

用於估計各個關節點在影像中出現的機率位置。在每一個訓練階段中，OpenPose 會同時產生關節點熱圖與 PAFs，並將前一階段的輸出結果遞迴（Iteratively）輸入下一階段，以進一步優化預測精度。這種多階段卷積結構（Multi-Stage CNN Architecture）使模型能逐步修正錯誤並提升定位準確度。最終，透過結合多個階段的結果，模型能在保持高精度的同時維持即時推論能力。此外，OpenPose 採用多分支（Multi-Branch）網路架構，可同時偵測不同身體部位（如身體骨架、手部與臉部）的關鍵點集合。這種設計不僅提升整體檢測覆蓋率，也使其能在多人場景下保持穩定運算效能。憑藉此結構，OpenPose 成為運動姿態分析、動作辨識與生物力學研究中最廣泛應用的 HPE 模型之一。

五、投籃動作分類模型

為有效辨識並分類籃球選手在投籃過程中的動作模式，本研究比較多種具代表性的分類模型，涵蓋傳統機器學習方法與深度學習架構兩大類。

（一）傳統機器學習模型

傳統分類方法包括支援向量分類器（Support Vector Classifier, SVC）（Hearst, Dumais, Osuna, Platt, and Scholkopf, 1998）與邏輯迴歸（Logistic Regression, LR）（Zou, Hu, Tian, and Shen, 2019）。這兩種方法具有理論穩定性與實務可行性，已被廣泛應用於各領域的分類任務中。SVC 透過在高維空間中尋找最佳分離超平面（Optimal Separating Hyperplane）以區分不同類別，特別適合處理非線性邊界的問題；而 LR 則以機率形式建模輸出類別，具備良好的可解釋性與效率。隨著資料量的成長與運算能力的提升，集成學習（Ensemble Learning）方法逐漸受到重視。其中，隨機森林（Random Forest, RF）（Breiman, 2001）結合多棵決策樹（Decision Trees）進行投票或平均預測，以提升模型的穩定性與準確度，同時有效緩解過擬合問題。RF 因其優異的泛化能力與對異質資料的適應性，已被廣泛應用於分類與迴歸任務中（Gupta, Gupta, Kumar, and Sardana, 2021; Sheykhmousa, Mahdianpari, Ghanbari, Mohammadimaneh, Ghamisi, and Homayouni, 2020）。

（二）深度學習模型

近年來，深度學習技術的發展使模型能夠自動從大量非結構化資料中學習複雜特徵表示，成為解決分類問題的強大工具。其中，多層感知器（Multilayer Perceptron, MLP）（Murtagh, 1991）作為最早的前饋神經網路之一，透過多層非線性轉換實現特徵映射，在影像辨識（Song, Lam, Han, and Li, 2020）、語音處理（L. Yang, Xie, Wen, and He, 2021）、自然語言處理（Wang, Yu, Lai, and Zhang, 2020）與預測任務中均展現出良好表現。另一重要模型為長短期記憶網路（Long Short-Term Memory, LSTM）（Hochreiter and Schmidhuber, 1997），其結構中引入遺忘門（Forget Gate）、輸入門（Input Gate）與輸出門（Output Gate），能有效捕捉長期依賴關係，避免梯度消失問題。LSTM 特別適用於處理序列資料，如語言模型（Jorge, Giménez, Silvestre-Cerdà, Civera, Sanchis, and Juan, 2022）、時間序列預測（Karim et al., 2018）與動作辨識等任務，能從連續的時間步中學習動作變化特徵。

（三）模型比較與研究定位

綜合上述方法，各模型在分類任務中的效能表現受資料特性與任務需求影響而異。傳統機器學習模型在特徵維度較低或資料量有限時具備效率與可解釋性優勢；而深度學習模型則能在大量時序資料中自動提取潛在特徵，展現更高的分類精度與泛化能力。因此，本研究將比較多種傳統與深度學習模型在投籃姿態分類任務中的表現，評估其準確率、穩定性與實際應用潛力，以尋找最適合應用於籃球訓練輔助系統之分類演算法。

六、HPE 與分類任務相關研究

過去有關籃球動作分析的研究多依賴穿戴式感測裝置（Wearable Sensors）蒐集人體運動資訊，並結合機器學習演算法進行資料分析。此類方法雖能提供高精度的動作參數，但通常需仰賴昂貴設備與複雜的資料蒐集流程，限制其在一般訓練環境中的應用。例如，Guo,

Brown, Chan, Chan, and Cheung (2023) 利用單一慣性感測器蒐集多項運動參數，並採用短時傅立葉轉換 (Short-Time Fourier Transform, STFT) 將時域訊號轉換為頻譜圖，接著輸入 CNN 進行特徵學習與分類分析。研究結果顯示，透過單一感測器與深度學習演算法即可有效區分籃球選手的技術水準，證實其在運動表現分析與訓練輔助上的可行性與有效性。然而，此類基於感測器的方式仍存在成本高、布設繁瑣與資料同步困難等限制。

隨著深度學習與電腦視覺技術的進步，研究者逐漸轉向使用攝影機為基礎的 HPE 方法，以取代傳統的穿戴式感測技術。Badiola-Bengoa and Mendez-Zorrilla (2021) 針對 HPE 技術在運動科學領域的應用進行了系統性回顧，指出透過影像可自動預測人體關節位置，並以骨架模型重建動作過程。近年來，隨著深度學習模型的引入，HPE 的精確度與即時性能均獲得顯著提升。其中，OpenPose (Cao et al., 2017) 作為通用型 HPE 系統，已被廣泛應用於不同運動分析情境。Yan et al. (2023) 即利用 OpenPose 從籃球選手的影像中擷取骨架關鍵點資訊，並以 LR 模型預測罰球是否命中。研究結果顯示，OpenPose 能以簡易攝影設備取代傳統動作捕捉系統，達成高精度的姿態辨識，並大幅降低實驗成本與時間。該研究成功展示了基於 OpenPose 的即時姿態分析與資料生成機制，不僅簡化了姿態分析流程，也驗證了其在靜態條件下偵測人體骨架關鍵點的高準確度與實用性。

綜合前述研究可知，結合 HPE 與分類模型的技術可為籃球投籃姿態分析提供新的方向與應用潛力。特別是以 OpenPose 為核心的視覺化方法，能在非侵入式、低成本且高效率的條件下，實現動作分析自動化，為後續運動訓練與表現評估奠定基礎。

參、研究方法與步驟

一、罰球動作優劣辨識系統架構

本研究所設計之罰球動作優劣辨識系統流程如圖 4 所示。系統以影片輸入為起點，對畫面進行逐幀 (Frame-by-Frame) 處理，並依序經過影像辨識階段與特徵工程階段 (Feature Engineering Stage)，以完成投籃動作的分析與分類。在影像辨識階段，系統首先利用經微調後的 YOLOv5 模型進行目標偵測，包括球員與籃球的位置識別。該模型能從輸入影像中自動辨識並框選出人物與球體區域，並據此篩選出具代表性的關鍵幀 (Key Frame) 作為後續分析的輸入資料。圖中上半部顯示在投籃過程中拍攝到球員動作的畫面，包括正面與側面視角；下半部則呈現籃球投射軌跡的影像。接著進入特徵工程階段。此階段首先透過 OpenPose 偵測球員關鍵關節點座標 (Feature Points)，並根據偵測結果計算各關節角度 (Posture Angles)，以量化投籃姿勢特徵。同時，系統亦利用 YOLOv5 模型分析籃球飛行軌跡，進行投射角度的計算，藉此反映球體運動軌跡與姿態動作間的關聯性。最後，系統將姿態角度與球體角度等特徵整合後，輸入至分類模型 (Classification Model)，以進行罰球動作優劣判斷。分類結果分為「Good」與「Bad」兩類，用以評估投籃動作的整體穩定性與技術表現。

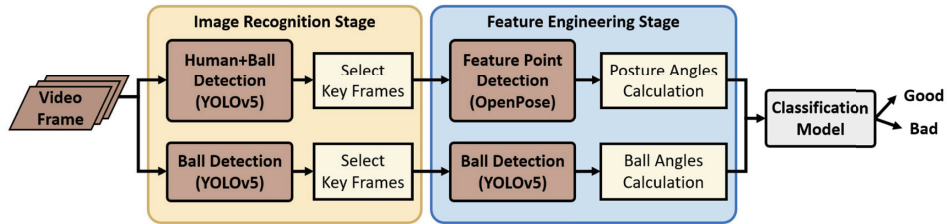


圖 4 罰球動作優劣辨識系統架構圖

資料來源：本研究自行整理。

二、資料收集

本研究所使用之資料為研究參與者進行罰球動作時所錄製之影片。每段影片長度約為 3 ~ 6 秒，影格率 (Frame Rate) 為每秒 30 幀 (Frames per Second [fps])。研究對象為年齡介於 19 ~ 25 歲之學生，身高範圍約為 170 ± 15 公分，涵蓋不同籃球技術程度之受試者，以確保樣本多樣性。為確保資料的一致性與可靠性，所有影片均以高畫質智慧型手機拍攝，並於相同場地條件下進行錄製。投籃拍攝配置如圖 5 所示，採用兩個固定視角進行拍攝：

- (一) 側面視角 (a)：拍攝位置設置於三分線 (b) 外側，距離禁區 (c) 約 4.05 公尺，鏡頭高度固定為 1.30 公尺。此角度用於觀察投籃過程中的動作連貫性與姿態變化。
- (二) 正面視角 (d)：拍攝位置設置於籃框底線正下方，距離罰球線 (e) 約 5.80 公尺，鏡頭高度為 1.60 公尺。此角度可清晰捕捉投籃者全身正面動作與球體運動軌跡。

所有拍攝過程均在光線穩定、背景單純的環境下進行，以降低外部干擾對影像辨識的影響。在進行後續分析前，本研究對原始影片進行裁切 (Cropping) 與格式轉換 (Format Conversion) 等預處理作業，以優化後續人體偵測與 HPE 模型的輸入品質，確保資料分析的準確性與穩定性。

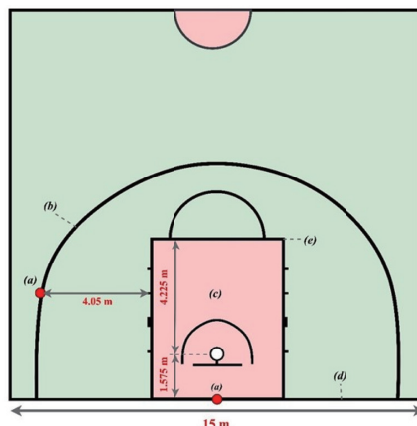


圖 5 球場名稱與拍攝位置說明：

(a) 拍攝位置；(b) 三分線；(c) 限制區；(d) 底線；(e) 罰球線

資料來源：本研究自行整理。

三、資料預處理

本研究的資料預處理流程可分為多個階段，旨在提升影像品質與模型運算效率，並確保輸入資料的一致性與準確性。在研究初期階段，直接以原始高解析度影像作為後續模型分析的輸入資料。然而，在實驗過程中發現，原始影像常包含非目標人物或背景雜訊，例如場邊人員、觀眾席或牆面陰影等，這些不相關資訊容易干擾模型的特徵提取過程，導致辨識準確度下降與運算速度降低。為改善上述問題，本研究首先採用手動影像裁切與尺寸調整方式，僅保留包含目標人物的區域，以降低背景干擾並縮短模型推論時間。雖然此方法能有效提升辨識效率，但在大量資料處理情境下，其人工操作成本較高，難以滿足自動化需求。因此，本研究進一步導入 YOLOv5 模型進行自動化預處理。該方法能直接於原始影像中偵測並定位籃球選手位置，並自動裁切出包含主要動作區域的影像框 (Bounding Box)，以取代人工裁切程序。最終，我們整合上述兩種方法，利用 YOLOv5 模型進行自動化區域裁切與輸出。此流程能快速產生僅含投籃者的影像資料，兼顧精確度與效率，並為後續 HPE (OpenPose) 與分類模型訓練提供更乾淨且結構化的輸入資料。

四、投籃動作與球體軌跡偵測

本研究採用 YOLOv5 與 OpenPose 作為核心偵測模型，分別負責投籃動作與球體軌跡的辨識與特徵擷取，並搭配全連接神經網路 (Fully Connected Neural Network, FCNN) 進行姿態分類，以判斷罰球動作之優劣。

(一) YOLOv5 的應用

本研究首先利用已標註之影像資料對 YOLOv5 模型進行微調 (Fine-Tuning)，使其能準確偵測投籃動作中人物與籃球的位置。在影像辨識流程中 (如圖 4 上、下兩部分所示)，YOLOv5 透過影像框取得人物與籃球的座標資訊。當籃球的 y 座標值小於人物的 y 座標值 (即籃球位於畫面中高於人物時)，系統將此畫面視為投籃關鍵幀。在球體飛行約 10 幀後，擷取第 2 個關鍵幀，以捕捉投射過程。於特徵工程階段，從上述兩個關鍵幀中提取籃球中心點座標 P_1 與 P_2 ，計算其位移向量 V_1 ，再以平行參考向量 (Reference Vector) V_2 計算兩者夾角 θ ，以獲得球體投射角度。相關公式如下：

$$V_1 = P_2 - P_1 \quad (1)$$

$$\theta = \arccos\left(\frac{V_1 \cdot V_2}{\|V_1\| \|V_2\|}\right) \quad (2)$$

$$\theta_{deg} = \frac{\theta}{\pi} \times 180 \quad (3)$$

其中，式 (1) 為球體在連續兩關鍵幀間的位移向量；式 (2) 為位移向量 V_1 與參考向

量 V_2 之間的夾角；式 (3) 則將弧度值轉換為角度值，以便於實際投射角度分析。在此， V_2 為平行參考向量，用於建立基準方向。為了衡量籃球拋射角度相對於地面水平線的夾角，本研究將地面方向視為水平方向向量，並定義 $V_2 = (1, 0)$ 。在三維座標系中，則可定義為 $V_2 = (1, 0, 0)$ 。因此， V_1 代表球體在實際飛行過程中的運動方向，而 V_2 則作為水平基準線。兩者的夾角 θ 即為球體投射角 (Projection Angle)，用以描述投籃拋射的仰角。

此外，為提升 HPE 的準確性，本研究採用級聯式偵測流程 (Cascade Detection Pipeline)，先以 YOLOv5 進行人物偵測，再將裁切後的影像輸入至 OpenPose。此流程可有效排除非目標人物與背景雜訊，減少 HPE 模型在多人場景下的誤差，進而提升姿勢分析的準確度與穩定性。在影像選取部分，根據第二節所述的兩種拍攝視角：

1. 側面視角：從投籃幀往後擷取 5 個關鍵幀，每隔 2 幀取樣一次；
2. 正面視角：從投籃幀往後擷取 3 個關鍵幀，每隔 3 幀取樣一次。

所有關鍵幀皆以 YOLOv5 偵測人物與球體位置，根據偵測框進行影像裁切與尺寸標準化，並將處理後影像輸入 OpenPose 進行姿態辨識。

(二) OpenPose 的應用

經由 YOLOv5 篩選與裁切後的關鍵幀影像將輸入 OpenPose 以進行人體骨架關鍵點偵測。模型輸出共包含 18 個關節點，如圖 6 所示，包括頭部、軀幹、手臂與下肢等主要節點。根據先前研究 (Okubo and Hubbard, 2016)，本研究主要觀察手臂與膝蓋角度變化作為投籃姿勢特徵。在側面視角中，選取 5 個關鍵幀進行手臂與腿部角度計算，完整描述從持球至出手的動作過程；而在正面視角中，由於腿部角度變化較小，僅使用 3 個關鍵幀計算手臂角度，並排除腿部資料。

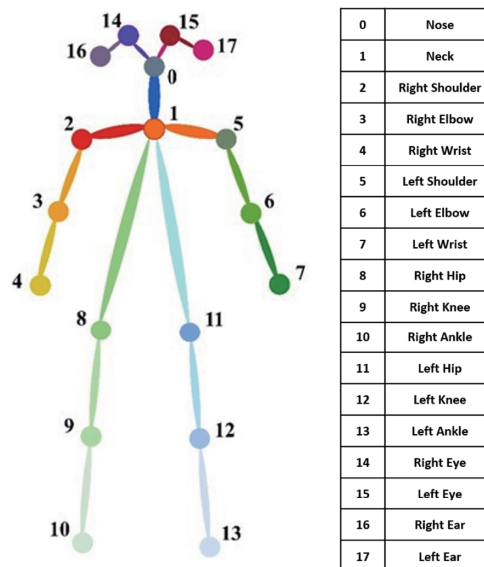


圖 6 人體骨架關節點

資料來源：本研究自行整理。

從 OpenPose 輸出之關節點座標中，分別選取對應關節（例如肩—肘—腕或臀—膝—踝）建立向量，利用下列公式計算兩向量夾角：

$$V_1 = P_2 - P_1, V_2 = P_3 - P_1 \quad (4)$$

$$\theta = \arccos\left(\frac{V_1 \cdot V_2}{|V_1| |V_2|}\right) \quad (5)$$

$$\theta_{deg} = \frac{\theta}{\pi} \times 180 \quad (6)$$

計算所得角度值即為手臂或腿部關節彎曲程度之量化表示。透過對多個關鍵幀進行角度計算，可建立代表整個投籃過程的多維姿態特徵向量，作為後續分類模型輸入資料。

五、資料分類與分類模型評估

為了更精確地分析籃球投籃動作，本研究整合 YOLOv5 與 OpenPose 進行特徵擷取與關鍵幀選取，包含「人物偵測」、「關鍵幀選取」、「關節角度計算」與「特徵生成」四個步驟。首先，透過 YOLOv5 模型偵測人與球的位置，並依據座標關係選取代表投籃過程的關鍵幀。在側面視角中，系統會找出球的座標完全高於人體的影格作為出手關鍵幀（第 5 幀），再往前回推 4 幀，共選取 5 個關鍵幀，以呈現從舉球到出手的完整投籃過程。於正面視角中，由於身體角度變化較小且主要觀察手部動作，因此選取 3 個關鍵幀，表示投籃動作的起始、中間與出手階段。在球體軌跡部分，選取出手幀及其往後第 5 幀 2 個關鍵幀，以計算球之運動軌跡與出手角度。接著，使用 OpenPose 偵測畫面中人物之關節點座標，並依據各節點間的向量關係計算肢體關節角度。側面畫面主要擷取手臂（arm_angle_1 至 arm_angle_5）與腿部（leg_angle_1 至 leg_angle_5）的 5 組角度變化；正面畫面則擷取 3 組手臂角度（arm_angle_1 至 arm_angle_3）；球體軌跡部分則計算出手角度（Shot Release Angle）。最終整合上述 3 個視角的特徵，共生成 14 維姿態與軌跡特徵資料（10 個側面人體手臂與腿部角度、3 個正面手臂角度及 1 個出手角度），輸入至分類模型進行動作優劣辨識，如表 1 所示。本研究選擇 SVC 作為主要分類器。SVC 的基本原理是在高維空間中尋找最佳分離超平面，以最大化不同類別樣本之間的時間（Margin），從而達成精準分類。由於 SVC 在小樣本資料集中具有穩定的分類能力，且特別適用於二元分類問題，因此非常契合本研究「動作優良與否」的分析需求。

表 1 數據特徵欄位表

視角	Data Feature Fields
Side View (Human)	arm_angle_1, arm_angle_2, arm_angle_3, arm_angle_4, arm_angle_5, leg_angle_1, leg_angle_2, leg_angle_3, leg_angle_4, leg_angle_5
Front View (Human)	arm_angle_1, arm_angle_2, arm_angle_3
Side View (Ball)	shot release angle

資料來源：本研究自行整理。

為全面評估分類模型的表現，本研究採用以下四項常見指標：準確率 (Accuracy)、召回率 (Recall)、精確率 (Precision) 與 F1 分數 (F1-Score)。其計算公式如下所示：

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

$$F1-Score = \frac{2 \times (Precision \times Recall)}{Precision + Recall} \quad (10)$$

其中， TP (True Positive) 代表正確分類為「良好動作」的樣本數； FP (False Positive) 為錯誤分類為良好動作的樣本； TN (True Negative) 為正確分類為「不良動作」的樣本； FN (False Negative) 則為誤判為不良動作的樣本。

綜合上述指標，可從不同角度衡量模型在識別正確樣本與避免誤判方面的能力。準確率反映整體預測正確率；精確率著重於分類結果的可靠性；召回率衡量模型捕捉所有正樣本的能力；而 F1 分數則綜合兩者，作為平衡評估指標。這些指標共同構成模型效能的完整評估框架，有助於確保本研究分類結果的可信度與穩定性。

肆、研究結果

一、實驗數據

本研究共蒐集約 30 位受試者的罰球動作影片資料，總計 300 筆影像樣本，每筆資料包含相同動作之正面與側面視角影像。資料拍攝與蒐集條件均依據前一節所述設定進行，以確保樣本的一致性與可比較性。本研究在不同階段採用了三種資料前處理方法。當比較使用原始影像尺寸與整合 YOLOv5 模型定位目標區域之後的偵測結果 (如圖 7 所示) 時，可觀察到兩者間的顯著差異。整合 YOLOv5 的方法能有效提升關鍵點偵測準確度，特別是在多人場景或背景複雜的情況下。然而，部分樣本仍出現少數關鍵點無法成功辨識的情形。相對地，透過手動裁切影像所得結果在關鍵點連續性與完整性上更為穩定。

綜合考量自動化與準確性，本研究最終採取結合式處理流程：

- (一) 先以 YOLOv5 偵測影像中人物位置並獲取其座標；
- (二) 以該座標區域作為影像尺寸調整範圍；

(三) 輸出經自動裁切後的影像，再進行 OpenPose 人體關鍵點偵測。

此混合流程兼顧了 YOLOv5 的定位效率與手動裁切的穩定性，使後續 HPE 能同時達到準確與高效的平衡。

該流程最終產生四種類型的實驗資料（如圖 8 所示），分別對應不同的影像前處理方式，包括原始影像輸入、手動裁切影像、YOLOv5 自動裁切影像，以及結合 YOLOv5 與 OpenPose 之複合方法。本研究針對這四種類型資料進行後續實驗比較與效能評估，以驗證所提出之 HPE 流程在準確度與運算效率上的改良效果。

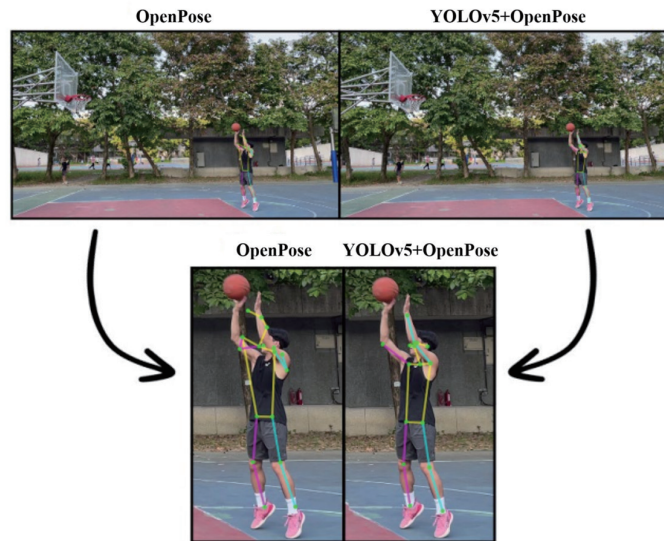


圖 7 比較 OpenPose 與 YOLOv5 + OpenPose 在人體關鍵點偵測效果

資料來源：本研究自行整理。

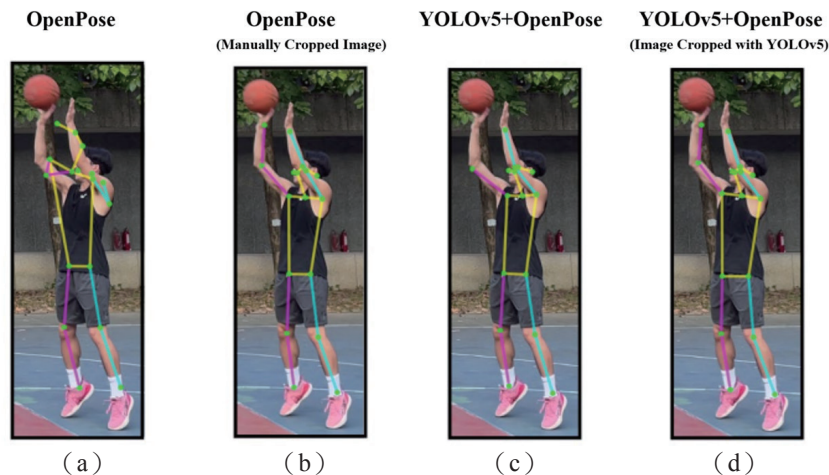


圖 8 四種類型實驗數據的範例：(a) 使用 OpenPose 的原始影像；(b) 使用 OpenPose 手動裁切影像；(c) 使用結合 YOLOv5 的 OpenPose 處理的原始影像；(d) 使用 YOLOv5 自動裁切後並搭配 OpenPose 的影像

資料來源：本研究自行整理。

(一) HPE 處理效率分析

本研究首先針對所蒐集之罰球影像資料集，比較不同前處理方式對 HPE 模型運算效率的影響，其中資料分割比例為 80% 作為訓練集、20% 作為測試集。四種前處理方法分別為：

1. OpenPose (原始影像輸入)；
2. OpenPose (手動裁切影像)；
3. YOLOv5 + OpenPose；
4. YOLOv5 + OpenPose (使用 YOLOv5 自動裁切影像)。

實驗結果如圖 9 與表 2 所示。可觀察到，無論影像拍攝角度為正面或側面，整合 YOLOv5 的方法均可顯著降低每幀影像之平均處理時間。這是因為 YOLOv5 能於初始階段快速識別並定位人體區域，大幅減少 OpenPose 在後續階段需處理的像素範圍，從而有效縮短運算時間並提升整體處理效率。

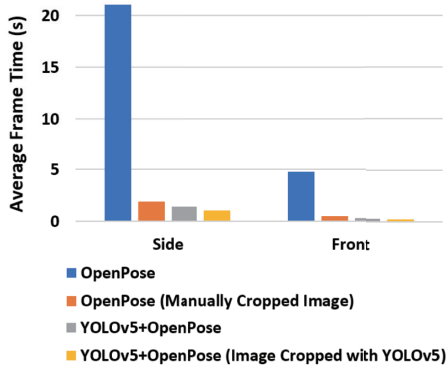


圖 9 每幀影片的平均處理時間

資料來源：本研究自行整理。

表 2 影片每幀平均處理時間

Method	Side	Front
OpenPose	21.022	4.798
OpenPose (Manually Cropped Image)	1.898	0.493
YOLOv5 + OpenPose	1.446	0.336
YOLOv5 + OpenPose (Image Cropped with YOLOv5)	1.061	0.257

資料來源：本研究自行整理。

由表 3 可見，僅使用 OpenPose 時平均處理時間最長，側面影像達 21.022 秒。相較之下，整合 YOLOv5 進行區域裁切後，處理時間降至 1.061 秒（側面）與 0.257 秒（正面），顯示本研究方法可有效提升處理速度超過 20 倍以上。此外，為驗證本研究方法在不同 HPE 架構下的效能表現，我們將結果與 MediaPipe 模型進行比較。MediaPipe 採用輕量化設計，適用於行動裝置與即時應用。比較結果如表 3 所示。雖然 MediaPipe 在處理效率上仍具優勢，但由於其僅偵測 33 個全身關鍵點，而 OpenPose 在啟用身體、手部與臉部模組後可偵測約 135 個關鍵點，因此在動作細節分析與運動學研究中的適用性相對較低。本研究方法在維持完整姿態結構資訊的同時，亦兼顧了良好的即時性能。

表 3 本研究方法與 MediaPipe 之處理效率比較

Method	Side Average Processing Time per Frame (s)	Front Average Processing Time per Frame (s)
YOLOv5 + OpenPose (Image Cropped with YOLOv5)	1.061	0.257
MediaPipe	0.582	0.163

資料來源：本研究自行整理。

(二) 關鍵點偵測準確度分析

為進一步驗證整合 YOLOv5 與 OpenPose 對於投籃動作關鍵點偵測之準確度提升，本研究採用兩項常用評估指標：物件關鍵點相似度 (Object Keypoint Similarity, OKS) (Lin et al., 2014) 與正確關鍵點比例 (Percentage of Correct Keypoints, PCK) (Y. Yang and Ramanan, 2011)。

OKS 用於衡量預測關鍵點與真實關鍵點位置之相似程度，其公式如下：

$$OKS = \frac{\sum_i [\exp(-d_i^2 / 2s^2 k_i^2) \cdot \delta(v_i > 0)]}{\sum_i [\delta(v_i > 0)]} \quad (11)$$

其中， d_i 為預測與真實關鍵點的歐氏距離； s 為尺度參數（通常取人體高度）； k_i 為關鍵點權重； $\delta(v_i > 0)$ 為可見性指示函數。OKS 值越高，表示偵測結果越接近真實標註。

PCK 則用於評估預測關鍵點是否落於真實位置的距離閾值內，其計算公式如下：

$$PCK = \frac{1}{N} \sum_{i=1}^N 1(\text{dist}(p_i, t_i) < \alpha \cdot \max(w, h)) \quad (12)$$

其中， N 為關鍵點總數， p_i 與 t_i 分別為預測與真實關鍵點座標， α 為預設閾值比例常數， w 與 h 為人體邊界框寬與高。PCK 值越高，代表模型在位置辨識上的準確率越佳。

綜合上述結果可得，整合 YOLOv5 的 HPE 流程不僅在處理效率上顯著提升，同時在關鍵點辨識準確度上亦能維持高水準表現，顯示本研究方法在運動動作分析領域中具有實際應用潛力。

(三) 關鍵點偵測準確度比較

如表 4 所示，YOLOv5 與 OpenPose 的整合方法在關鍵點偵測準確度方面表現顯著優於單一模型。整合後的架構在 OKS 及多數 PCK 閾值下皆有明顯提升，特別在高解析度影像中

表現更為突出。此方法不僅能有效提升處理速度，亦強化系統於複雜環境（如光線不足與背景干擾）下的穩定性與適應性，顯示其於即時應用與運動分析系統開發中的潛在價值。

進一步比較不同 HPE 系統的效能，MediaPipe 作為一種輕量化 HPE 工具，將人體檢測與關鍵點預測整合於同一流程中。由表 5 可見，MediaPipe 在單人 HPE 情境下的準確度略高於本研究所提出之方法。然而，本研究所設計的模組化架構具備高度彈性，能根據應用需求調整與優化模型組合，以兼顧運算效率與姿態解析度。

表 4 人體關鍵點偵測評估數據

Method	OKS	PCK@0.1	PCK@0.2	PCK@0.5
OpenPose	0.579	0.539	0.615	0.615
OpenPose (Manually Cropped Image)	0.914	0.769	1.000	1.000
YOLOv5 + OpenPose	0.801	0.769	0.846	0.846
YOLOv5 + OpenPose (Image Cropped with YOLOv5)	0.871	0.846	0.923	0.923

註：OKS：物件關鍵點相似度 (Object Keypoint Similarity)；PCK：正確關鍵點比例 (Percentage of Correct Keypoints)；PCK@0.1 代表極高精度要求，常用於高解析度影像；PCK@0.2 與 PCK@0.5 分別對應高精度與中等精度標準。

資料來源：本研究自行整理。

表 5 兩種方法對人體關鍵點檢測的評估比較

Method	OKS	PCK@0.1	PCK@0.2	PCK@0.5
YOLOv5 + OpenPose (Image Cropped with YOLOv5)	0.871	0.846	0.923	0.923
MediaPipe	0.920	0.917	1.000	1.000

註：OKS：物件關鍵點相似度 (Object Keypoint Similarity)；PCK：正確關鍵點比例 (Percentage of Correct Keypoints)；PCK@0.1 代表極高精度要求，常用於高解析度影像；PCK@0.2 與 PCK@0.5 分別對應高精度與中等精度標準。

資料來源：本研究自行整理。

雖然 MediaPipe 在單人場景下能提供更高的即時效能與準確率，但由於其僅偵測 33 個全身關鍵點，而 OpenPose 可擴展至偵測多達 135 個關節與特徵點（含身體、手部與臉部），因此在動作細節分析與運動學研究中的應用價值更高。綜合結果顯示，本研究提出之 YOLOv5 + OpenPose 整合架構在保持高準確度的同時，顯著提升了處理速度與場景適應性，具備應用於即時運動姿態分析系統的潛力。

二、分類模型實驗結果

本研究以 FCNN (Basha, Dubey, Pulabaigari, and Mukherjee, 2020; Sainath, Vinyals, Senior, and Sak, 2015) 為基礎架構，評估不同 HPE 與資料前處理方法對投籃動作分類效能的影響。進一步，我們亦比較 MLP、LSTM 及傳統機器學習模型 (RF、LR、SVC) 之分類表現，以全面探討不同模型架構於姿態分類任務的適用性。為評估分類器的整體效能，本研究比較四種資料前處理方式與 MediaPipe 所產生之關鍵點輸出，分析其在分類任務中的準確性表現。

根據不同樣本數的訓練結果（表 6），當訓練樣本數為 300 時，僅使用 OpenPose 進行 HPE 導致 84 筆樣本無法正確計算關節角度，而整合 YOLOv5 後，錯誤偵測樣本顯著下降至 24 筆，最終有效樣本數達 290 筆。

表 6 不同資料集採樣結果的準確率

Data Preprocessing Methods	Amount of Data	Identification Errors Number	Classification Accuracy (%)
OpenPose	100	16	69.05
	200	42	70.89
	300	84	71.30
OpenPose (Manually Cropped Image)	100	6	87.41
	200	12	88.74
	300	16	89.35
YOLOv5 + OpenPose	100	10	86.30
	200	18	87.69
	300	24	90.76
YOLOv5 + OpenPose (Image Cropped with YOLOv5)	100	5	88.42
	200	8	90.69
	300	10	92.41
MediaPipe	100	3	79.38
	200	5	82.56
	300	9	82.82

資料來源：本研究自行整理。

根據表 6 結果，後續分析以有效樣本數 290 筆進行。整合 YOLOv5 與 OpenPose 的影像處理方法在分類效能上表現顯著優於單一模型。採用 FCNN 模型對處理後的姿態資料進行訓練時，準確率達 90.76%，如圖 10 所示，相較於僅使用 OpenPose 的情境提升約 19 個百分點。

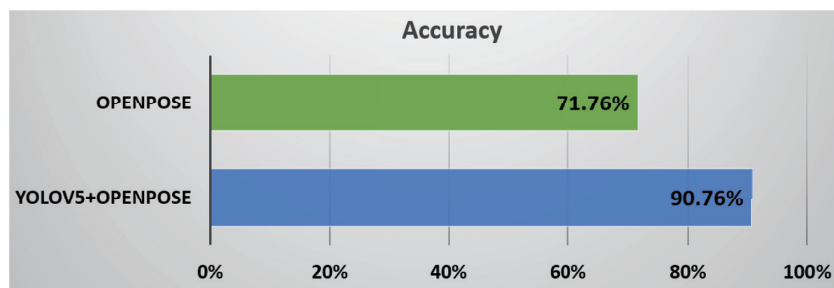


圖 10 結合 YOLOv5 後的準確率提升百分比

資料來源：本研究自行整理。

此外，如圖 11 所示，採用 YOLOv5 自動裁切影像的方法進一步提升準確率至 92.41%，

而 MediaPipe 的分類準確率僅為 82.82%。此結果顯示，整合 YOLOv5 與 OpenPose 的方法不僅能顯著提升姿態估算的整體準確性，亦能克服單一方法在關鍵點遺漏與角度估算誤差方面的限制。

為了進一步驗證本研究姿態特徵資料在不同分類器下的效能表現，我們分別採用多種分類模型進行實驗。每一種分類模型均進行超參數調校，以尋求最佳的精度組合，並比較各模型於相同資料條件下的最終分類準確率，以判定最適合的方案。

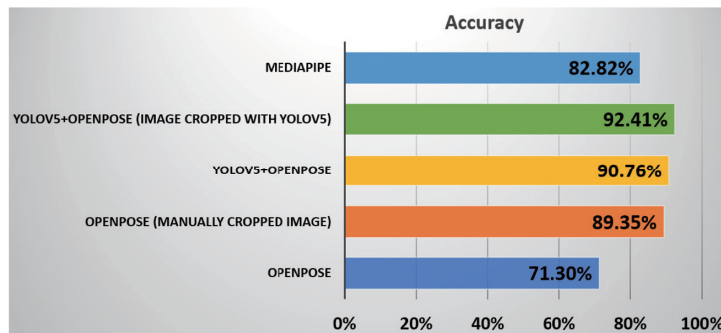


圖 11 不同前處理方法在 FCNN 分類中的準確率比較

資料來源：本研究自行整理。

(一) FCNN 模型結果

本研究設計之 FCNN 架構包含 3 層全連接層 (Fully Connected Layers)，每層各含 10 個神經元。輸入層至隱藏層使用 ReLU (Rectified Linear Unit) 激活函數，輸出層則採用 Sigmoid 函數以進行二元分類，輸出結果轉換為 0 到 1 之間的機率值。如表 7 所示，透過實驗不同神經元數量後發現，10 個神經元能在穩定性與準確率之間達成最佳平衡。

表 7 神經網路中神經元數量的設定

Data Preprocessing Methods	Amount of Data	Identification Errors Number	Classification Accuracy (%)
YOLOv5 + OpenPose (Image Cropped with YOLOv5)	100	10	88.42
	200	10	90.69
	300	10	92.41
MediaPipe	100	10	79.38
	200	10	82.56
	300	10	82.82
YOLOv5 + OpenPose (Image Cropped with YOLOv5)	100	100	40.00
	200	100	53.12
	300	100	52.76
MediaPipe	100	100	41.24
	200	100	53.33
	300	100	53.23

資料來源：本研究自行整理。

在深度學習模型訓練過程中，損失函數與優化器的選擇對收斂速度及最終效能影響顯著。表 8 顯示不同組合的比較結果，最佳表現出現在 BCELoss + Adam 優化器（學習率 0.01）的設定下，準確率達 92.41%，F1 分數為 89.11%。

表 8 FCNN 模型的實驗結果

Classification Model	Parameters		Learning Rate	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
	Loss Function	Optimizer					
FCNN	BCELoss	Adam	0.001	88.28	82.00	83.67	82.83
	BCELoss	Adam	0.010	92.41	86.54	91.84	89.11
	BCELoss	Adam	0.100	53.45	34.96	43.88	38.91
	BCELoss	SGD	0.010	64.14	47.12	50.00	48.51
	MSELoss	Adam	0.010	77.59	70.37	58.16	63.69
	CrossEntropyLoss	Adam	0.010	75.86	61.67	75.51	67.89

註：FCNN：全連接神經網路（Fully Connected Neural Network）。

資料來源：本研究自行整理。

（二）MLP 模型結果

在 MLP 實驗中（表 9），不同激活函數與優化器的組合顯示顯著差異。結果指出，使用“identity”或“tanh”激活函數時可獲得最高分類準確率，且“lbfgs”優化器於小型資料集上具有良好收斂特性。最終在僅含單隱藏層與 50 個神經元的情況下，達成 100% 準確率。

（三）LSTM 模型結果

為捕捉時間序列特徵，本研究構建整合式 LSTM 分類器，輸入特徵包含側面手部角度（5 個時間步），側面腿部角度（5 個時間步），正面手部角度（3 個時間步），投籃釋放角度（1 項特徵）。4 組 LSTM 模組分別提取時間序列特徵後，將最終時間步輸出串接，並通過全連接層與 Sigmoid 函數進行二元分類。根據表 10 結果，2 層 LSTM 與 Hidden_size = 40 可達最佳效能，準確率達 96.55%，F1 分數 94.79%。

（四）RF 模型結果

RF 模型測試不同 n_estimators、max_depth 與 max_features 組合後（表 11），在本研究資料集中未出現顯著性能差異。整體表現穩定，最高準確率達 98.97%。

（五）LR 模型結果

在 LR 模型中，調整正則化強度 C、懲罰項（L1, L2）與不同優化器後（表 12），結果顯示 L2 正則化下的模型在準確率與穩定性均最佳。最終達 99.66% 準確率與 99.49% F1 分數。

表 9 MLP 模型的實驗結果

Classification Model	Parameters			Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
	Hidden_layer_sizes	Activation	Solver				
MLP	(100,)*	'relu'	'adam'	99.31	98.98	98.98	98.98
	(100, 150,)	'relu'	'adam'	99.31	98.98	98.98	98.98
	(150,)	'identity'	'adam'	100.00	100.00	100.00	100.00
	(150,)	'logistic'	'adam'	99.31	98.98	98.98	98.98
	(100,100,)	'logistic'	'adam'	100.00	100.00	100.00	100.00
	(150,)	'tanh'	'adam'	99.66	98.99	100.00	99.49
	(100,100,)	'tanh'	'adam'	100.00	100.00	100.00	100.00
	(50,)	'relu'	'lbfgs'	99.31	98.98	98.98	98.98
	(150,)	'relu'	'lbfgs'	99.31	98.98	98.98	98.98
	(100, 200,)	'relu'	'lbfgs'	99.31	98.98	98.98	98.98
	(100, 150, 100,)	'relu'	'lbfgs'	99.31	98.98	98.98	98.98
	(100,)	'identity'	'lbfgs'	99.66	98.99	100.00	99.49
	(150,)	'identity'	'lbfgs'	100.00	100.00	100.00	100.00
	(100,)	'logistic'	'lbfgs'	99.31	98.00	100.00	98.99
	(50,)	'tanh'	'lbfgs'	100.00	100.00	100.00	100.00
	(150, 50, 50,)	'relu'	'sgd'	99.31	98.00	100.00	98.99

註：MLP：多層感知器（Multilayer Perceptron）；*表示作為比較基準之預設參數配置。

資料來源：本研究自行整理。

表 10 LSTM 模型的實驗結果

Classification Model	Parameters		Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
	num_layers	hidden_size				
LSTM	$n = 1$	$n = 16$	77.24	92.11	35.71	51.47
	$n = 2$	$n = 16$	94.83	95.60	88.78	92.06
	$n = 3$	$n = 16$	86.21	95.31	62.24	75.31
	$n = 2$	$n = 32$	95.52	94.74	91.84	93.26
	$n = 2$	$n = 40$	96.55	96.81	92.86	94.79
	$n = 2$	$n = 48$	95.17	93.75	91.84	92.78
	$n = 2$	$n = 64$	93.79	95.45	85.71	90.32

註：LSTM：長短期記憶網路（Long Short-Term Memory）。

資料來源：本研究自行整理。

表 11 RF 模型的實驗結果

Classification Model	Parameters			Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
	n_estimators	max_depth	max_features				
RF	10	None	'sqrt'	98.51	98.86	96.43	97.62
	100*	None*	'sqrt'	98.69	98.16	97.96	98.06
	10	None	'log2'	98.62	97.96	97.96	97.96
	100	None	'log2'	98.97	98.97	97.96	98.46
	10	None	None	97.90	96.38	97.45	96.91
	100	None	None	98.07	96.67	97.65	97.16

註：RF：隨機森林（Random Forest）；*表示作為比較基準之預設參數配置。
資料來源：本研究自行整理。

表 12 LR 模型的實驗結果

Classification Model	Parameters			Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
	C	penalty	solver				
LR	0.1	'L2'	'lbfgs'	98.97	97.98	98.98	98.48
	1.0*	'L2'	'lbfgs'	99.66	98.99	100.00	99.49
	10.0	'L2'	'lbfgs'	99.66	98.99	100.00	99.49
	0.1	'L2'	'sag'	98.97	97.98	98.98	98.48
	1.0	'L2'	'sag'	99.66	98.99	100.00	99.49
	10.0	'L2'	'sag'	99.31	98.00	100.00	98.99
	0.1	'L2'	'newton-cg'	98.97	97.98	98.98	98.48
	1.0	'L2'	'newton-cg'	99.66	98.99	100.00	99.49
	0.1	'L2'	'liblinear'	97.24	93.27	98.98	96.04
	1.0	'L2'	'liblinear'	98.62	97.00	98.98	97.98
	10.0	'L2'	'liblinear'	99.66	98.99	100.00	99.49
	0.1	'L1'	'liblinear'	96.55	90.74	100.00	95.15
	1.0	'L1'	'liblinear'	99.31	98.00	100.00	98.99
	10.0	'L1'	'liblinear'	99.66	98.99	100.00	99.49
	0.1	'L2'	'saga'	98.97	97.98	98.98	98.48
	1.0	'L2'	'saga'	99.31	98.00	100.00	98.99
	0.1	'L1'	'saga'	98.62	97.00	98.98	97.98
	1.0	'L1'	'saga'	99.31	98.00	100.00	98.99

註：LR：邏輯迴歸（Logistic Regression）；*表示作為比較基準之預設參數配置。
資料來源：本研究自行整理。

(六) SVC 模型結果

本研究測試 SVC 模型不同核函數 (Kernel) 與 C 值的影響 (表 13)。結果顯示, 使用線性 (Linear) 核的模型能完美區分類別, 達成 100% 準確率與 F1 分數, 顯示本研究資料具高度線性可分性, SVC 在此任務中為最具效率且穩定的分類器。

表 13 SVC 模型的實驗結果

Classification Model	Parameters			Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
	C	kernel	gamma				
SVC	0.1	'rbf'	'scale'	96.55	98.89	90.82	94.68
	1.0*	'rbf'	'scale'	99.31	98.98	98.98	98.98
	10.0	'rbf'	'scale'	99.66	100.00	98.98	99.49
	0.1	'linear'	—	100.00	100.00	100.00	100.00
	1.0	'linear'	—	99.31	98.98	98.98	98.98
	1.0	'poly'	'scale'	96.55	100.00	89.80	94.62
	10.0	'poly'	'scale'	98.97	98.97	97.96	98.46
	0.1	'sigmoid'	'scale'	97.93	95.10	98.98	97.00
	1.0	'sigmoid'	'scale'	98.97	97.03	100.00	98.49

註: SVC: 支援向量分類器 (Support Vector Classifier); * 表示作為比較基準之預設參數配置。
資料來源: 本研究自行整理。

(七) 方法比較與總結

表 14 顯示本研究所提出之 YOLOv5 + OpenPose 架構與 MediaPipe 在特徵提取與分類表現上的比較。可見在所有分類模型下, 本研究方法均優於 MediaPipe, 顯示其在特徵穩定性與分類精度方面的優勢。

表 14 不同分類器下兩種方法之分類準確率比較

Method	FCNN	MLP	LSTM	RF	LR	SVC
Our Method	92.41	100.00	96.55	98.97	99.66	100.00
MediaPipe	82.82	87.97	87.63	92.10	87.97	91.41

註: FCNN: 全連接神經網路 (Fully Connected Neural Network); MLP: 多層感知器 (Multilayer Perceptron); LSTM: 長短期記憶網路 (Long Short-Term Memory); RF: 隨機森林 (Random Forest); LR: 邏輯迴歸 (Logistic Regression); SVC: 支援向量分類器 (Support Vector Classifier)。

資料來源: 本研究自行整理。

表 15 彙整各分類器最佳表現結果。可觀察到 LSTM 模型在捕捉時間序列依賴性方面表現突出, 顯示時間資訊在動作識別中的關鍵性。傳統分類器如 MLP 與 SVC 亦展現極高準確率, 特別是 SVC 模型在本研究資料集中達到完美分類表現, 兼具高效能與可解釋性。

表 15 各分類器最佳結果比較

Classification Model	FCNN	MLP	LSTM	RF	LR	SVC
Accuracy (%)	92.41	100.00	96.55	98.97	99.66	100.00
Precision (%)	86.54	100.00	96.81	98.97	98.99	100.00
Recall (%)	91.84	100.00	92.86	97.96	100.00	100.00
F1-Score (%)	89.11	100.00	94.79	98.46	99.49	100.00

註：FCNN：全連接神經網路（Fully Connected Neural Network）；MLP：多層感知器（Multilayer Perceptron）；LSTM：長短期記憶網路（Long Short-Term Memory）；RF：隨機森林（Random Forest）；LR：邏輯迴歸（Logistic Regression）；SVC：支援向量分類器（Support Vector Classifier）。

資料來源：本研究自行整理。

綜合分析結果顯示，LSTM 能有效捕捉投籃動作的時間序列特性，而 SVC 則展現最高的分類效能與泛化能力。考量分類精度、運算效率及模型可解釋性，本研究最終選擇 SVC 作為本系統之主要分類模型。該模型能在多種線性與非線性特徵空間下保持穩定表現，為運動姿態分析與訓練輔助系統中極具應用潛力的方案。

伍、結論與未來研究方向

本研究結果顯示，雖然 OpenPose 本身已具備穩定且高效的 HPE 能力，但結合 YOLOv5 之後，對於籃球投籃等複雜動作的辨識與分析準確性均有顯著提升。該整合方法不僅有效改善了關鍵點偵測的精確度與處理效率，亦大幅提升系統在動態動作與複雜背景下的適應性。對於運動分析與運動科學領域的研究人員與實務教練而言，本研究所提出之架構提供了一項具備即時性與高可擴展性的工具，可作為科學化訓練與運動表現評估的重要輔助系統。本研究採用的兩階段模組化設計（YOLOv5 物件偵測 + OpenPose HPE），在系統性與靈活性上具有明顯優勢。該設計允許各階段獨立優化，使 HPE 能在高品質邊界框範圍內執行，有效提升人體關鍵點預測的穩定性與準確度。此外，透過多種分類模型的比較分析結果可知，SVC 分類器在本研究資料集上達成 100% 的分類準確率，展現了其於運動動作分類任務中的卓越表現。此結果驗證深度學習技術在運動姿態識別與訓練分析應用上的潛力，亦顯示本系統可作為教練進行動作矯正與技術精進的輔助依據，從而促進運動訓練的智能化與數據化發展。儘管本研究之系統在 HPE 與分類效能上均有顯著成果，但仍存在以下技術限制：

- 一、光線變化影響準確度：在光照不足、背光或強反光場景中，關鍵點偵測精度仍可能下降。
- 二、多人遮擋問題：當畫面中出現多位運動員或肢體重疊時，模型辨識結果可能產生錯誤對應。
- 三、即時運算負擔：YOLOv5 + OpenPose 結合雖提升準確率，但相對增加運算成本，需進一步優化以適應行動裝置或邊緣運算環境。

在未來研究方面，本研究提出以下幾項可行之改進方向，以提升現有方法的精度與應用廣度：

- 一、技術層面改進：導入更先進的深度學習架構（如 Transformer-Based 或 Graph Neural Network 模型），並結合精密的特徵提取與資料增強技術，以進一步提升肢體關鍵點辨識的準確度與對複雜動作的解析能力。
- 二、分類模型優化：透過整合多模態資訊（如動作序列時間特徵與球體軌跡動態特徵）以及引入輕量化神經網絡，可在兼顧計算效率的同時提升分類效能，實現更即時的分析能力。
- 三、跨領域應用：未來可將本研究所建立之模型架構擴展應用於其他領域，例如不同運動項目（如網球、羽球、高爾夫）、物理治療與復健訓練、人因工程分析，甚至互動式體感遊戲與虛擬實境（Virtual Reality, VR）／擴增實境（Augmented Reality, AR）系統。此舉不僅能驗證模型於多場景下的通用性與穩定性，亦可推動運動姿態分析技術的跨領域發展與實務應用價值。

綜上所述，本研究成功建構出一個結合 YOLOv5 與 OpenPose 之運動 HPE 與分類系統，並藉由多模型比較驗證其準確性與效能。未來隨著深度學習技術與運動感測設備的持續進步，本研究方法可望進一步拓展至智慧訓練輔助、動作矯正回饋系統與運動表現評估平臺等領域，為智慧運動科技的發展奠定基礎。

參考文獻

- Andriluka M., Roth S., and Schiele B., 2009, "Pictorial Structures Revisited: People Detection and Articulated Pose Estimation," in *2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Los Alamitos, CA: IEEE Computer Society Press, 1014-1021. doi:10.1109/CVPR.2009.5206754
- Awais M., Iqbal M. J., Ahmad I., Alassafi M. O., Alghamdi R., Basher M., and Waqas M., 2019, "Real-Time Surveillance through Face Recognition Using HOG and Feedforward Neural Networks," *IEEE Access*, 7, 121236-121244. doi:10.1109/ACCESS.2019.2937810
- Bachir N. and Memon Q. A., 2024, "Benchmarking YOLOv5 Models for Improved Human Detection in Search and Rescue Missions," *Journal of Electronic Science and Technology*, 22(1), 100243. doi:10.1016/j.jnlest.2024.100243
- Badiola-Bengoa A. and Mendez-Zorrilla A., 2021, "A Systematic Review of the Application of Camera-Based Human Pose Estimation in the Field of Sport and Physical Exercise," *Sensors*, 21(18), 5996. doi:10.3390/s21185996
- Bardou D., Zhang K., and Ahmad S. M., 2018, "Classification of Breast Cancer Based on Histology Images Using Convolutional Neural Networks," *IEEE Access*, 6, 24680-24693. doi:10.1109/ACCESS.2018.2831280
- Basha S. H. S., Dubey S. R., Pulabaigari V., and Mukherjee S., 2020, "Impact of Fully Connected Layers on Performance of Convolutional Neural Networks for Image Classification," *Neurocomputing*, 378, 112-119. doi:10.1016/j.neucom.2019.10.008

- Breiman L., 2001, "Random Forests," *Machine Learning*, 45, 5-32. doi:10.1023/A:1010933404324
- Cao Z., Simon T., Wei S.-E., and Sheikh Y., 2017, "Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Los Alamitos, CA: IEEE Computer Society Press, 1302-1310. doi:10.1109/CVPR.2017.143
- Fang H.-S., Li J., Tang H., Xu C., Zhu H., Xiu Y., Li Y.-L., and Lu C., 2022, "AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(6), 7157-7173. doi:10.1109/TPAMI.2022.3222784
- Girshick R., 2015, "Fast R-CNN," in *2015 IEEE International Conference on Computer Vision (ICCV)*, Los Alamitos, CA: IEEE Computer Society Press, 1440-1448. doi:10.1109/ICCV.2015.169
- Girshick R., Donahue J., Darrell T., and Malik J., 2014, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Los Alamitos, CA: IEEE Computer Society Press, 580-587. doi:10.1109/CVPR.2014.81
- Go R. and Aoki Y., 2016, "Flexible Top-View Human Pose Estimation for Detection System via CNN," in *2016 IEEE 5th Global Conference on Consumer Electronics*, Los Alamitos, CA: IEEE Computer Society Press, 537-540. doi:10.1109/GCCE.2016.7800406
- Guo X., Brown E., Chan P. P. K., Chan R. H. M., and Cheung R. T. H., 2023, "Skill Level Classification in Basketball Free-Throws Using a Single Inertial Sensor," *Applied Sciences*, 13(9), 5401. doi:10.3390/app13095401
- Gupta V. K., Gupta A., Kumar D., and Sardana A., 2021, "Prediction of COVID-19 Confirmed, Death, and Cured Cases in India Using Random Forest Model," *Big Data Mining and Analytics*, 4(2), 116-123. doi:10.26599/BDMA.2020.9020016
- Hearst M. A., Dumais S. T., Osuna E., Platt J., and Scholkopf B., 1998, "Support Vector Machines," *IEEE Intelligent Systems and Their Applications*, 13(4), 18-28. doi:10.1109/5254.708428
- Hochreiter S. and Schmidhuber J., 1997, "Long Short-Term Memory," *Neural Computation*, 9(8), 1735-1780. doi:10.1162/neco.1997.9.8.1735
- Ji R., 2020, "Research on Basketball Shooting Action Based on Image Feature Extraction and Machine Learning," *IEEE Access*, 8, 138743-138751. doi:10.1109/ACCESS.2020.3012456
- Johnson S. and Everingham M., 2011, "Learning Effective Human Pose Estimation from Inaccurate Annotation," in *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition*, Los Alamitos, CA: IEEE Computer Society Press, 1465-1472. doi:10.1109/CVPR.2011.5995318
- Jorge J., Giménez A., Silvestre-Cerdà J. A., Civera J., Sanchis A., and Juan A., 2022, "Live Streaming Speech Recognition Using Deep Bidirectional LSTM Acoustic Models and Inter-

- polated Language Models,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 30, 148-161. doi:10.1109/TASLP.2021.3133216
- Jubayer F., Soeb J. A., Mojumder A. N., Paul M. K., Barua P., Kayshar S., Akter S. S., Rahman M., and Islam A., 2021, “Detection of Mold on the Food Surface Using YOLOv5,” *Current Research in Food Science*, 4, 724-728. doi:10.1016/j.crf.2021.10.003
- Karim F., Majumdar S., Darabi H., and Chen S., 2018, “LSTM Fully Convolutional Networks for Time Series Classification,” *IEEE Access*, 6, 1662-1669. doi:10.1109/ACCESS.2017.2779939
- Ker J., Wang L., Rao J., and Lim T., 2018, “Deep Learning Applications in Medical Image Analysis,” *IEEE Access*, 6, 9375-9389. doi:10.1109/ACCESS.2017.2788044
- Lamane M., Tabaa M., and Klilou A., 2022, “Classification of Targets Detected by mmWave Radar Using YOLOv5,” *Procedia Computer Science*, 203, 426-431. doi:10.1016/j.procs.2022.07.056
- Li H., Xiong P., An J., and Wang L., 2018, “Pyramid Attention Network for Semantic Segmentation,” *arXiv*, <https://doi.org/10.48550/arXiv.1805.10180> (accessed August 10, 2024).
- Li L., Mu X., Li S., and Peng H., 2020, “A Review of Face Recognition Technology,” *IEEE Access*, 8, 139110-139120. doi:10.1109/ACCESS.2020.3011028
- Lin T.-Y., Dollár P., Girshick R., He K., Hariharan B., and Belongie S., 2017, “Feature Pyramid Networks for Object Detection,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Los Alamitos, CA: IEEE Computer Society Press, 936-944. doi:10.1109/CVPR.2017.106
- Lin T.-Y., Maire M., Belongie S., Hays J., Perona P., Ramanan D., Dollár P., and Zitnick C. L., 2014, “Microsoft COCO: Common Objects in Context,” in *Computer Vision—ECCV 2014*, Cham, Switzerland: Springer, 740-755. doi:10.1007/978-3-319-10602-1_48
- Muhammad K., Ullah A., Lloret J., Del Ser J., and de Albuquerque V. H. C., 2021, “Deep Learning for Safe Autonomous Driving: Current Challenges and Future Directions,” *IEEE Transactions on Intelligent Transportation Systems*, 22(7), 4316-4336. doi:10.1109/TITS.2020.3032227
- Munea T. L., Jembre Y. Z., Weldegebriel H. T., Chen L., Huang C., and Yang C., 2020, “The Progress of Human Pose Estimation: A Survey and Taxonomy of Models Applied in 2D Human Pose Estimation,” *IEEE Access*, 8, 133330-133348. doi:10.1109/ACCESS.2020.3010248
- Murtagh F., 1991, “Multilayer Perceptrons for Classification and Regression,” *Neurocomputing*, 2(5-6), 183-197. doi:10.1016/0925-2312(91)90023-5
- Nakazawa T. and Kulkarni D. V., 2018, “Wafer Map Defect Pattern Classification and Image Retrieval Using Convolutional Neural Network,” *IEEE Transactions on Semiconductor Manufacturing*, 31(2), 309-314. doi:10.1109/TSM.2018.2795466
- Neubeck A. and Van Gool L., 2006, “Efficient Non-Maximum Suppression,” in *18th International Conference on Pattern Recognition (ICPR’06)*, Los Alamitos, CA: IEEE Computer Society Press, 850-855. doi:10.1109/ICPR.2006.479

- Okazaki V. H. A., Rodacki A. L. F., and Satern M. N., 2015, "A Review on the Basketball Jump Shot," *Sports Biomechanics*, 14(2), 190-205. doi:10.1080/14763141.2015.1052541
- Okubo H. and Hubbard M., 2016, "Comparison of Shooting Arm Motions in Basketball," *Procedia Engineering*, 147, 133-138. doi:10.1016/j.proeng.2016.06.202
- Oproescu M., Mihaila I., Bizon N., Tudosoiu D., Iana V. G., Fleancu L. J., Cojanu F., Potop V., Mihai I., and Popescu D. C., 2023, "Detection and Analysis of Motricity and Somato-Functional Indices of 3×3 Basketball Players to Customize Sports Training," in *2023 15th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, Los Alamitos, CA: IEEE Computer Society Press, 525-530. doi:10.1109/ECAI58194.2023.10194186
- O'Shea K. and Nash R., 2015, "An Introduction to Convolutional Neural Networks," *arXiv*, <https://doi.org/10.48550/arXiv.1511.08458> (accessed August 10, 2024).
- Pishchulin L., Insafutdinov E., Tang S., Andres B., Andriluka M., Gehler P. V., and Schiele B., 2016, "DeepCut: Joint Subset Partition and Labeling for Multi Person Pose Estimation," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Los Alamitos, CA: IEEE Computer Society Press, 4929-4937. doi:10.1109/CVPR.2016.533
- Rajchl M., Lee M. C. H., Oktay O., Kamnitsas K., Passerat-Palmbach J., Bai W., Damodaram M., Rutherford M. A., Hajnal J. V., Kainz B., and Rueckert D., 2017, "DeepCut: Object Segmentation from Bounding Box Annotations Using Convolutional Neural Networks," *IEEE Transactions on Medical Imaging*, 36(2), 674-683. doi:10.1109/TMI.2016.2621185
- Redmon J., Divvala S., Girshick R., and Farhadi A., 2016, "You Only Look Once: Unified, Real-Time Object Detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Los Alamitos, CA: IEEE Computer Society Press, 779-788. doi:10.1109/CVPR.2016.91
- Redmon J. and Farhadi A., 2017, "YOLO9000: Better, Faster, Stronger," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Los Alamitos, CA: IEEE Computer Society Press, 6517-6525. doi:10.1109/CVPR.2017.690
- Redmon J. and Farhadi A., 2018, "YOLOv3: An Incremental Improvement," *arXiv*, <https://doi.org/10.48550/arXiv.1804.02767> (accessed August 10, 2024).
- Ren S., He K., Girshick R., and Sun J., 2017, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137-1149. doi:10.1109/TPAMI.2016.2577031
- Rezatofighi H., Tsoi N., Gwak J., Sadeghian A., Reid I., and Savarese S., 2019, "Generalized Intersection over Union: A Metric and a Loss for Bounding Box Regression," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Los Alamitos, CA: IEEE Computer Society Press, 658-666. doi:10.1109/CVPR.2019.00075
- Sainath T. N., Vinyals O., Senior A., and Sak H., 2015, "Convolutional, Long Short-Term Memory, Fully Connected Deep Neural Networks," in *2015 IEEE International Conference on Acous-*

- tics, Speech and Signal Processing (ICASSP)*, Los Alamitos, CA: IEEE Computer Society Press, 4580-4584. doi:10.1109/ICASSP.2015.7178838
- Seo J. and Park H., 2019, "Object Recognition in Very Low Resolution Images Using Deep Collaborative Learning," *IEEE Access*, 7, 134071-134082. doi:10.1109/ACCESS.2019.2941005
- Sheykhmousa M., Mahdianpari M., Ghanbari H., Mohammadimanesh F., Ghamisi P., and Homayouni S., 2020, "Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 6308-6325. doi:10.1109/JSTARS.2020.3026724
- Song S., Lam J. C. K., Han Y., and Li V. O. K., 2020, "ResNet-LSTM for Real-Time PM2.5 and PM10 Estimation Using Sequential Smartphone Images," *IEEE Access*, 8, 220069-220082. doi:10.1109/ACCESS.2020.3042278
- Toshev A. and Szegedy C., 2014, "DeepPose: Human Pose Estimation via Deep Neural Networks," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Los Alamitos, CA: IEEE Computer Society Press, 1653-1660. doi:10.1109/CVPR.2014.214
- Tran C. M. and Silverberg L. M., 2008, "Optimal Release Conditions for the Free Throw in Men's Basketball," *Journal of Sports Sciences*, 26(11), 1147-1155. doi:10.1080/02640410802004948
- Wang J., Yu L.-C., Lai K. R., and Zhang X., 2020, "Tree-Structured Regional CNN-LSTM Model for Dimensional Sentiment Analysis," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 28, 581-591. doi:10.1109/TASLP.2019.2959251
- Wei S.-E., Ramakrishna V., Kanade T., and Sheikh Y., 2016, "Convolutional Pose Machines," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Los Alamitos, CA: IEEE Computer Society Press, 4724-4732. doi:10.1109/CVPR.2016.511
- Yan W., Jiang X., and Liu P., 2023, "A Review of Basketball Shooting Analysis Based on Artificial Intelligence," *IEEE Access*, 11, 87344-87365. doi:10.1109/ACCESS.2023.3304631
- Yang L., Xie K., Wen C., and He J.-B., 2021, "Speech Emotion Analysis of Netizens Based on Bidirectional LSTM and PGCDNB," *IEEE Access*, 9, 59860-59872. doi:10.1109/ACCESS.2021.3073234
- Yang Y. and Ramanan D., 2011, "Articulated Pose Estimation with Flexible Mixtures-of-Parts," in *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition*, Los Alamitos, CA: IEEE Computer Society Press, 1385-1392. doi:10.1109/CVPR.2011.5995741
- Zhang P., Zhong Y., and Li X., 2019, "SlimYOLOv3: Narrower, Faster and Better for Real-Time UAV Applications," in *2019 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, Los Alamitos, CA: IEEE Computer Society Press, 37-45. doi:10.1109/ICCVW.2019.00011
- Zhou Y., Mok P. Y., and Zhou S., 2019, "A Part-Based Deep Neural Network Cascade Model for Human Parsing," *IEEE Access*, 7, 160101-160111. doi:10.1109/ACCESS.2019.2951182

Zou X., Hu Y., Tian Z., and Shen K., 2019, “Logistic Regression Model Optimization and Case Analysis,” in *2019 IEEE 7th International Conference on Computer Science and Network Technology (ICCSNT)*, Los Alamitos, CA: IEEE Computer Society Press, 135-139. doi:10.1109/ICCSNT47585.2019.8962457