

Article

Exploration of Applying Pose Estimation Techniques in Table Tennis

Chih-Hung Wu ^{1,*}, Te-Cheng Wu ² and Wen-Bin Lin ^{3,*}

¹ Department of Digital Content and Technology, National Taichung University of Education, Taichung 403, Taiwan

² Physical Education Office, National Tsing Hua University, Hsinchu 300, Taiwan

³ Physical Education Center, Taipei National University of the Arts, Taipei City 112, Taiwan

* Correspondence: chwu@mail.ntcu.edu.tw (C.-H.W.); winbeing@ccnia.tnu.edu.tw (W.-B.L.)

Highlights:

Originality/Contribution: The newly developed computer vision pose estimation technique in artificial intelligence (AI) is an emerging technology with potential advantages, such as high efficiency and contactless detection, for improving competitive advantage in the sports industry. The contributions of this study are described below.

What are the main findings?

- Our study examines how to implement AI for table tennis pose recognition.
- Applications, benefits, and limitations of pose recognition in table tennis are discussed.

What are the implications of the main findings?

- A two-stage process for an effective AI pose estimation method is proposed.
- Perspectives of training and tactics for applying pose estimation are discussed.

Abstract: The newly developed computer vision pose estimation technique in artificial intelligence (AI) is an emerging technology with potential advantages, such as high efficiency and contactless detection, for improving competitive advantage in the sports industry. The related literature is currently lacking an integrated and comprehensive discussion about the applications and limitations of using the pose estimation technique. The purpose of this study was to apply AI pose estimation techniques, and to discuss the concepts, possible applications, and limitations of these techniques in table tennis. This study implemented the OpenPose pose algorithm in a real-world video of a table tennis game. The research results show that the pose estimation algorithm performs well in estimating table tennis players' poses from the video in a graphics processing unit (GPU)-accelerated environment. This study proposes an innovative two-stage AI pose estimation method for effectively addressing the current difficulties in applying AI to table tennis players' pose estimation. Finally, this study provides several recommendations, benefits, and various perspectives (training vs. tactics) of table tennis and pose estimation limitations for the sports industry.

Keywords: pose estimation; computer vision; training and tactics analysis; table tennis



Citation: Wu, C.-H.; Wu, T.-C.; Lin, W.-B. Exploration of Applying Pose Estimation Techniques in Table Tennis. *Appl. Sci.* **2023**, *13*, 1896. <https://doi.org/10.3390/app13031896>

Academic Editors: Božidar Šarler, Haiping Liu, Jian Zhang and Jiaqi Li

Received: 5 December 2022

Revised: 27 January 2023

Accepted: 30 January 2023

Published: 1 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Sports science is a cross-disciplinary applied science, and its research and development results and benefits can be directly applied to the public; it can have a deep and specific impact in competitive sports, national health, preventive medicine, or enhancing the economic development of the sports industry [1]. The Ministry of Science and Technology of Taiwan [2] has noted that the sports industry is already mature in Europe and the U.S., with significant economic scale and benefits, and that the annual growth rate of the global

sports industry is estimated to exceed 6%. However, compared with the relatively well-developed European and American countries, there is still considerable room for growth in the Asian sports industry and market. The future of the sports industry can be further developed through close cooperation with sports science research.

The integration of computer vision in artificial intelligence (AI) technology into sports training and competition is important for two reasons. One is the application of AI to bring sports viewing to a new level of development and enhance the visual perception of sports viewing. Second, the application of gesture recognition in technical analysis to improve the competitiveness of the game is a new trend in the development of technology integration in sports. With the emphasis on precision sports research and academic research, the application of “training” and “tactical information collection” has become crucial. With the rapid advancement of technology, the importance of sports technology in assisting with accurate scientific research is growing. The rapid development of wearable devices in recent years has resulted in a boom in applications in sports fields due to an increase in the accuracy and miniaturization of sensor elements. The sensor or wearable device is used to detect sports posture, and the acceleration information and three-axis sensor information of the sensor are collected to analyze the athlete’s movement [3]. However, the limitation is that it must be worn; the requirement to wear the relevant device results in some differences from the real sports situation, and due to device limitations and costs, some sports items or situations will be difficult to popularize and apply. Nevertheless, with the development of posture recognition technology in computer vision, if the athletes’ posture in the video can be directly analyzed through the computer, benefits of real-time, noncontact analysis can be achieved. When analyzing the posture, the athletes do not need to wear any device, the situation is the same as the competition situation, and the video of the competition can be analyzed afterward. Therefore, introducing this technology will reduce the cost and achieve popularization.

Accurate real-time recognition of sports posture will greatly benefit the information collection at the “training level” for posture correction, and at the “competition level” for analysis of technical and tactical use. Therefore, technology-assisted precision sports research has become a crucial research issue. Technology can be introduced and integrated into sports technology and assistance in a timely manner, which will enhance sports and improve the national sports culture and athletic strength. The purposes of this study were to analyze the applications and limitations of pose estimation/recognition for motion analysis in computer vision in AI. This study also discusses past research, practical applications, limitations, and future trends of pose recognition technology.

2. Literature Review

2.1. Computer Vision Technique for Gesture Recognition Applications

Due to the development of technology, there is a long history of development of algorithmic techniques for tracking human movement using images [4,5]. However, in the past, studies often relied on a single image or sensor to provide visual data, which also led to inconvenience in the operation of the subject and limitations in the recording of the object [6,7]. Recent advances in image recognition technology have produced a variety of visual recognition systems that make it possible to more easily perform pose estimation on test subjects; thus, these systems are gradually being applied to motion pose estimation [8,9].

Pose estimation, a computer vision technology in AI, has good potential for development and application in sports and fitness activities, motion analysis, and 3D clothing fitting. Currently, various techniques have been developed for body pose estimation, for instance, the Mask R-CNN technique [10] or AlphaPose [11–13], which is now available in v0.3.0. A multipose estimation system can simultaneously achieve a mean average precision (mAP) of over 60 and multiple object tracker accuracy (MOTA) of over 50 on the PoseTrack Challenge dataset [14]. OpenPose can simultaneously integrate body posture estimation, face estimation, and hand and leg posture estimation [15]. The OpenPose tech-

nology provides a real-time multiperson pose recognition system. It can instantly analyze the body, hand, face, and foot posture key point estimation. It provides up to 135 posture key points in total. The related programs and descriptions are available on Carnegie Mellon University GitHub (<https://github.com/CMU-Perceptual-Computing-Lab/openpose>, access date: 1 October 2022), and its technical development is based on published research and extensions [15,16]. Figure 1 shows the execution results of two studies developed for face and hand gesture recognition [15].

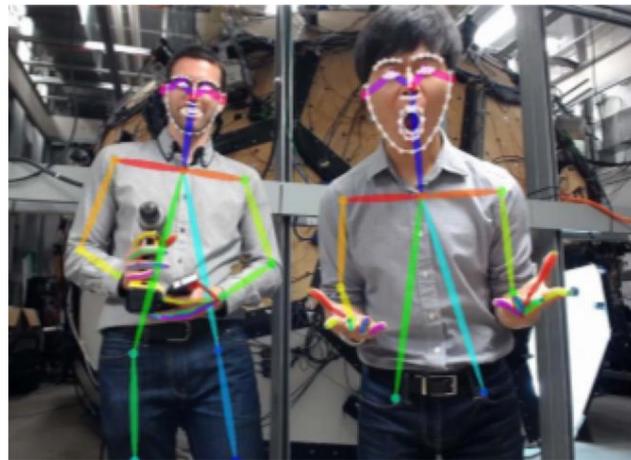


Figure 1. OpenPose detects the key points of body posture, facial expression, hands, and legs of multiple people at the same time. Photo credits: <https://arxiv.org/pdf/1812.08008.pdf> (access date: 1 October 2022).

2.2. Posture Estimation during Movement

Using posture estimation techniques in sports remains an emerging research area. Deep learning techniques for human activity recognition are a challenging task in computer vision [17,18]. The advancement of technology enables the application of deep learning techniques to real-time motion pose detection through pose estimation in computer vision [19]. To apply deep learning to sports, scholars have used a hybrid deep learning architecture that combines a convolutional neural network (CNN) and a long short-term memory (LSTM) network for yoga posture recognition [20], pingpong ball drop detection, and billiard ball drop detection [21]. Body sensor networks (BSNs) are used to collect motion data, and sensors are placed on the upper arm, lower arm, and back to collect the above motion information. Then, dimension reduction by principal component analysis, and finally, a support vector machine (SVM) for machine learning, were used to analyze the table tennis strokes detection [3]; the SVM analyzed the table tennis strokes videos. A 240 fps image was captured with an RGB high-speed camera, which was reduced to 30 fps to accelerate the stance and ball estimations. The stance estimation was performed with a residual CNN network to estimate the player's body stance. The LSTM model predicted the ball drop using only 10 joint positions in two dimensions of the upper body [19].

In addition to body posture estimation, most current literature on computer vision applications in table tennis focuses on applying racket posture estimation to table tennis robots. Previous research has proposed the detection of the image of the racket area during the motion of a table tennis ball with a modified LineSegmentDetector (LSD) algorithm to analyze the vertices and rectangular areas of the racket image, and, finally, integrate the center and pose of the 3D racket area to obtain the pose and position of the racket using a PnP positioning method to implement the application in a table tennis robot [22]. Chen et al. [23] used a high-speed monocular vision system to track the trajectory of the racket to estimate the ball rotation and proposed a novel and effective feature filtering method to screen out important features of the racket's posture during a collision. A new combined HSV and RGB color space algorithm was used to estimate the racquet stance

using real-time racquet image detection [24]. The estimation of the image's rectangular area was shown to be effective. Recently, some scholars [21] used RGB camera images to analyze the sender's motions through a long short-term pose prediction network to accurately predict the landing point of the serve [18]. To enhance the learning of video features, a topological sparse encoder was constructed for semi-supervised learning, which could effectively enhance the application of computer vision technology to table tennis video pose recognition.

3. Methodology

This study analyzed the applications and limitations of pose estimation/recognition for motion analysis in computer vision in AI. This paper discusses past research, practical applications, limitations, and future trends of pose recognition technology. To select the related literature, the following databases were searched: Science Direct, IEEE Xplore, ISI Web of Science, Airiti Library, and Google Scholar. Because the technology is developing rapidly, many of the latest technical development studies are published on arXiv and GitHub. Therefore, we also searched for the latest studies on arXiv.org and GitHub. The search period was up to 2022, and the keywords of pose estimation, pose recognition, and table tennis were used to search and filter, in order to adopt the appropriate pose recognition tool in this study. Final, this study adopted the OpenPose algorithm to develop our pose recognition system for table tennis.

3.1. Pose Estimation System Development

The OpenPose algorithm currently supports platforms such as Ubuntu (14, 16), Windows (8, 10), and Mac OSX, embedded systems such as the Nvidia TX2, and various computing hardware environments including graphics processing unit (GPU) graphics cards, CUDA GPUs (Nvidia GPUs, Nvidia, Santa Clara, CA, USA), OpenCL GPUs (AMD GPUs, AMD, Santa Clara, CA, USA), and individual central processing unit (CPU) computing environments.

The input video sources include single photos (image), videos (video), webcam, and IP camera streams [15]. The video feeds include a single photo (image), video, webcam, and IP camera streaming. After OpenPose acquires the input data, the core of the computation includes three main modules: (1) body + leg pose recognition, (2) hand pose recognition, and (3) face recognition. The body posture model is trained with COCO [25] and MPII Human Pose [26] datasets. The output comprises the original picture + key points (PNG, JPG), original video + key points (AVI), and key point storage (JSON, XML, YML format) [15]; 2D multiplayer key points are instantly recognized. A total of 15, 18, or 25 key points are recognized on the body and legs, 21 on each hand, and 70 on the face [27]. Figure 2 shows the overall posture recognition system flow.

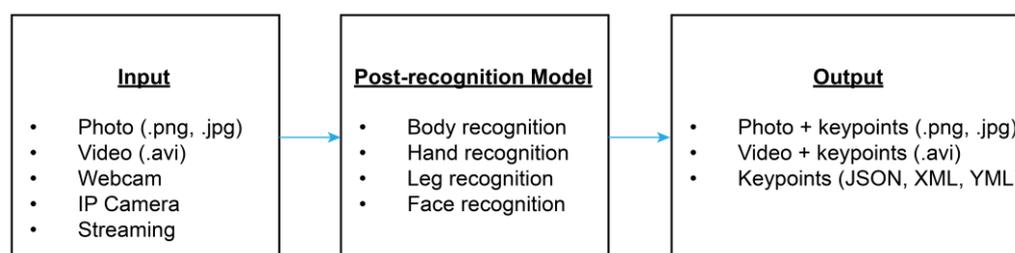


Figure 2. Flow of the OpenPose gesture recognition system.

3.2. Posture Analysis Technology Application and Limitation Analysis

This study was an empirical analysis investigation of gesture recognition technology applied to motion video analysis. The host specifications are among the high-end devices of the current personal host, and an RTX2070 high-end graphics card can handle computer identification in an environment requiring high computing power.

The test host specifications are as follows:

Intel i7-8700 CPU @3.2 G

64 G RAM

RTX 2070 graphics card

Windows 10 Pro operating system

The actual processing time for photos and videos is shown in Table 1 and Figure 3 for both CPU and GPU versions. It shows that the CPU version takes 107.73 s and the GPU version takes 5.07 s to process 22 photos in a batch. If a built-in sample video file (1.33 MB, 4 s video length) is processed, the CPU version takes 1313 s and the GPU version takes only 12.46 s. This study also used a video clip of table tennis player Lin Yun-Ju’s game on the YouTube website for AI gesture recognition (23.4 MB, 1 min 21 s), which took 7681.55 s for the CPU version and 62 s for the GPU version. This shows that, with the current technology, the GPU environment is considerably faster than the CPU version, and the improvement is apparent. The time required to use GPU is within the acceptable time range. Therefore, when using AI for computer vision processing, a high-end graphics card and an environment for GPU installation is required to fully utilize the GPU’s computing power and improve computing efficiency.

Table 1. Executive performance analysis.

Material	CPU Version	GPU Version
Photo processing speed (22 photos)	107.7 s	5.07 s
Video processing speed Video.avi (4 s) 1.33 MB	1313 s	12.46 s
Race Video Processing Speed Lin.avi (1 min 21 s) 23.4 MB	7681.6 s	62 s

Note: Lin.avi video source: <https://www.youtube.com/watch?v=nyIE5d0cRxI>. Video Title: Lin Yun-Ju The Silent Assassin (avi format) (access date: 1 September 2022)

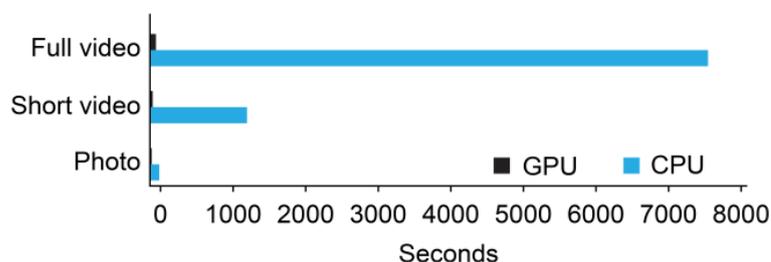


Figure 3. Execution performance analysis (unit: seconds).

4. Analysis and Discussion of Results

4.1. Comparative Analysis of Operational Performance

FPS (frames per second), or frame rate, denotes the number of consecutive images (frames) that are captured or displayed per second. A higher FPS enables smoother AI-based real-time recognition of a table tennis player’s posture in a video. The CPU computes a person’s pose in a video in real time. The result is shown in Figure 4. It shows that the CPU can only process 0.2 frames per second, so it is extremely slow and cannot smoothly complete the real-time pose recognition. In contrast, the GPU can compute 20.5 frames per second in real time (Figure 4), and this speed can be even faster with higher-end graphics cards. Therefore, with GPU computing, smooth computing can be achieved (Figure 5).



Figure 4. CPU version to compute the pose of a person in a video in real time (0.2 fps).



Figure 5. GPU version for real-time computation of the pose of a person in a video (up to 20.5 fps).

4.2. Analysis of Current Limitations in the Use of Posture Recognition

This study analyzed computer gesture recognition using the video highlights of the competition and found that there are several limitations in the current application. The limitations are as follows:

1. The human pose may not be recognized correctly in certain action angles and situations. For example, the pose of the player at the bottom of Figure 6 (yellow dotted box) is not correctly recognized.
2. Interference from off-court figures. For example, in Figures 7 and 8, the referee, off-court coaches, spectators, and others are also recognized for their movements and postures. Since posture recognition should only target the players, this causes unnecessary interference.
3. Other restrictions and interference. This study also identified the following limitations and interference:



Figure 6. Some postures may not be recognized correctly.



Figure 7. Interference of off-field persons (referee, coach).

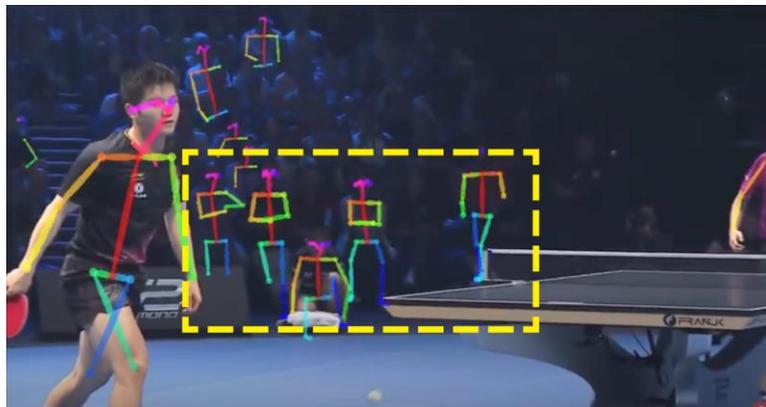


Figure 8. Interference of off-site persons (audience).

Estimation of footwork under the table: When the player's body is blocked by the table, sometimes the player's posture may not be recognized correctly; for example, the top half of the player in Figure 9 (yellow dotted box).

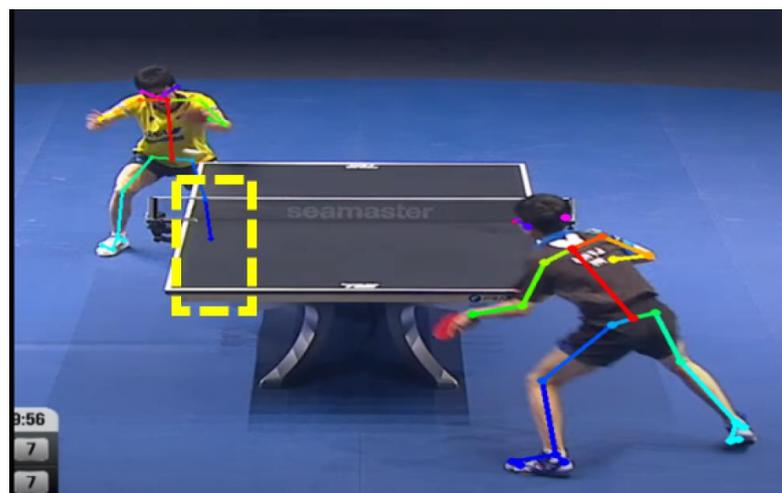


Figure 9. Estimation of foot posture under the table.

Interference in pose estimation due to video angle: Different camera angles in different games may cause the pose to not be interpreted smoothly. An example is shown in Figure 10 (yellow dashed box).

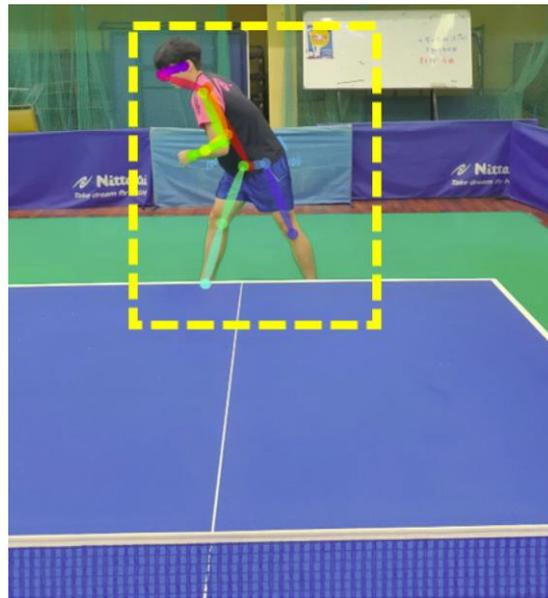


Figure 10. The pose caused by the angle of the film is not easily interpreted.

Interference in motion pose estimation recordings caused by replaying images: The video may be replayed in the highlight reel. Replays should not be included in pose recognition; only pose recognition during the competition is required. The solution is to remove the motion pose recognition from the highlight reel.

4.3. Practical Application to Table Tennis Strokes Analysis

This study used OpenPose to analyze the posture of table tennis strokes and was applied to a video of a game and actual photographs. The photos were taken of the National Tsing Hua University table tennis team, and the players were invited to demonstrate their strokes. The actual results of this study using AI in table tennis are shown in Figure 11 below. The left picture is the original history photo, and the right picture is the processed photo after the AI posture analysis. The human skeleton and joint points analyzed by AI are plotted on the photo to show the results of the AI analysis.



Figure 11. (A) Original photo. (B) Photo after AI analysis. The original photo and photo after AI analysis.

In this study, four major representative table tennis player's strokes were selected, namely, forehand loop (Figure 12A), backhand flick (Figure 12B), cut (Figure 13A), and chop (Figure 13B). This shows that the current AI posture estimation algorithm can analyze the skeleton and joints.

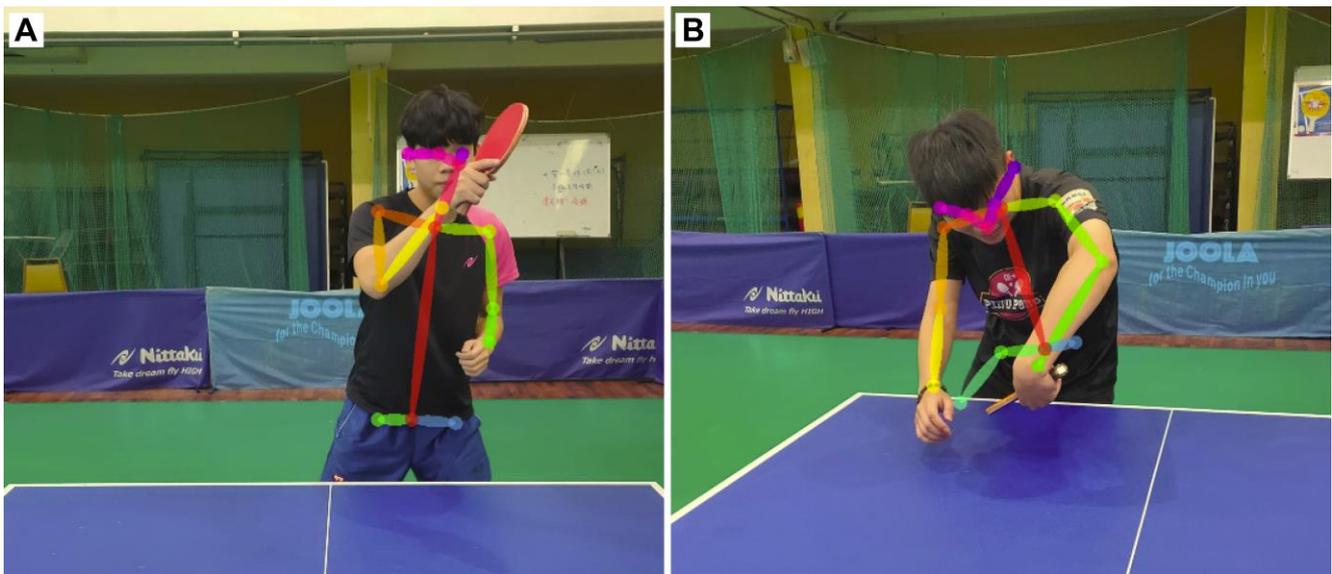


Figure 12. (A) Forehand loop (after analysis). (B) Backhand flick (after analysis). The forehand loop and backhand flick after analysis.

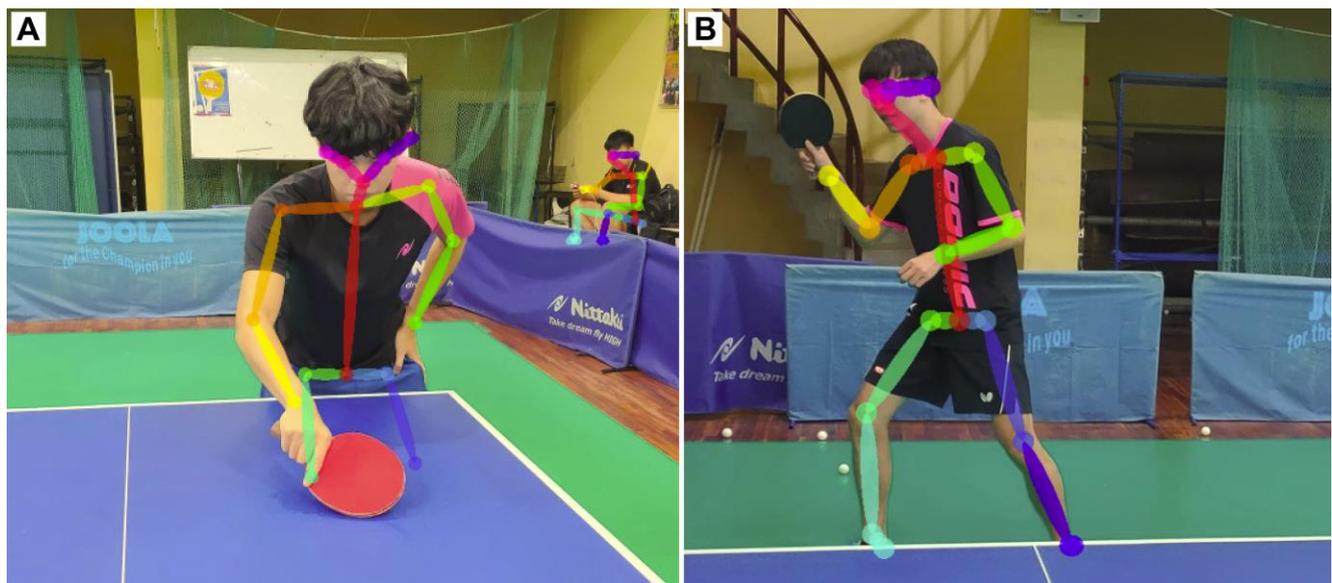


Figure 13. (A) Cut (after analysis). (B) Chop (after analysis). The cut and the chop after analysis.

4.4. Innovative Practices: Two-Stage AI Pose Recognition Process

This study proposes a prototype of an innovative two-stage AI pose recognition procedure to solve the current difficulties of AI applications for pose recognition. Initially, the OpenPose multiplayer pose estimation model architecture (e.g., Figure 14) or other multiplayer pose estimation algorithms can be applied to analyze the video and the key points of the pose in the video. For example, the multiperson pose estimation model architecture for OpenPose is shown in Figure 14. The input to the model is a color image of the dimensions $h \times w$. The output is an array of matrices that contain the confidence maps of key points and partial affinity heatmaps for each key point pair [27]. The specific key points as defined by OpenPose [27] can be found in Figure 15. The key points of the major joints of the human body as defined by OpenPose are illustrated in Figure 15A. Secondly, the pose key point data are saved. The pose key points (e.g., Figure 15B) are used as input variable X and provided to the deep learning network, such as the CNN, for pose

recognition. The pose is first defined by assigning different values to different poses, such as 1 for the forehand loop, 2 for the backhand flip, 3 for the cut, and 4 for the chop, as the output variable Y . The current pose can be a specific table tennis player’s motion (e.g., push, cut, loop, or chop).

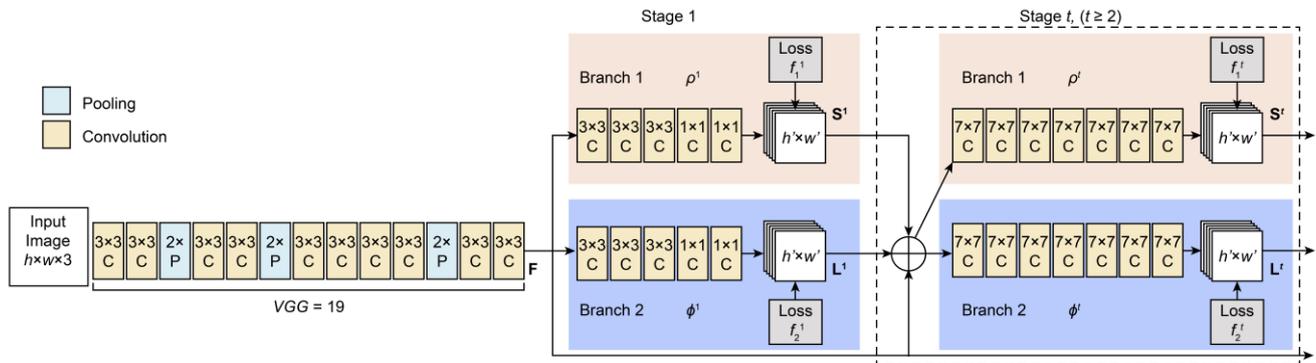


Figure 14. Multiperson pose estimation model architecture for OpenPose. Source: <https://learnopencv.com/multi-person-pose-estimation-in-opencv-using-openpose/> (access date: 1 October 2022).

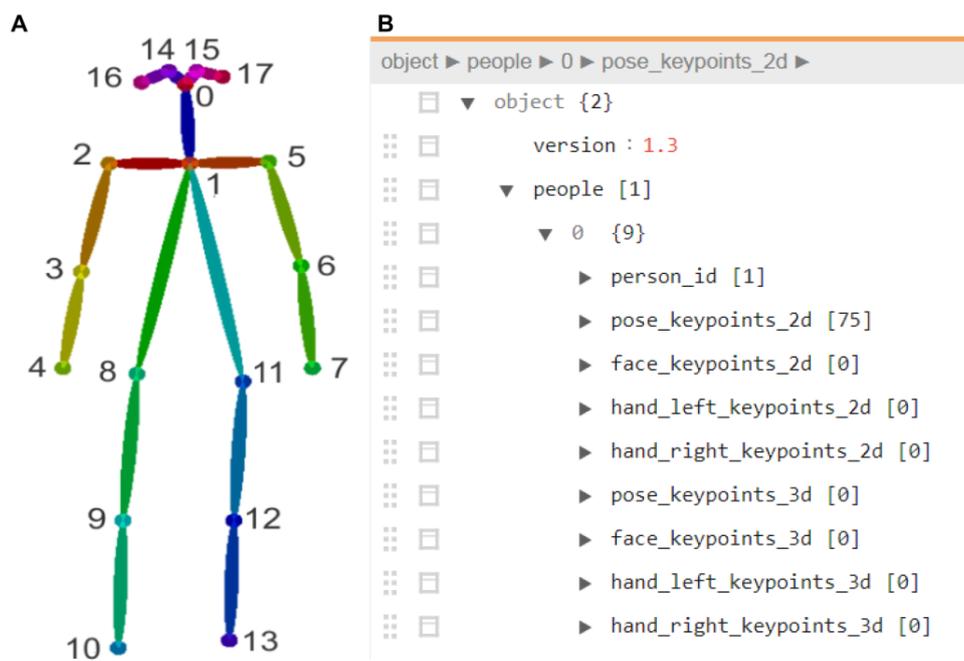


Figure 15. (A) Schematic diagram of the key points analyzed by the AI. (B) Posture key point data. Schematic diagram of the key points analyzed and the posture key point data by the AI.

In the above two-stage process, the first step is to record and analyze the gesture characteristics of a specific table tennis player’s pose so that it can be manually annotated and provided to deep network learning to build a prediction model. However, in table tennis, each action contains a series of movements, and AI video analysis will be performed for each gesture in each frame, thereby creating a problem of repeated recognition of technical movements by AI. This study used two of the table tennis player’s poses—backhand flip and backhand chop—as examples. The main characteristics of these two poses were recorded, and the following diagram shows the continuous poses of the backhand flip (Figure 16) and chop (Figure 17), showing that each pose is composed of several restorative poses (from the start pose to the end pose). Therefore, it is challenging to apply AI to gesture recognition. This study selected essential pose features for each type of table tennis player’s pose, such as the backhand flip in Figure 16B and backhand chop in Figure 17A,

then manually annotated them and provided them to the AI pose recognition model for training. The use in the subsequent video analysis solved the difficulties of AI application in table tennis pose recognition. The two-stage AI pose recognition process is shown in Figure 18.

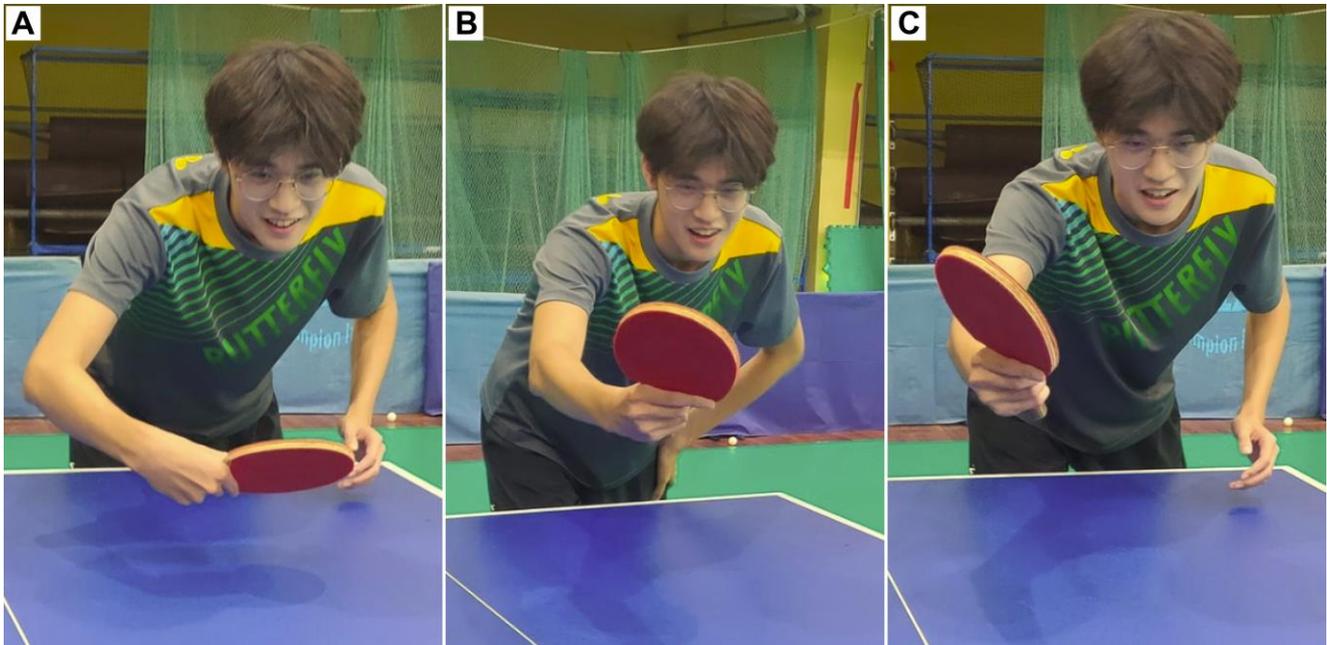


Figure 16. (A) Backhand flip action (start). (B) Backhand flip action. (C) Backhand flip action (end). The backhand flip action from the start to the end.

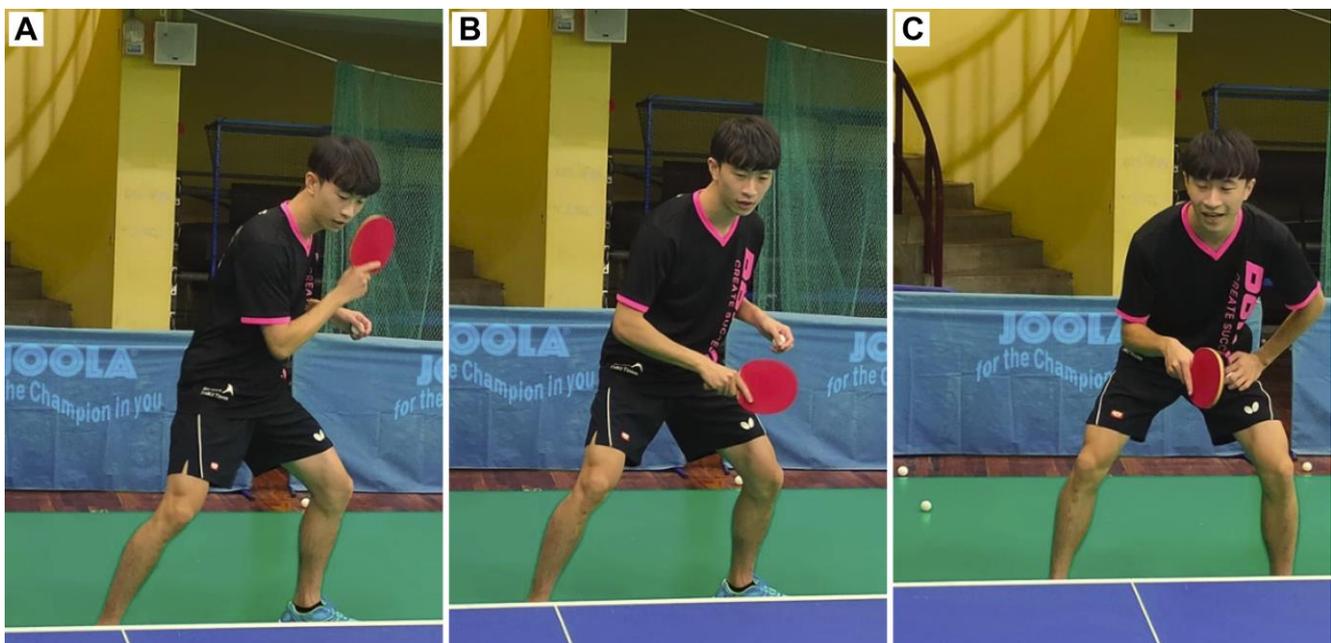


Figure 17. (A) Chop action (start). (B) Chop action. (C) Chop action (end). The chop action from the start to the end.

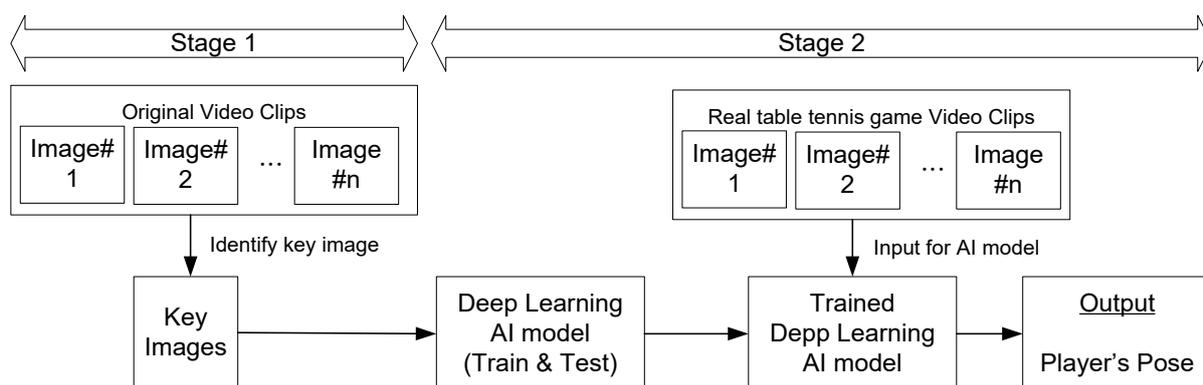


Figure 18. Two-stage AI pose recognition process.

4.5. Discussion

Taking the practical application of sports as an example, if AI gesture recognition technology is applied to the movement analysis of table tennis, this study suggests that it can be divided into two different aspects for analysis and exploration, namely “training” and “game technique analysis.” Because of the different emphasis on these two aspects of training and competition (Table 2), AI motion and posture recognition technologies should be designed and improved according to the actual needs of these aspects to achieve the expected results and objectives. Integrating AI technology practices and concepts in competitive sports aligns with Lin [28], who uses big data analysis and data management to provide “training focus,” “tactical application,” and “technical analysis” to improve athletic performance, thus providing specific directions and suggestions for developing technology-assisted contemporary competitive sports practices.

Table 2. The main analysis points of the AI application in the training and tactical analysis of the game of table tennis.

Application	Training	Tactical Analysis
Analysis highlights	<ul style="list-style-type: none"> - Posture analysis of a single technical movement - Types of major steps used - Range of anterior and posterior left and right shifts 	<ul style="list-style-type: none"> - Combination of two board techniques (tactics) - Basic stance information (side-to-side, middle) - Space information (highlight forehand, backhand, or balance) - Type of play (offensive, defensive) - Shift speed, the trajectory of technical interface

5. Conclusions

This study explored the current computer vision technologies in AI. It introduced one of these techniques, gesture recognition, and examined how it can be applied to motion video analysis. This study applied this technique to real-world video, analyzed its performance, and explored possible limitations, as described below.

Due to its rapid development, AI, particularly deep learning approaches such as CNNs and LSTM networks, will be extremely suitable for motion analysis applications. According to empirical studies, GPU performance is significantly better than CPU-only performance. Therefore, when using this technology, a high-end computing power host paired with a high-end graphics card should be built to achieve a smooth real-time computing effect. Currently, there are several limitations and disturbances in applying gesture recognition in real-world analysis that must be addressed. The motion can be corrected in the pre-production of the video or through different systems. The interference factors found in this

study are: (1) some poses cannot be estimated; (2) the interference of off-camera characters on pose recognition; (3) the estimation of foot pose under the table; (4) the interference of pose estimation due to the angle of the film; and (5) the interference of pose estimation during the recording of replay movements. In the second stage, the values of the key points were used as a depth network for pose recognition to overcome the current difficulties in applying AI to pose recognition.

The benefits of incorporating computer vision into AI technology and applying it to sports training and athletic practices are at least two-fold. First, this technology enhances the spectacle of the sport. The analysis of the player's biomechanics and power structure, and the process of each technique, can be presented to the audience in real time through posture analysis. Real-time data analysis can provide information for sports assistance, such as technical play analysis, which can improve the quality of ball commentary and significantly increase the visual effect and professionalism of sports viewing. Secondly, the big data stored after analysis can be used for subsequent training and tactical analysis of the game, which will improve the overall competitiveness of the game through the accumulation of time and information. Because OpenPose is currently a more mature algorithm for AI pose recognition, this paper focused only on it. In addition to OpenPose, some scholars have developed different gesture recognition algorithms, which could be compared in depth in future research.

Overall, AI pose recognition technology is still not widely used in sports on a global scale, but this study sees extensive room for actual technology development. AI systems should be designed with the expertise of sports experts to address the actual needs and ideals. In sports, AI should be combined and applied with other technologies across different fields to continue to create new research areas and directions.

Author Contributions: Conceptualization, C.-H.W., T.-C.W. and W.-B.L.; methodology, C.-H.W.; software, C.-H.W.; data curation, C.-H.W. and T.-C.W.; writing—original draft preparation, C.-H.W. and T.-C.W.; writing—review and editing, C.-H.W. and W.-B.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Ministry of Science and Technology Taiwan, grant number MOST 108-2511-H-142 -007-MY2 and MOST 110-2511-H-142 -008-MY2.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: The current study was supported by the Ministry of Science and Technology of Taiwan. (MOST 108-2511-H-142 -007 -MY2 and MOST 110-2511-H-142 -008-MY2) The authors thank Chien-Chih Wang, Te-Hao Chou, Yao-Chen Chiu, and Wen-Chieh Lo for providing photos on this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yu, C.; Huang, T.-Y.; Ma, H.-P. Motion Analysis of Football Kick Based on an IMU Sensor. *Sensors* **2022**, *22*, 6244. [CrossRef] [PubMed]
2. The Ministry of Science and Technology of Taiwan. 2019. Available online: <https://www.nstc.gov.tw/folksonomy/detail/177379c3-0061-43bb-ab33-c966df9edc73?l=ch> (accessed on 1 October 2022).
3. Liu, R.; Wang, Z.; Shi, X.; Zhao, H.; Qiu, S.; Li, J.; Yang, N. Table tennis stroke recognition based on body sensor network. In *Internet and Distributed Computing Systems*; Montella, R., Ciaramella, A., Fortino, G., Guerrieri, A., Liotta, A., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 1–10.
4. Balan, A.O.; Sigal, L.; Black, M.J.; Davis, J.E.; Haussecker, H.W. Detailed human shape and pose from images. In Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 17–22 June 2007; pp. 1–8. [CrossRef]

5. Bregler, C.; Malik, J. Tracking people with twists and exponential maps. In Proceedings of the 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No.98CB36231), Santa Barbara, CA, USA, 23–25 June 1998; pp. 8–15. [\[CrossRef\]](#)
6. Wei, X.; Chai, J. VideoMocap: Modeling physically realistic human motion from monocular video sequences. *ACM Trans. Graph.* **2010**, *29*, 42. [\[CrossRef\]](#)
7. Ye, G.; Liu, Y.; Hasler, N.; Ji, X.; Dai, Q.; Theobalt, C. Performance capture of interacting characters with handheld kinects. In *Computer Vision—ECCV 2012*; Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 828–841.
8. Huang, K.; Sui, T.; Wu, H. 3D human pose estimation with multi-scale graph convolution and hierarchical body pooling. *Multimed. Syst.* **2022**, *28*, 403–412. [\[CrossRef\]](#)
9. Kim, B.; Choo, Y.; Jeong, H.I.; Kim, C.I.; Shin, S.; Kim, J. Multi-resolution fusion network for human pose estimation in low-resolution images. *KSII Trans. Internet Inf. Syst.* **2022**, *16*, 2328–2344. [\[CrossRef\]](#)
10. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2980–2988. [\[CrossRef\]](#)
11. Fang, H.S.; Xie, S.; Tai, Y.W.; Lu, C. RMPE: Regional multi-person pose estimation. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2353–2362.
12. Li, J.; Wang, C.; Zhu, H.; Mao, Y.; Fang, H.S.; Lu, C. CrowdPose: Efficient crowded scenes pose estimation and a new benchmark. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 10855–10864.
13. Xiu, Y.; Li, J.; Wang, H.; Fang, Y.; Lu, C. Pose flow: Efficient online pose tracking. *arXiv* **2018**, arXiv:1802.00977.
14. AlphaPose. Alphapose Github. Available online: <https://github.com/MVIG-SJTU/AlphaPose> (accessed on 1 October 2022).
15. Cao, Z.; Hidalgo, G.; Simon, T.; Wei, S.E.; Sheikh, Y. OpenPose: Realtime multi-person 2D pose estimation using part affinity fields. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *43*, 172–186. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Simon, T.; Joo, H.; Matthews, I.; Sheikh, Y. Hand keypoint detection in single images using multiview bootstrapping. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 4645–4653.
17. Gao, Z.; Zhang, H.; Liu, A.A.; Xu, G.; Xue, Y. Human action recognition on depth dataset. *Neural Comput. Appl.* **2016**, *27*, 2047–2054. [\[CrossRef\]](#)
18. Xu, S.; Liang, L.; Ji, C. Gesture recognition for human–machine interaction in table tennis video based on deep semantic understanding. *Signal Process. Image Commun.* **2020**, *81*, 115688. [\[CrossRef\]](#)
19. Wu, E.; Perteneder, F.; Koike, H. Real-time table tennis forecasting system based on long short-term pose prediction network. In *SIGGRAPH Asia 2019 Posters*; Association for Computing Machinery: Brisbane, QLD, Australia, 2019.
20. Yadav, S.K.; Singh, A.; Gupta, A.; Raheja, J.L. Real-time Yoga recognition using deep learning. *Neural Comput. Appl.* **2019**, *31*, 9349–9361. [\[CrossRef\]](#)
21. Wu, E.; Koike, H. FuturePong: Real-time table tennis trajectory forecasting using pose prediction network. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*; Association for Computing Machinery: Honolulu, HI, USA, 2020; pp. 1–8.
22. Wang, Q.; Shi, L. Pose estimation based on PnP algorithm for the racket of table tennis robot. In Proceedings of the 2013 25th Chinese Control and Decision Conference (CCDC), Guiyang, China, 25–27 May 2013. [\[CrossRef\]](#)
23. Chen, G.; Xu, D.; Fang, Z.; Jiang, Z.; Tan, M. Visual Measurement of the Racket Trajectory in Spinning Ball Striking for Table Tennis Player. *IEEE Trans. Instrum. Meas.* **2013**, *62*, 2901–2911. [\[CrossRef\]](#)
24. Kun, Z.; ZaoJun, F.; JianRan, L.; Min, T. An adaptive way to detect the racket of the table tennis robot based on HSV and RGB. In Proceedings of the 2015 34th Chinese Control Conference (CCC), Hangzhou, China, 28–30 July 2015; pp. 5936–5940.
25. Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft COCO: Common objects in context. In *Computer Vision—ECCV 2014*; Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T., Eds.; Springer International Publishing: Cham, Switzerland, 2014; pp. 740–755.
26. Andriluka, M.; Pishchulin, L.; Gehler, P.; Schiele, B. 2D human pose estimation: New benchmark and state of the art analysis. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–24 June 2014; pp. 3686–3693.
27. Hidalgo, G.; Cao, Z.; Simon, T.; Wei, S.E.; Joo, H.; Sheikh, Y. CMU openpose Github. Available online: <https://github.com/CMU-Perceptual-Computing-Lab/openpose> (accessed on 1 October 2022).
28. Lin, W.B.; Yeh, S.W.; Yang, C.W. A Study of efficiency management for players and teams in CPBL from the viewpoint of data science. *Phys. Educ. J.* **2017**, *50*, 91–107. (In Chinese) [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.