Check for updates

Commercial vision sensors and AI-based pose estimation frameworks for markerless motion analysis in sports and exercises: a mini review

Saeid Edriss[1], Cristian Romagnoli[2]*, Lucio Caprioli[1], Vincenzo Bonaiuto[1], Elvira Padua[2] and Giuseppe Annino[1,3]

[1]Department of Industrial Engineering, Sports Engineering Laboratory, University of Rome Tor Vergata, Rome, Italy, [2]Department of Human Science and Promotion of Quality of Life, Human Performance, Sport Training, Health Education Laboratory, San Raffaele Open University, Rome, Italy, [3]Department of Medicine Systems, Human Performance Laboratory, Centre of Space Bio-Medicine, University of Rome Tor Vergata, Rome, Italy

Kinematic and biomechanical analysis in monitoring human movement to assess athletes' or patients' motor control behaviors. Traditional motion capture systems provide high accuracy but are expensive and complex for the public. Recent advancements in markerless systems using videos captured with commercial RGB, depth, and infrared cameras, such as Microsoft Kinect, StereoLabs ZED Camera, and Intel RealSense, enable the acquisition of high-quality videos for 2D and 3D kinematic analyses. On the other hand, open-source frameworks like OpenPose, MediaPipe, AlphaPose, and DensePose are the new generation of 2D or 3D mesh-based markerless motion tools that utilize standard cameras in motion analysis through real-time and offline pose estimation models in sports, clinical, and gaming applications. The review examined studies that focused on the validity and reliability of these technologies compared to gold-standard systems, specifically in sports and exercise applications. Additionally, it discusses the optimal setup and perspectives for achieving accurate results in these studies. The findings suggest that 2D systems offer economic and straightforward solutions, but they still face limitations in capturing out-of-plane movements and environmental factors. Merging vision sensors with built-in artificial intelligence and machine learning software to create 2D-to-3D pose estimation is highlighted as a promising method to address these challenges, supporting the broader adoption of markerless motion analysis in future kinematic and biomechanical research.

KEYWORDS

markerless motion capture, vision sensors, pose estimation, human movement analysis, sports biomechanics, kinematic analysis, artificial intelligence in sports, sports technology

1 Introduction

Understanding musculoskeletal movement is an important approach in monitoring human movements and motor behaviors (Schmidt and Lee, 2019). Kinematic and biomechanical analysis helps evaluate these parameters to measure and

capture the degrees of freedom of joints and limbs independently, diagnose injuries, and analyze the performance of athletes and patients (Dao and Tho, 2014). Numerous studies using motion capture systems (MoCap), wearable sensors, and technologies have become essential tools in laboratories and field tests for investigating specific biomechanical or kinematic issues. While these technologies serve as the gold standard for monitoring individuals in 3D simulations, their high costs and the setup complexity present significant limitations (Edriss et al., 2024a; Patrizi et al., 2016).

In the recent decade, body simulation in 2D and 3D has undergone rapid evolution, driven by artificial intelligence (AI), and has combined with the sports, medicine, animation, and gaming industries (Khan et al., 2024; Dai and Li, 2024). Kinematic monitoring and analysis represent a contemporary method for simulating human body motion, especially in 2D (Seo et al., 2019). Numerous studies have examined the validity of various devices and software designed to monitor specific motions. While 3D data offers more information than 2D data, the choice largely depends on the intended application. 2D information can be suitable for particular applications, especially concerning ease of use and cost. However, 2D simulation has limitations like restricted access to certain angles and the capturing of movements from specific perspectives. Despite these limitations, these systems provide significant advantages, including low costs for registering videos with one or a few cameras, quick and easy to setup and operate (Caprioli et al., 2024). Kinematics is valuable for characterizing changes in specific joint angles to monitor how motion performs and loads in land and water environments (Cronin et al., 2019). Markerless systems use RGB, depth, or infrared cameras, with Microsoft launching the successful commercial system supported by a pose estimation (PE) device in 2010 (Holte et al., 2012). Now, vision sensors, including depth RGB-D and stereo cameras, are being used. Depth RGB cameras capture actual 3D spatial information through active sensing methods such as infrared projection, stereo vision, or Time-of-Flight technology, and stereo cameras use two RGB lenses to estimate depth through disparity. Later, in 2017, OpenPose, as one of the first PE frameworks, was released to extract body pose from standard 2D videos (Qiao et al., 2017).

This review explores the development of commercial technologies in real-time and offline markerless MoCap, starting with earlier systems using vision sensors that enable direct 3D motion tracking, followed by newer PE frameworks that analyze standard videos with AI, emphasizing the importance of their integration. Each methodology has its exclusive capabilities, and familiarity with each tool is beneficial; however, recent research has increasingly combined both to enhance functionality and usability. Conversely, PE frameworks estimate joint positions from 2D RGB videos using machine learning (ML) algorithms. Consequently, studies since 2010 on kinematic hardware and software for 2D and 3D analysis have focused on their validity and reliability compared to gold-standard systems. The selected experimental validation and technical reports focused on body motion in English are sourced from IEEE, PubMed, and Google Scholar since 2012, with PE tools from 2018 onward. The following keywords and combinations were used: "markerless-MoCap", "vision-sensors", "pose-estimation", "MediaPipe", "OpenPose", "RealSense", "Kinect", "ZED-camera", and "sports-tracking-analysis." The review also highlights the chosen gold-standards for these systems' validation, optimal perspectives,

and setups to achieve the best results in sports applications, emphasizing the practical benefits and uses of these technologies in human activity research and athletic performance, and emphasizing the need for future vision sensor generation to integrate AI-based PE frameworks for broader usability.

2 Hardware (vision sensors)

2.1 Kinect

Microsoft designed and released Kinect K1 (in 2010) and K2 (in 2014), which consist of one IR emitter, one IR camera, and one RGB camera, to gain depth and color images. To strongly enter the gaming industry and provide human connections as a part of Xbox consoles, targeted the companies that designed the Kinect, which has lately been used as a markerless, affordable, and portable MoCap sensor (Bilesan et al., 2019). Kinect can perform PE using its built-in depth sensor and the associated software development kits (SDKs). Many studies explored the accuracy of the two Kinects in real-time skeletal tracking during human movements (Ganea et al., 2014; Kurillo et al., 2022). However, according to (Wang et al., 2015), regarding MoCap selection as a gold standard, the K2 generation is beneficial in terms of joint estimation accuracy and more robust in handling track occlusion and body rotation. In another study, a Vicon MoCap was compared with Kinect in gait analysis, as participants walked and jogged on a treadmill. The results show the validity of the Kinect in the sagittal plane for stride timing in kinematic analysis (Pfister et al., 2014).

Depending on the coaches' purpose, estimated pose through the Kinect somatosensory camera can be sufficient in physical activity recognition, such as qigong movement detection (Fan et al., 2024). Or by requiring more complex analysis processes, their implementation with other codes or devices, such as 3DSMAX, can invent novel methods for monitoring the sports training seasons (Shi, 2014). Additionally, regarding simulating and estimating the body joints and limbs, the skeletal body structure, Asteriadis et al. used Kinect in body estimation in 3D by minimizing the noise (Asteriadis et al., 2013).

The Kinect device alone lacks accuracy in defining orthogonal axes unrelated to joint centers, limiting its ability to measure internal or external angular rotations and displacement. Adding depth sensors, video cameras, or multiple devices to make a MoCap-based Kinect can improve the tracking process and angular measurement accuracy (Clark et al., 2012; Napoli et al., 2017).

2.2 StereoLabs ZED camera

The ZED Camera by StereoLabs, released in 2015, was followed by the next-generation, including ZED2 (2019) and ZED2i (2021), which were enhanced with depth sensing and AI features (Sarıalioğlu et al., 2024). ZED cameras are equipped with two high-resolution lenses to capture in-depth information, making them suitable commercial vision hardware for integrating open pose libraries in kinematic and biomechanical studies (Avogaro et al., 2023). A comparative analysis of the ZED2 camera with other commercial vision systems, including Azure Kinect and RealSense

D455, utilizing their respective SDKs, shows higher performance in human pose tracking at distances exceeding 3 m. This advantage aids PE in larger spaces or dynamic environments, such as field tests (Ramasubramanian et al., 2024). Another suggests that Nuitrack and MediaPipe software integrate with ZED2i to perform better in upper and lower extremity features for pose tracking (Aharony et al., 2024). Combining ZEDmini with a Virtual Reality (VR) headset and an augmented reality (AR) prototype resulted in a device capable of recognizing body formation and sports equipment such as a tennis racket and basketball hoop. This novelty aids players with low vision to enhance their ability to identify the players and objects (Lee et al., 2024). The advantages of ZED cameras as hardware for exercise monitoring, assessing gait impairment in Neurodegenerative disorders, and joint angle measurement include their ease of use, affordability, and ability to deliver reliable results in both 2D and 3D kinematic analyses of exercises, such as squats (Aliprandi et al., 2023; Zanela et al., 2022).

2.3 Intel RealSense

Intel RealSense cameras combine high-resolution RGB imaging with depth perception, offering models like the D400 series for general depth sensing, the L515 for precise LiDAR tasks, and the T265 for motion tracking with simultaneous localization and mapping (SLAM) capabilities (Moghari et al., 2024; Tsykunov et al., 2020). By capturing depth maps and RGB frames, they integrate seamlessly with pose-estimation software framework tools, mapping 2D keypoints to real-world coordinates. Their lightweight design, robust SDK, and real-time performance make them versatile for indoor and outdoor applications, outperforming infrared-based systems like Kinect in varied lighting conditions (Maddipatla et al., 2023).

An Intel RealSense camera is used in head motion tracking using custom software while participants walk indoors and outdoors. They walked in both slow and fast modes, and their head motion tracking was compared with that of a perambulator containing inertial measurement units (IMUs) and a distance counter, serving as the gold standard. The accurate results of this gait study showed that while the participants walked slowly (Hausamann et al., 2021). Another study compared the Intel RealSense camera to the Kinect in slow walking gait analysis. The results indicated that while the Intel RealSense D435 is valid for measuring certain spatiotemporal variables, the Kinect proved more reliable for collecting skeletal data due to its more robust RGB-D camera (Mejia-Trujillo et al., 2019). MediaPipe landmarks and two Intel RealSense cameras were used to estimate angular parameters in five exercises, including squats, knee, ankle, hip extension, and shoulder elevation, compared with OptiTrack as the reference. The results showed an error of less than 20 degrees, demonstrating an adequate estimation quality. Still, further improvements in both software and hardware can enhance the system's accuracy (Pilla-Barroso et al., 2024). These generations of cameras are increasingly being used in sports performance analysis through integration with AI-based markerless MoCap software. Intel RealSense depth cameras serve as sufficient registration tools in markerless PE, especially for closer range (under 3 m) (Ramasubramanian et al., 2024; Jacobsson et al., 2023).

3 Software (pose estimation models)

3.1 OpenPose

OpenPose is an open-source library developed by Carnegie Mellon University for 2D PE of the skeletal structure by detecting 25 keypoints on the body (Hidalgo et al., 2019). A real-time detection system may focus on specific angles between joints, the center of mass, or a point of displacement (Cronin et al., 2024; Needham et al., 2021). To analyze the running performance and timing, OpenPose evaluated it with the Coco dataset regarding human PE to recognize hurdles athletes' wrists, ankles, hips, and knees (Jafarzadeh et al., 2021; Sharma et al., 2022). The Coco dataset recognizes 18 pre-trained landmarks with x, y, and v values, where x and y indicate coordinates, and v shows the visibility of the landmarks (Duan et al., 2023).

PE relies on video quality and frame rates, meaning that weather conditions and lighting impact the accuracy of ML systems. A study evaluating the AI accuracy of OpenPose as a PE model under various environmental and lighting conditions, while participants performed stretching exercises, found that the accuracy rates depend on the environment and setup context (Song and Chen, 2024).

In a study using 3D motion analysis as the reference standard, an OpenPose-based motion analysis was used to compare knee valgus angular data during the drop vertical jump test. The strong correlation among these data indicated a sufficient ML accuracy level of the PE models in measuring angular movements (Ino et al., 2024). In baseball, analyzing a hitter's limb angles and hip distance aids in swing performance (Li et al., 2021). AI-driven gesture recognition supports biomechanical analysis in training techniques and game strategy evaluation. This approach was a key objective in a study analyzing tennis player performance through skeletal PE (Wu et al., 2023).

Furthermore, PE models' integration with wearable sensors or MoCap can facilitate more complex processes through Taekwondo kinematic analysis through a PE model, 8 body joint angles in players' upper and lower body registered by Contemplas cameras, and the Vicon system. OpenPose was used to process recordings from Contemplas cameras to extract 2D human skeleton data, then triangulated to obtain 3D joint angles to compare world-class and master-class players' kicking performance (Fukushima et al., 2024). Plus, optimal and kinetic analysis of fouling and shooting in basketball player performance was conducted using video-based (2D) methods with a lightweight deep learning (DL) architecture, where OpenPose was used as the PE model (Xu, 2024). This open-source framework is being used in athletes' 3D estimating with integration with sensors or the Internet of Things (IoT) in various sports such as gymnastics (Ren et al., 2025).

3.2 MediaPipe

MediaPipe is an open-source framework developed by Google for 2D human PE (Pham et al., 2022). It enhances AI applications by detecting keypoints on the body, face, and hands in images or live video streams to offer gesture recognition, motion analysis, and interactive systems (Saini, 2024). MediaPipe is a broader framework
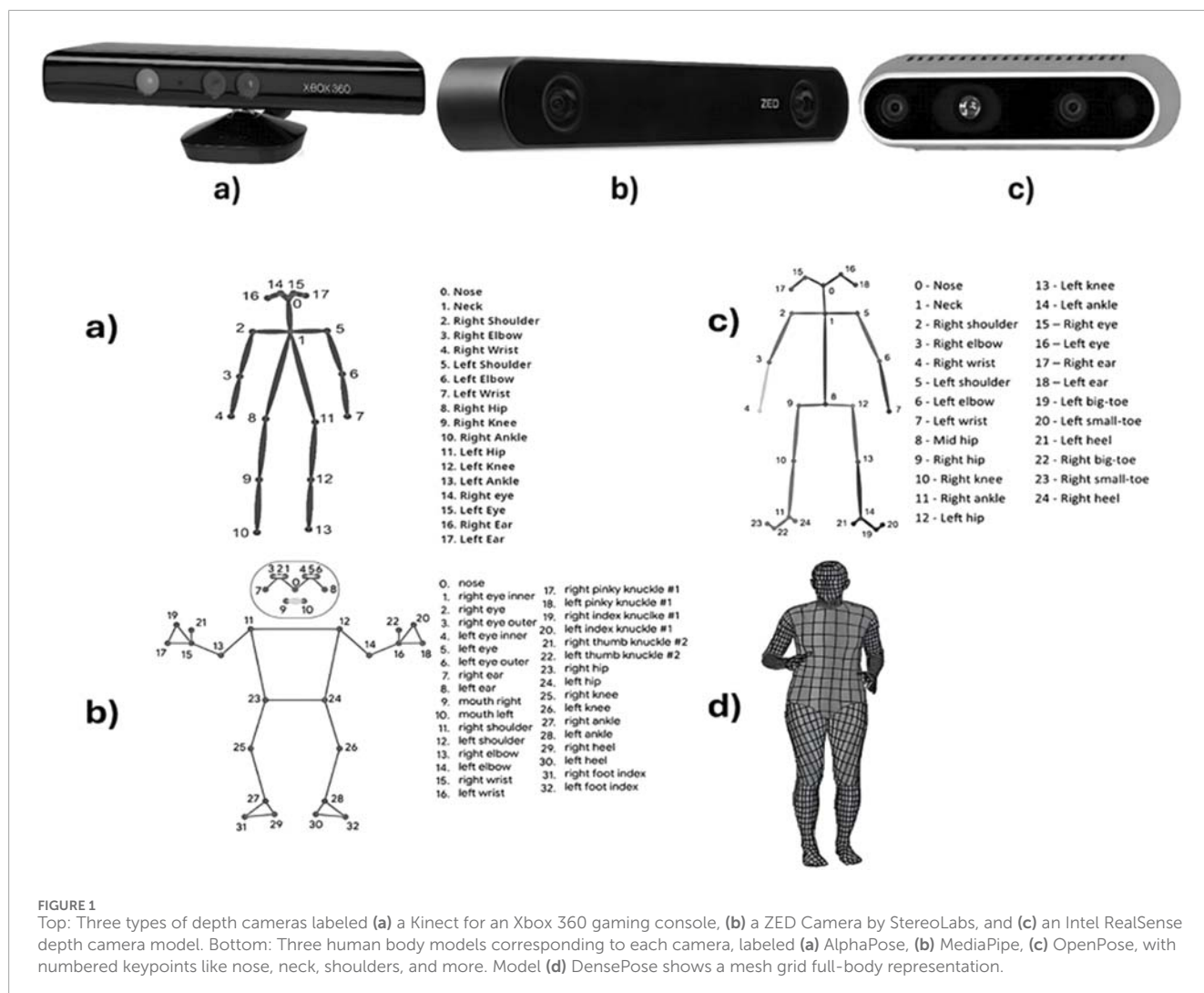
FIGURE 1
Top: Three types of depth cameras labeled **(a)** a Kinect for an Xbox 360 gaming console, **(b)** a ZED Camera by StereoLabs, and **(c)** an Intel RealSense depth camera model. Bottom: Three human body models corresponding to each camera, labeled **(a)** AlphaPose, **(b)** MediaPipe, **(c)** OpenPose, with numbered keypoints like nose, neck, shoulders, and more. Model **(d)** DensePose shows a mesh grid full-body representation.

that includes BlazePose and other ML solutions, where BlazePose specializes in detecting 33 keypoints (Wang, 2024).

An important finding shows that when assessing PE quality using MediaPipe from a frontal and lateral view, an error margin of less than 25% of the range of motion is optimal, with specific tolerances: under 20 degrees for movements greater than 90°, around 10 degrees for movements up to 40°, and below 10° for static angles (Pilla-Barroso et al., 2024).

The validity of MediaPipe BlazePose in measuring joint angles compared with the IMU-based MoCap in 2D image analysis, and the results showed that the difference in accuracy is within 10% (Pau et al., 2021). In addition, an ML tool using MediaPipe simulated kinematic points, assuming shoulder, knee, hip, and ankle joints for artistic swimming analysis. Regarding the 2D kinematic simulation frames, the leg deviation angles were measured during execution and body position in preparation to aid officiating systems and performance monitoring (Edriss et al., 2024b).

In a video performance analysis using an ML tool with MediaPipe, the knee angles of advanced and beginner pickleball players were measured during dink shots. This kinematic analysis

assessed the differences in knee bending and body flexion between the two groups, suggesting an appropriate positioning formation and preparation for dinking (Edriss et al., 2025).

An application using MediaPipe, compared with Kinect, was used to measure knee angles for athletes under a motion test. The results showed the higher capacity of the MediaPipe framework in analyzing ACL injury risk by calculating lower limb kinematics through a smartphone (Babouras et al., 2024). However, a limitation of the MediaPipe source is the lack of ability to recognize multiple people in a frame (Dong and Yan, 2024).

3.3 AlphaPose

AlphaPose is an open-source tool that provides PE by detecting 25 keypoints. It was developed by the Chinese University of Hong Kong (Fang et al., 2023). Like OpenPose, AlphaPose provides a high keypoint detection rate and is highly interchangeable with ML models for biomechanical PE parameters like ground reaction forces (Mundt et al., 2023).

TABLE 1 Provides a comparative overview of depth cameras and pose estimation frameworks, reviewing their typical error margins, envir onmental constraints, reliability, and suitability for recording players' performance, with a focus on body limbs, setup complexity, and suitability for recording.

System	Type	Error range	Environment suitability	Depth range (m)	2D/3d output	Reliability (limb Focus)	Setup complexity	Key strengths and limitations
Kinect v2	Depth Camera	~10°–20° (Pfister et al., 2014)	Indoor only	0.5–4.5 (Kurillo et al., 2022)	3D	Both (higher in the sagittal plane)	Low	Built-in skeletal tracking
ZED 2i	Stereo Camera	~5°–15° (Aharony et al., 2024)	Indoor/Outdoor, >3 m	0.2–20 (Aharony et al., 2024)	3D	Lower limb (wide FOV applications)	Medium-High	Higher performance for wide views
RealSense	Depth Camera	~10°–20° (Pilla-Barroso et al., 2024)	Indoor/Outdoor	0.2–10 (Aharony et al., 2024)	2D/3D	Lower limb (slow gait, squat)	Medium	Valid for close-range tracking
OpenPose	Software	~5°–15° (Ino et al., 2024)	Indoor/Outdoor	Camera-limited	2D (pseudo-3D)	Upper limb (fine movement detection)	Medium	Good multi-person support
MediaPipe	Software	~<10° (Pilla-Barroso et al., 2024)	Indoor/Outdoor	Camera-limited	2D	Lower limb (squats, gait)	Low	Fastest runtime, low memory use, suitable for mobile apps
AlphaPose	Software	~15°–25° (Hulleck et al., 2023)	Indoor/Outdoor	Camera-limited	2D	Upper limb (combat/sport gestures)	Medium	Higher RMSE in some studies
DensePose	Software	~10°–20° (Suzuki et al., 2024)	Indoor preferred	Camera-limited	2D/3D mesh	Both (AR/VR mesh mapping)	High	Ideal for multi-person and VR/AR

The numbers refer to the references mentioned in the text.

This framework is used in various ways, including gait and player kinematic analyses. For example, in gait analysis, participants walked at different speeds, and their estimated poses were captured using AlphaPose (Wei et al., 2023). Additionally, AlphaPose captured accurate skeleton PE of badminton players to classify their movements. These results achieved an 80% prediction accuracy, effectively analyzing the sequential nature of badminton actions (Liang and Nyamasvisva, 2023).

Several researchers are working on PE framework development through the DL tool, such as YOLO (Yang et al., 2024; Zhao et al., 2024). Although the validity of AlphaPose is sufficient, its accuracy compared to the data of other frameworks is considered. For example, a study compared MediaPipe BlazePose and AlphaPose with Vicon MoCap as the reference standard in gait joint kinematics analysis. This study emphasized that MediaPipe shows a lower mean square error (RMSE), indicating more accurate joint kinematic measurements. However, AlphaPose showed greater variability with a higher range of motion and RMSE (Hulleck et al., 2023).

3.4 DensePose

DensePose is a PE model developed by Facebook AI Research that maps human pixels in an image to the body's 3D surface (Gu et al., 2022; Lovanshi and Tiwari, 2022; Mahajan et al., 2023). The advantages of DensePose, compared to other frameworks, are its perfect ability for multi-person recognition and its capability to estimate the body by keypoints or mesh (Güler et al., 2018; Guler and Kokkinos, 2019; Zhang et al., 2021). DensePose is effective for applications in augmented reality, virtual try-on, and immersive human-body interaction scenarios (Islam et al., 2024; Zhu and Song, 2021). Athletes' biomechanical and kinematic analyses can be conducted by designing a 2D PE from videos using DensePose, which is pre-trained on the COCO keypoint dataset. For monocular 3D PE, MotionAGFormer, pre-trained on the Human3.6M dataset, can be utilized. DensePose requires more technical skills to set up due to its powerful GPU resources and custom code for mesh mapping. Studies indicate that joint angular measurements obtained through DensePose also provide reliable data (Hellstén et al., 2022; Suzuki et al., 2024).

Figure 1i illustrates the shape of the vision sensors, and Figure 1ii demonstrates the human skeletal structure as detected by the PE models.

Table 1 shows a brief comparison between hardware and software to achieve a clearer idea of their strength and pros.

4 Discussion

Many researchers believe future sports motion studies will rely on markerless motion analysis systems (Ismail et al., 2016). Therefore, much research highlighted the accuracy and reliability of the body pose simulation device and software in the kinematic analysis procedures (Clark et al., 2012), particularly for motions within a single plane. Mainly, the articles validated body simulation tools by gold-standard IMUs or MoCaps such as Vicon or OptiTrack, comparing quantitative methods like RMSE in joint angles and joint position errors. Some studies applied qualitative validation, such as expert visual assessment or usability scoring. However, high limitations exist because 2D PE restricts their ability to capture out-of-plane movements, such as joint flexion or rotation, as they lack depth information (Pfister et al., 2014; Annino et al., 2023). Additionally, proper setup requires careful camera or device setup procedures to reduce their limitation (Yang and Park, 2024). Plus, PE accuracy depends on video quality, frame rates, and environmental conditions (outdoor, indoor, or natural or artificial lights), with studies showing varying OpenPose performance across different scenarios (Song and Chen, 2024).

Despite these drawbacks, 2D PE methods are beneficial due to affordability, ease of use, and relatively simple setup compared to 3D systems such as Vicon (Menychtas et al., 2023). Regarding the investigation aims, the camera and device can be set preferably if the motions are in the same plane, and it is beneficial to use the 2D PE methods to assess clinical or physical kinematic or biomechanical aspects (Gozlan et al., 2024; Hamill et al., 2012; Winter, 2009).

Regarding the articles, each hardware and software tool has limitations. Kinect and RealSense are constrained by lighting conditions, occlusions, and limited field of view, especially in dynamic environments. ZED cameras may require more computational resources. OpenPose and MediaPipe rely heavily on camera angles and lighting and may struggle with occlusions or out-of-plane movements. AlphaPose exhibits higher RMSE and variability in certain tasks, while DensePose requires significant computational resources. While these hardware and software tools can function independently, leveraging RGB video and PE software (for example, Kinect with OpenPose or ZED with MediaPipe) offers a simple way to obtain 3D body landmarks and points of interest for the upper and lower body analysis, with high accuracy (Ramasubramanian et al., 2024; Liu and Chang, 2022).

5 Conclusion

PE software has strengths compared to others; thus, selecting the best framework is not straightforward. For instance, some research illustrated MediaPipe's advantage over OpenPose in offering lower deviation in keypoints detection, especially in feet and wrists, along with faster runtime (Latyshev et al., 2024). On the other hand, another study emphasizes OpenPose's superiority over MediaPipe BlazePose in accurately detecting clinically relevant keypoints closer to anatomical joint centers (Mroz et al., 2021). In conclusion, while both models show promise, further improvements are necessary, and future studies will increasingly rely on AI-based markerless PE tools.

These methods provide a cost-effective and practical solution for studies focusing on movements confined to a single plane. They are particularly beneficial for assessing clinical, physical, and kinematic or biomechanical parameters when high precision across multiple planes is not required (Stenum et al., 2021). Moreover, AI-based 2D-to-3D PE techniques are a fast-developing method to reduce these limitations, further expanding the potential applications of 2D systems (Fortini et al., 2023). We suggest that future studies aim to evaluate protocols and integrate vision sensors with PE

framework software to bridge the gap in the development and validation of the 2D-to-3D PE device in real-world sports, thereby accessing biomedical data such as joint angles and providing real-time feedback.

Author contributions

SE: Writing – original draft, Conceptualization, Writing – review and editing. CR: Supervision, Writing – original draft, Writing – review and editing, Visualization. LC: Writing – review and editing. VB: Writing – review and editing, Supervision. EP: Writing – review and editing, Supervision. GA: Supervision, Writing – review and editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aharony, N., Meshurer, A., Krakovski, M., Parmet, Y., Melzer, I., and Edan, Y. (2024). Comparative analysis of cameras and software tools for skeleton tracking. *IEEE Sensors J.* 24 (20), 32302–32312. doi:10.1109/JSEN.2024.3450754

Aliprandi, P., Girardi, L., Martinelli, G., De Natale, F., Bisagno, N., and Conci, N. (2023). A case study for the automatic supervision of body-weight exercises: the squat. In: 2023 IEEE international workshop on sport, technology and research (STAR); 2024 July 8-10: IEEE. p. 13–16. doi:10.1109/STAR58331.2023.10302651

Annino, G., Bonaiuto, V., Campoli, F., Caprioli, L., Edriss, S., Emilio, P., et al. (2023). Assessing sports performances using an artificial intelligence-driven system. In: 2023 IEEE international workshop on sport, technology and research (STAR); 2023 September 14-16: IEEE. p. 98–103.

Asteriadis, S., Chatzitofis, A., Zarpalas, D., Alexiadis, D. S., and Daras, P. (2013). Estimating human motion from multiple kinect sensors. In: Proceedings of the 6th international conference on computer vision/computer graphics collaboration techniques and applications, in MIRAGE '13; 1998 January 7-7. New York, NY, USA: Association for Computing Machinery. p. 1–6. doi:10.1145/2466715.2466727

Avogaro, A., Cunico, F., Rosenhahn, B., and Setti, F. (2023). Markerless human pose estimation for biomedical applications: a survey. *Front. Comput. Sci.* 5. doi:10.3389/fcomp.2023.1153160

Babouras, A., Abdelnour, P., Fevens, T., and Martineau, P. A. (2024). Comparing novel smartphone pose estimation frameworks with the kinect V2 for knee tracking during athletic stress tests. *Int. J. CARS* 19 (7), 1321–1328. doi:10.1007/s11548-024-03156-5

Bilesan, A., Behzadipour, S., Tsujita, T., Komizunai, S., and Konno, A. (2019). Markerless human motion tracking using microsoft kinect SDK and inverse kinematics. In: 2019 12th Asian control conference (ASCC); 2019 June 09-12; Kitakyushu, Japan: IEEE. p. 504–509. Available online at: https://ieeexplore.ieee.org/document/8765061/?arnumber=8765061 (Accessed December 23, 2024).

Caprioli, L., Campoli, F., Edriss, S., Padua, E., Najlaoui, A., Giuseppe, A., et al. (2024). Impact distance detection in tennis forehand by an inertial system. In: Proceedings of the 12th international conference on sport sciences research and technology support. Porto, Portugal: SciTePress. p. 21–22. Available online at: https://www.scitepress.org/Papers/2024/130731/130731.pdf (Accessed January 17, 2025).

Clark, R. A., Pua, Y. H., Fortin, K., Ritchie, C., Webster, K. E., Denehy, L., et al. (2012). Validity of the microsoft kinect for assessment of postural control. *Gait. Posture* 36 (3), 372–377. doi:10.1016/j.gaitpost.2012.03.033

Cronin, N. J., Rantalainen, T., Ahtiainen, J. P., Hynynen, E., and Waller, B. (2019). Markerless 2D kinematic analysis of underwater running: a deep learning approach. *J. Biomechanics* 87, 75–82. doi:10.1016/j.jbiomech.2019.02.021

Cronin, N. J., Walker, J., Tucker, C. B., Nicholson, G., Cooke, M., Merlino, S., et al. (2024). Feasibility of OpenPose markerless motion analysis in a real athletics competition. *Front. Sports Act. Living* 5 (Jan). doi:10.3389/fspor.2023.1298003

Dai, F., and Li, Z. (2024). Research on 2D animation simulation based on artificial intelligence and biomechanical modeling. *EAI Endorsed Trans. Perv. Health Tech.* 10. doi:10.4108/eetpht.10.5907

Dao, T. T., and Tho, M.-C. H. B. (2014). *Biomechanics of the musculoskeletal system: modeling of data uncertainty and knowledge*. Hoboken, New Jersey: John Wiley and Sons.

Dong, K., and Yan, W. Q. (2024). Player performance analysis in table tennis through human action recognition. *Computers* 13 (12), 332. doi:10.3390/computers13120332

Duan, C., Hu, B., Liu, W., and Song, J. (2023). Motion capture for sporting events based on graph convolutional neural networks and single target pose estimation algorithms. *Appl. Sci.* 13 (13), 7611. doi:10.3390/app13137611

Edriss, S., Caprioli, L., Campoli, F., Manzi, V., Padua, E., Bonaiuto, V., et al. (2024b). Advancing artistic swimming officiating and performance assessment: a computer vision study using MediaPipe. *Int. J. Comput. Sci. Sport* 23 (2), 35–47. doi:10.2478/ijcss-2024-0010

Edriss, S., Romagnoli, C., Caprioli, L., Zanela, A., Panichi, E., Campoli, F., et al. (2024a). The role of emergent technologies in the dynamic and kinematic assessment of human movement in sport and clinical applications. *Appl. Sci.* 14 (3), 1012. doi:10.3390/app14031012

Edriss, S., Romagnoli, C., Maurizi, M., Caprioli, L., Bonaiuto, V., and Annino, G. (2025). Pose estimation for pickleball players' kinematic analysis through MediaPipe-based deep learning: a pilot study. *J. Sports Sci.*, 1–11. doi:10.1080/02640414.2025.2524283

Fan, Z., and Sun, K. (2024). Kinect based recognition and detection of fitness qigong movements. *ICIC Int. 学会* 06. doi:10.24507/ijicic.20.06.1837

Fang, H.-S., Li, J., Tang, H., Xu, C., Zhu, H., Xiu, Y., et al. (2023). AlphaPose: whole-body regional multi-person pose estimation and tracking in real-time. *IEEE Trans. Pattern Anal. Mach. Intell.* 45 (6), 7157–7173. doi:10.1109/TPAMI.2022.3222784

Fortini, L., Leonori, M., Gandarias, J. M., de Momi, E., and Ajoudani, A. (2023) Markerless 3D human pose tracking through multiple cameras and AI: enabling high accuracy, robustness, and real-time performance. *arXiv Pre print*. arXiv:2303.18119. doi:10.48550/arXiv.2303.18119

Fukushima, T., Haggenmueller, K., and Lames, M. (2024). Validity of OpenPose key point recognition and performance analysis in taekwondo. In: H. Zhang, M. Lames, A. Baca, and Y. Wu, editors. Proceedings of the 14th international symposium on computer science in sport (IACSS 2023); Singapore: Springer Nature. p. 68–76. doi:10.1007/978-981-97-2898-5_8

Ganea, D., Mereuta, E., and Mereuta, C. (2014). Human body kinematics and the kinect sensor. *Appl. Mech. Mater.* 555, 707–712. doi:10.4028/www.scientific.net/AMM.555.707

Gozlan, Y., Falisse, A., Uhlrich, S., Gatti, A., Black, M., and Chaudhari, A. (2024). OpenCapBench: a benchmark to bridge pose estimation and biomechanics. *arXiv preprint*. arXiv:2406.09788.

Gu, D., Yun, Y., Tuan, T. T., and Ahn, H. (2022). Dense-Pose2SMPL: 3D human body shape estimation from a single and multiple images and its performance study. *IEEE Access* 10, 75859–75871. doi:10.1109/access.2022.3191644

Guler, R. A., and Kokkinos, I. (2019). Holopose: Holistic 3d human reconstruction in-the-wild. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2019 June 15-20; Long Beach, CA, USA: IEEE. p. 10884–10894.

Güler, R. A., Neverova, N., and Kokkinos, I. (2018). Densepose: dense human pose estimation in the wild. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2024 June 16-22; Seattle, WA, USA: IEEE. p. 7297–7306.

Hamill, J, Gorton, G., and Masso, P. (2012). Clinical biomechanics: contributions to the medical treatment of physical abnormalities. *Kinesiol. Rev.* 1, 17–23. doi:10.1123/krj.1.1.17

Hausamann, P., Sinnott, C. B., Daumer, M., and MacNeilage, P. R. (2021). Evaluation of the intel RealSense T265 for tracking natural human head motion. *Sci. Rep.* 11 (1), 12486. doi:10.1038/s41598-021-91861-5

Hellstén, T., Karlsson, J., Häggblom, C., and Kettunen, J. (2022). Towards accurate computer vision-based marker less human joint localization for rehabilitation purposes. In: Scandinavian conference on health informatics; 2022 August 22-24. p. 21–25. doi:10.3384/ecp187004

Hidalgo, G., Raaj, Y., Idrees, H., Xiang, D., Joo, H., Simon, T., et al. (2019). Single-network whole-body pose estimation. In: Proceedings of the IEEE/CVF international conference on computer vision; 2023 October 1-6: IEEE, 6982–6991.

Holte, M. B., Tran, C., Trivedi, M. M., and Moeslund, T. B. (2012). Human pose estimation and activity recognition from multi-view videos: comparative explorations of recent developments. *IEEE J. Sel. Top. Signal Process.* 6 (5), 538–552. doi:10.1109/JSTSP.2012.2196975

Hulleck, A. A., Mohseni, M., Hantash, M. K. A., Katmah, R., Almadani, M., Arjmand, N., et al. (2023). Accuracy of computer vision-based pose estimation algorithms in predicting joint kinematics during gait. *Res. Square.* doi:10.21203/rs.3.rs-3239200/v1

Ino, T., Samukawa, M., Ishida, T., Wada, N., Koshino, Y., Kasahara, S., et al. (2024). Validity and reliability of OpenPose-Based motion analysis in measuring knee valgus during drop vertical jump test. *J. Sports Sci. Med.* 23 (3), 515–525. doi:10.52082/jssm.2024.515

Islam, T., Miron, A., Liu, X., and Li, Y. (2024). Deep learning in virtual try-on: a comprehensive survey. *IEEE Access* 12, 29475–29502. doi:10.1109/ACCESS.2024.3368612

Ismail, S. I., Osman, E., Sulaiman, N., and Adnan, R. (2016). Comparison between marker-less kinect-based and conventional 2D motion analysis system on vertical jump kinematic properties measured from sagittal view. In: P. Chung, A. Soltoggio, C. W. Dawson, Q. Meng, and M. Pain, editors. Proceedings of the 10th international symposium on computer science in sports (ISCSS). Cham: Springer International Publishing. p. 11–17. doi:10.1007/978-3-319-24560-7_2

Jacobsson, M., Willén, J., and Swarén, M. (2023). A drone-mounted depth camera-based motion capture system for sports performance analysis. In: H. Degen, and S. Ntoa, editors. *Artificial intelligence.* Cham: Springer Nature Switzerland. p. 489–503. doi:10.1007/978-3-031-35894-4_36

Jafarzadeh, P., Virjonen, P., Nevalainen, P., Farahnakian, F., and Heikkonen, J. (2021). Pose estimation of hurdles athletes using OpenPose. In: 2021 international conference on electrical, computer, communications and mechatronics engineering (ICECCME); 2021 June 12-13; Kuala Lumpur, Malaysia: IEEE. p. 1–6. doi:10.1109/ICECCME52200.2021.9591066

Khan, G., Maraha, H., Li, Q., and Pimbblet, K. (2024). A brief review of recent advances in AI-Based 3D modeling and reconstruction in medical, education, surveillance and entertainment. In: 2024 29th international conference on automation and computing (ICAC); 2024 August 28-30: IEEE. p. 1–6. doi:10.1109/ICAC61394.2024.10718770

Kurillo, G., Hemingway, E., Cheng, M.-L., and Cheng, L. (2022). Evaluating the accuracy of the azure kinect and kinect v2. *Sensors (Basel)* 22 (7), 2469. doi:10.3390/s22072469

Latyshev, M., Lopatenko, G., Shandryhos, V., Yarmoliuk, O., Pryimak, M., and Kvasnytsia, I. (2024). Computer vision technologies for human pose estimation in exercise: accuracy and practicality. *Soc. Integration. Educ. Proc. Int. Sci. Conf.* 2, 626–636. doi:10.17770/sie2024vol2.7842

Lee, J., Li, Y., Bunarto, D., Lee, E., Wang, O. H., Rodriguez, A., et al. (2024). Towards AI-Powered AR for enhancing sports playability for people with low vision: an exploration of ARSports: In 2024 IEEE international symposium on mixed and augmented reality adjunct (ISMAR-Adjunct); 2024 Oct 21-25. p. 228–233. doi:10.1109/ISMAR-Adjunct64951.2024.00055

Li, Y.-C., Chang, C.-T., Cheng, C.-C., and Huang, Y.-L. (2021). Baseball swing pose estimation using openpose: In 2021 IEEE international conference on robotics, automation and artificial intelligence (RAAI); 2021 April 21-23: IEEE. p. 6–9.

Liang, Z., and Nyamasvisva, T. E. (2023). Badminton action classification based on human skeleton data extracted by AlphaPose. In: 2023 international conference on sensing, measurement and data analytics in the era of artificial intelligence (ICSMD); 2023 Nov 2-4: IEEE. p. 1–4. doi:10.1109/ICSMD60522.2023.10490491

Liu, P.-L., and Chang, C.-C. (2022). Simple method integrating OpenPose and RGB-D camera for identifying 3D body landmark locations in various postures. *Int. J. Industrial Ergonomics* 91, 103354. doi:10.1016/j.ergon.2022.103354

Lovanshi, M., and Tiwari, V. (2022). Human pose estimation: benchmarking deep learning-based methods. In: 2022 IEEE conference on interdisciplinary approaches in technology and management for social innovation (IATMSI); 2022 Dec 21-23: IEEE. p. 1–6. doi:10.1109/IATMSI56455.2022.10119324

Maddipatla, Y., Li, J., Zheng, Y., and Li, B. (2023). Tracking and visualization of benchtop assembly components using a RGBD camera. In: *Presented at the ASME 2023 18th international manufacturing science and engineering conference.* New York, NY: American Society of Mechanical Engineers Digital Collection. doi:10.1115/MSEC2023-104622

Mahajan, P., Gupta, S., and Bhanushali, D. K. (2023). Body pose estimation using deep learning. *Int. J.* 11, 1419, 1424. doi:10.22214/ijraset.2023.49688

Mejia-Trujillo, J. D., Castaño-Pino, Y. J., Navarro, A., Arango-Paredes, J. D., Rincón, D., Valderrama, J., et al. (2019). Kinect^{TM} and intel RealSense^{TM} D435 comparison: a preliminary study for motion analysis. In: 2019 IEEE international conference on E-health networking, application and services (HealthCom); 2019 Oct 14-16: IEEE. p. 1–4. doi:10.1109/HealthCom46333.2019.9009433

Menychtas, D., Petrou, N., Kansizoglou, I., Giannakou, E., Grekidis, A., Gasteratos, A., et al. (2023). Gait analysis comparison between manual marking, 2D pose estimation algorithms, and 3D marker-based system. *Front. Rehabil. Sci.* 4 (Sep). doi:10.3389/fresc.2023.1238134

Moghari, M. D., Noonan, P., Henry, D, Fulton, R. R., Young, N., Moore, K., et al. (2024). Characterisation of the intel RealSense D415 stereo depth camera for motion-corrected CT perfusion imaging. *arXiv Preprint.* arXiv:2403.16490. doi:10.48550/arXiv.2403.16490

Mroz, S., Baddour, N., McGuirk, C., Juneau, P., Tu, A., Cheung, K., et al. (2021). Comparing the quality of human pose estimation with BlazePose or OpenPose. In: 2021 4th international conference on bio-engineering for smart technologies (BioSMART); 2021 Dec 8-10: IEEE. p. 1–4. doi:10.1109/BioSMART54244.2021.9677850

Mundt, M., Born, Z., Goldacre, M., and Alderson, J. (2023). Estimating ground reaction forces from two-dimensional pose data: a biomechanics-based comparison of AlphaPose, BlazePose, and OpenPose. *Sensors* 23, 78. doi:10.3390/s23010078

Napoli, A., Glass, S., Ward, C., Tucker, C., and Obeid, I. (2017). Performance analysis of a generalized motion capture system using microsoft kinect 2.0. *Biomed. Signal Process. Control* 38, 265–280. doi:10.1016/j.bspc.2017.06.006

Needham, L., Evans, M., Cosker, D. P., and Colyer, S. L. (2021). Can markerless pose estimation algorithms estimate 3D mass centre positions and velocities during linear sprinting activities? *Sensors* 21, 2889. doi:10.3390/s21082889

Patrizi, A., Pennestrì, E., and Valentini, P. P. (2016). Comparison between low-cost marker-less and high-end marker-based motion capture systems for the computer-aided assessment of working ergonomics. *Ergonomics* 59 (1), 155–162. doi:10.1080/00140139.2015.1057238

Pauzi, A. S. B., et al. (2021). Movement estimation using mediapipe BlazePose. In: H. Badioze Zaman, A. F. Smeaton, T. K. Shih, S. Velastin, T. Terutoshi, B. N. Jørgensen, et al. editors. *Advances in visual informatics.* Cham: Springer International Publishing. p. 562–571. doi:10.1007/978-3-030-90235-3_49

Pfister, A., West, A. M., Bronner, S., and Noah, J. A. (2014). Comparative abilities of microsoft kinect and vicon 3D motion capture for gait analysis. *J. Med. Eng. and Technol.* 38 (5), 274–280. doi:10.3109/03091902.2014.909540

Pham, Q.-T., Nguyen, D.-A., Nguyen, T.-T., Nguyen, T. N., Nguyen, D.-T., Pham, D.-T., et al. (2022). A study on skeleton-based action recognition and its application to physical exercise recognition. In: Proceedings of the 11th international symposium on information and communication technology, in SoICT '22. New York, NY, USA: Association for Computing Machinery. p. 239–246. doi:10.1145/3568562.3568639

Pilla-Barroso, M., Jiménez-Ruiz, A. R., and Jiménez-Martín, A. (2024). Estimation of movement in physical exercise programs using depth cameras: validation against a gold standard. In: 2024 IEEE international symposium on medical measurements and applications (MeMeA); 2024 June 26-28: IEEE. p. 1–6. doi:10.1109/MeMeA60663.2024.10596727

Qiao, S., Wang, Y., and Li, J. (2017). Real-time human gesture grading based on OpenPose. In: 2017 10th international congress on image and signal processing, BioMedical engineering and informatics (CISP-BMEI); 2017 Oct 14-16 (Shanghai, China: IEEE. p. 1–6. doi:10.1109/CISP-BMEI.2017.8301910

Ramasubramanian, A. K., Kazasidis, M., Fay, B., and Papakostas, N. (2024). On the evaluation of diverse vision systems towards detecting human pose in collaborative robot applications. *Sensors* 24 (2), 578. doi:10.3390/s24020578

Ren, F., Ren, C., and Lyu, T. (2025). IoT-based 3D pose estimation and motion optimization for athletes: application of C3D and OpenPose. *Alexandria Eng. J.* 115, 210–221. doi:10.1016/j.aej.2024.10.079

Saini, S. (2024). Yoga with deep learning: linking mind and machine. *SN Comput. Sci.* 5 (4), 427. doi:10.1007/s42979-024-02784-7

Sarıalioğlu, O., Balcı, İ. C., Sayİn, ı., Çiçek, B., Kakulia, N., Saltürk, S., et al. (2024). A study on tennis ball ground impact point detection. In: 2024 11th international conference on electrical and electronics engineering (ICEEE); 2024 April 22-24: IEEE. p. 377–381. doi:10.1109/ICEEE62185.2024.10779300

Schmidt, R. A., and Lee, T. D. (2019). *Motor learning and performance 6th edition with web study guide-loose-leaf edition: from principles to application*. Champaign, IL: Human Kinetics Publishers.

Seo, J., Alwasel, A., Lee, S., Abdel-Rahman, E. M., and Haas, C. (2019). A comparative study of in-field motion capture approaches for body kinematics measurement in construction. *Robotica* 37 (5), 928–946. doi:10.1017/s0263574717000571

Sharma, P., Shah, B. B., and Prakash, C. (2022). A pilot study on human pose estimation for sports analysis. In: D. Gupta, R. S. Goswami, S. Banerjee, M. Tanvee, and R. B. Pachori, editors. *Pattern recognition and data analysis with applications*. Singapore: Springer Nature. p. 533–544. doi:10.1007/978-981-19-1520-8_43

Shi, Q. W. (2014). Research on the application of motion capture of kinect technology in the sport training. *Adv. Mater. Res.* 926 (930), 2714–2717. doi:10.4028/www.scientific.net/AMR.926-930.2714

Song, Z., and Chen, Z. (2024). Sports action detection and counting algorithm based on pose estimation and its application in physical education teaching. *Informatica* 48 (10). doi:10.31449/inf.v48i10.5918

Stenum, J., Cherry-Allen, K. M., Pyles, C. O., Reetzke, R. D., Vignos, M. F., and Roemmich, R. T. (2021). Applications of pose estimation in human health and performance across the lifespan. *Sensors* 21 (21), 7315. doi:10.3390/s21217315

Suzuki, T., Tanaka, R., Takeda, K., and Fujii, K. (2024). Pseudo-label based unsupervised fine-tuning of a monocular 3D pose estimation model for sports motions. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2024 June 16-22: IEEE. p. 3315–3324.

Tsykunov, E., Ilin, V., Perminov, S., Fedoseev, A., and Zainulina, E. (2020). Coupling of localization and depth data for mapping using intel RealSense T265 and D435i cameras. *arXiv Preprint*. arXiv:2004.00269. doi:10.48550/arXiv.2004.00269

Wang, Q., Kurillo, G., Ofli, F., and Bajcsy, R. (2015). Evaluation of pose tracking accuracy in the first and second generations of microsoft kinect. In: 2015 international conference on healthcare informatics; 2015 Oct 21-23: IEEE. p. 380–389. doi:10.1109/ICHI.2015.54

Wang, T. (2024). RETRACTED: intelligent long jump evaluation system integrating blazepose human pose assessment algorithm in higher education sports teaching. *Syst. Soft Comput.* 6, 200130. doi:10.1016/j.sasc.2024.200130

Wei, J., Fu, X., Wang, Z., Yao, J., Zhao, J., and Wei, H. (2023). Research on human gait characteristics at different walking speeds based on alphapose. In: Proceedings of the 2023 15th international conference on bioinformatics and biomedical technology, in ICBBT '23. New York, NY, USA: Association for Computing Machinery. p. 184–189. doi:10.1145/3608164.3608191

Winter, D. A. (2009). *Biomechanics and motor control of human movement*.Hoboken, New Jersey: John Wiley and Sons.

Wu, C.-H., Wu, T.-C., and Lin, W.-B. (2023). Exploration of applying pose estimation techniques in table tennis. *Appl. Sci.* 13 (3), 1896. doi:10.3390/app13031896

Xu, J. (2024). Basketball tactics analysis based on improved openpose algorithm and its application. *J. Electr. Syst.* 20 (3s), 374–383. doi:10.52783/jes.1303

Yang, J., and Park, K. (2024). Improving gait analysis techniques with markerless pose estimation based on smartphone location. *Bioengineering* 11 (2), 141. doi:10.3390/bioengineering11020141

Yang, Z., Chen, H., Cheng, M., Liu, W., Chen, Y., Cao, Y., et al. (2024). Design of a lightweight human pose estimation algorithm based on AlphaPose. In: 2024 international conference on networking, sensing and control (ICNSC); 2024 Oct 18-20. p. 1–6. doi:10.1109/ICNSC62968.2024.10760165

Zanela, A., Schirinzi, T., Mercuri, N. B., Stefani, A., Romagnoli, C., Annino, G., et al. (2022). Using a video device and a deep learning-based pose estimator to assess gait impairment in neurodegenerative related disorders: a pilot study. *Appl. Sci.* 12, 4642. doi:10.3390/app12094642

Zhang, D., Wu, Y., Guo, M., and Chen, Y. (2021). Deep learning methods for 3D human pose estimation under different supervision paradigms: a survey. *Electronics* 10 (18), 2267. doi:10.3390/electronics10182267

Zhao, J., Cao, Y., and Xiang, Y. (2024). Pose estimation method for construction machine based on improved AlphaPose model. *Eng. Constr. Archit. Manag.* 31 (3), 976–996. doi:10.1108/ecam-05-2022-0476

Zhu, X., and Song, Q. (2021). Joint model for human body part instance segmentation and DensePose estimation. In: Proceedings of the 2021 9th international conference on communications and broadband networking. New York, NY, USA: Association for Computing Machinery. p. 66–73. doi:10.1145/3456415.3456426