A Pose-estimation Physical movement Training System with User-created Content

Atima Tharatipyakul[1]
Suporn Pongnumkul[2]
National Electronics and Computer Technology Center
Pathumthani, Thailand
[1]atima.tharatipyakul@gmail.com
[2]suporn.pongnumkul@nectec.or.th

Dulyawat Wiriyaphong[3]
Sarunya Kanjanawattana[4]
Gun Bhakdisongkhram[5]
Suranaree University of Technology
Nakhonratchasima, Thailand
[3]m6401054@g.sut.ac.th, [4]sarunya.k@sut.ac.th, [5]gunbhak@sut.ac.th

***Abstract*—Physical movement training systems can offer structured, accessible, and personalized fitness or rehabilitation. While many existing systems pose estimation for real-time feedback and movement assessment, they often lack support for user-generated content. This study presents an online physical training system enabling trainers to upload custom videos and utilize pose estimation for real-time feedback and posture similarity assessment. A pilot feasibility study with five healthy participants yielded a System Usability Score of 78.5 and high usefulness ratings (4.8/5 for rehabilitation, 4.6/5 for frozen shoulder diagnosis). Posture similarity scores were closely aligned with expert ratings (0.16 difference). However, several challenges remain for clinical use. Future work should focus on improving 3D pose estimation accuracy, conducting broader clinical validation, developing vision-based usage guidelines, and integrating human expertise with AI to foster ecosystem adoption.**

***Index Terms*—Pose estimation, Movement guidance, Frozen shoulder assessment, Online health application**

I. Introduction

Physical body movement is important for both physical and mental well-being. Nowadays, online channels allow diverse groups—from fitness enthusiasts to rehabilitation patients—to practice remotely. However, while convenient, asynchronous training often leaves trainees uncertain about their form, and live sessions risk overburdening trainers. Recent advances in computer vision address these limitations by enabling systems to provide automated, real-time feedback on user postures, enhancing personalized care and accessibility.

Interests in using pose estimation with online training systems are rising [1], [2], with applications including rehabilitation and sports. However, most existing systems use data collected in controlled laboratories. Systems utilizing user-created content remain limited. Leveraging trainer-uploaded videos as a resource allows for scalable, expert-driven instructional material that reflects current expertise without relying on laboratory constraints.

In this paper, we define a **trainer** as an individual movement designer (e.g., physiotherapist or fitness instructor) and a **trainee** as the practitioner of such movements. We presents a practical system prototype where trainers upload demonstration videos as references for trainees. The system facilitates intuitive visual comparison as well as utilizes existing pose estimation library to quantitatively assess posture against trainer-defined criteria, providing real-time feedback to enhance movement accuracy. We contribute a pilot feasibility study and discuss challenges for deploying integrated pose-estimation systems.

II. Background

This section provides background and related works on pose estimation, physical movement training systems, and movement assessment.

Pose Estimation. Human pose estimation identifies body points such as elbows or wrists, using inertial measurement units (IMUs), visual markers, or marker-free systems [3], yielding 2D or 3D results [4]. Challenges persist in visual data: 2D estimation suffers from occlusion and lack of visual cues, though temporal data helps [5]. 3D estimation struggles with complex backgrounds. While multi-view integration [6] addresses these issues, it remains computationally expensive [7].

We focus on accessible, camera-based ML open-source libraries like OpenPose [8] and MediaPipe [9] that can detect real-time posture using only conventional cameras. We selected MediaPipe for its versatility and efficiency on resource-constrained devices. Recently, Saraswat and Malathi [10] demonstrated that MediaPipe could achieve 95.84% accuracy in fall detection. It is currently a dominant library for physical movement feedback [2]. Unlike works solely improving accuracy, this paper examines system limitations and address inaccuracies from both technological and human perspectives.

Physical movement training systems. Several physical movement training systems adopt pose estimation to analyze and understand human movement, such as PoseNet for kickboxing [11] and OpenPose for sports

This research was funded by National Electronics and Computer Technology Center (NECTEC), National Science and Technology Development Agency (NSTDA), Thailand. It is supported by (1) Suranaree University of Technology (SUT), (2) Thailand Science Research and Innovation (TSRI), and (3) National Science, Research and Innovation Fund (NSRF) (NRIIS Project Number 179264).

forms [12]. However, efficacy with diverse user-created content remains unclear. While Tharatipyakul et al [13] utilized YouTube content, the system lacked manual upload capabilities and relied on controlled settings. Recently, Jaya et al [14] achieved high accuracy in strength training using MediaPipe on expert-recorded data. Still, their assessment limited to repetition counter. This paper proposes a system enabling user-uploaded training videos, automated feedback via user-customized rules, and visual comparison. The practical prototype was evaluated in a pilot study of frozen shoulder assessment movement in clinical settings.

Movement assessment. Intelligent movement assessment, typically assessing user activity and delivering user feedback, is well-established [15], yet camera-based pose estimation remains relatively emerging [16], [17]. A review [2] categorizes assessment methods into: mathematical models (e.g., angular similarity), rule-based methods, and machine learning. In this paper, we implemented a mathematical formula and a rule-based method. Our contribution is an end-to-end workflow that enables non-technical trainers to create and customize evaluation rules directly from their videos and to provide both automated and manual feedback to trainees.

III. System Design and Implementation

We designed and implemented a system that enables users to create a training lesson, practice, assess, and review the movement using a laptop without the need for high-cost equipment.

Create: A trainer uploads a video recorded using any application and specifies movement assessment rules (e.g., target body parts and angle thresholds). The system then pre-processes the trainer's pose data to improve runtime efficiency. To support reliable assessment, the uploaded video must contain only a single person.

Practice: The trainee uses a webcam to practice while viewing the trainer's video. The system performs real-time pose estimation on the trainee, compares their movement to the trainer's, and provides automated visual feedback.

Assess: Post-practice, the system displays a summary of the assessment results (e.g., similarity score). The trainee can optionally submit their video recording and results report, along with questions, to the trainer for subsequent manual feedback.

Review: The trainee reviews their recorded performance, utilizing both the system's automated feedback (for rapid understanding) and the trainer's manual feedback (for nuanced guidance and confidence building).

We describe the core features that facilitate the practice, assessment, and review of movements in next subsections.

A. Pose Estimation

Pose estimation determines the positions and orientations of a person's joints in a video recording or live video feed. It tracks 33 body landmark locations, which serve as inputs for similarity comparison, angle measurement, range of motion measurement, and visualization.

The system employs MediaPipe [9] for pose estimation, leveraging its direct output. The server utilizes the BlazePose GHUM Heavy model (highest accuracy, slowest), while the client uses the faster, less accurate BlazePose GHUM Full model. The architecture is flexible, allowing for easy updates or replacement of the pose estimation method, provided it supplies the landmarks for the shoulders, elbows, wrists, hips, knees, and ankles. This ensures adaptability as technology advances.

Although the system primarily uses 2D coordinates, MediaPipe's intrinsic provision of 3D points enabled preliminary experimentation with 3D data (Section IV).

B. Pose Similarity Comparison

Pose comparison is essential in physical movement training for ensuring accurate alignment. Similarity analysis comprises three parts: defining body parts, visualizing target positions, and calculating the similarity score.

1) Define Interested Body Parts: Selecting specific body segments (e.g., arms, legs) for movement training comparison is crucial for increasing learning efficacy. First, this segment-specific feedback ensures high relevance to the practiced movement, minimizing distraction and enhancing effectiveness (e.g., focusing on arms during bicep curls, not legs). Second, it facilitates incremental part learning, where sub-tasks are mastered individually before integration into the whole task. This approach is superior to whole learning for complex movements [18].

Our system allows trainers to specify default interested body parts (e.g., left/right upper arm, hip line) during lesson creation. Trainees can further customize the parts for personalized feedback during practice.

2) Visualize Target Positions: The system offers real-time, frame-by-frame visual guidance by overlaying the trainer's reference pose onto the trainee's pose, enabling precise movement adjustments.

To account for body structural differences, the trainer's pose is mapped to the trainee's pose. This involves matching joints and projecting the trainer's body part vectors ($\vec{v_i}$) onto the trainee's corresponding vectors ($\vec{u_i}$). The resulting vector, $\vec{u}_{target_i}$, retains the magnitude of $\vec{u_i}$ but adopts the direction of $\vec{v_i}$ (Equation 1).

$$\vec{u}_{target_i} = \frac{\vec{v_i}}{\|\vec{v_i}\|} \times \|\vec{u_i}\| \tag{1}$$

The mapping process adheres to body structure hierarchy: upper body parts (e.g., upper arms) are mapped first, establishing new target joint locations (e.g., elbows) for subsequent lower parts (e.g., lower arms). The system calculates the angular difference for each body part between the trainer and trainee. To ensure the system captures only significant postural deviations, the feedback is shown only when the discrepancy exceeds a trainer-defined threshold.

The target vectors are visualized as red dashed lines, analogous to handwriting practice guidelines. Trainees should aim to align their body parts with these lines. Alternative visualization strategies, such as highlighting differences or marking incorrect parts, are also available.

3) Calculate Similarity Score: In addition to visual guidance, the system calculates a numeric similarity score to objectively grade trainee performance and track improvement. Scores are calculated per frame but displayed as a final average upon practice completion.

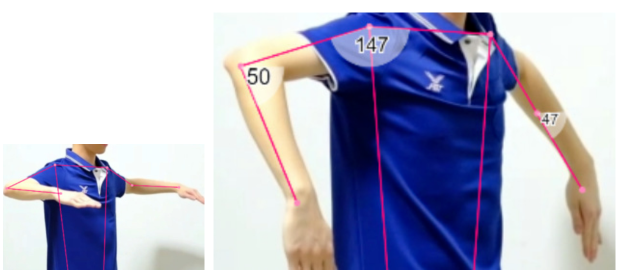
Cosine similarity was selected as the core metric due to its superior accuracy compared to Euclidean distance and raw angle difference [19]. The system averages the cosine similarity scores across all relevant body parts detected in the current frame. Undetected trainer body parts are ignored. If a trainer body part is detected but the corresponding trainee part is not, that part's score is set to zero to ensure thorough evaluation.

The final average score is presented along with a trainer-defined qualitative rating (e.g., Good, Excellent) to summarize the trainee's overall alignment.

C. Degree Measurements

Beyond pose similarity, the system implements degree measurements to assess the trainee's form against trainer-defined custom rules. Inspired by literature review [2], we identified two key measurements: angle of body parts and range of motion (ROM), useful across various movement types. The angle of body parts measures the angle between two lines within a frame where as ROM measures a body part's angle across frames.

Unlike systems with fixed conditions, our platform allows trainers to flexibly define the expected angles and ROM thresholds for any number of measurements during lesson creation. The system processes these rules frame-by-frame and visualizes the results during playback (Figure 1) as well as in the final summary report.



(a) (b)

Fig. 1: Visualization of the results when measuring the angle of body parts and range of motion (ROM) at 2nd second (a) and 7th second (b) of a video. *At* the 7th second, the angle between the right upper arm and shoulder is 147°. *Between* the 2nd to the 7th second, the ROM of the left and right lower arms are 47° and 50°, respectively.

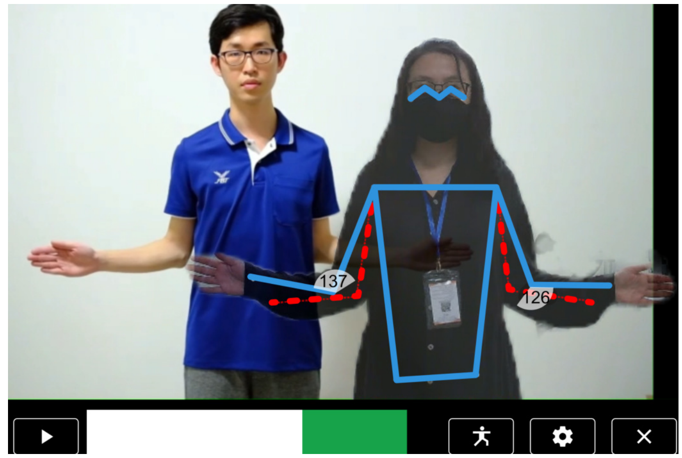


Fig. 2: Example of user interface of video playback.

D. User Interface for Video Playback

The system utilizes a unified video playback component (Figure 2) for practice, assessment, and review. All feedback—including pose estimation, pose similarity, and degree measurement results—is visualized directly onto the trainee's body. The trainer's video remains annotation-free for clear observation. To enhance focus, we use MediaPipe for background removal, allowing the trainer's and trainee's video feeds to be positioned closely within the central visual field. These design choices are based on prior experiment in a laboratory setting [13]. Additionally, measurement results are presented in a timeline view using color-coded bars: blue-gray for unmeasured, and green/red bars to indicate pass/fail against trainer-defined thresholds.

IV. Pilot Feasibility Study

As a case study, we experimented the system's ability to access frozen shoulder symptoms. This experiment, performed at the Rehabilitation Department of a local hospital, aimed to measure system error and identify challenges within a rehabilitation context.

A. Participants

Recruitment was conducted via word-of-mouth within the computer engineering department of a local university. A total of five healthy individuals (4 males, 1 female; age range: 23–31 years, $M = 25.8$, $SD = 3.35$) were enrolled.

B. Apparatus

The system was run on a laptop with 11th Gen Intel(R) Core(TM) i7-1165G7 processor at 2.80GHz, with 16 GB of RAM and Intel(R) Iris(R) Xe Graphics. We run the back-end in Docker on the laptop to eliminate Internet issues and run the front-end in a Chrome web browser. The feedback was shown only when the angular discrepancy exceeded 10°, chosen empirically for the pilot setup.

The laptop had a 14" monitor with 1920 × 1080 pixels resolution. Four Hoco DI01 web cameras were used as input devices. Figure 3 shows the experimental setup.

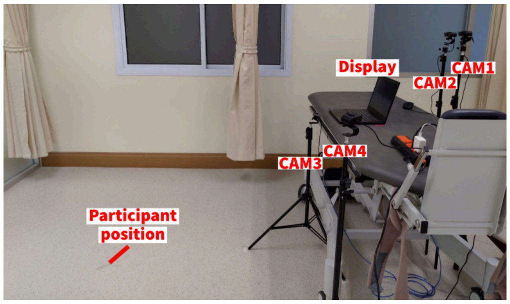


Fig. 3: Participants stood at the mark, then followed movements and received feedback via the laptop display and CAM1. Additional cameras, simulating typical spatial constraints, were used solely for evaluation.

We recorded five training videos in our native language: elevation through abduction, elevation through flexion, extension, external rotation, and internal rotation (VID 1-5, respectively). The movements and camera angle were imitated from a third-party video (https://youtu.be/cP4LLJie9kw), selected by searching for "shoulder range of motion" and limiting results to videos featuring a single person.

C. Procedure

After obtaining informed consent, the experimental session began with an explanation of the system features using a sample video. Participants were then allowed a brief period to freely interact with the system using this sample. Upon readiness, participants were instructed to perform a series of predefined movements at designated locations, prioritizing safety by avoiding forced motion.

For each movement, we played the video once while the participant stayed still to study the movement. The video played again, prompting the participant to begin following the movement. The participant movement was recorded. Once the video ended, the participant sustained the final posture while a researcher measured body angles using a goniometer. A traditional goniometry procedure followed, involving aligning the stationary and moving arms of the goniometer with predefined anatomical landmarks.

Finally, participants completed a post-experiment questionnaire covering demographics, system usability, usage issues, and general comments. The whole process took 30 - 40 minutes. After all participants completed the procedure, two researchers watched the recordings and individually rated whether the participant movement was similar to the trainer movement in 5-point scale.

D. Results

Measurement Error. We collected 2D and 3D landmark data. As described in the procedure, we tested two manual measurement methods. The first method, sustained posture, risks manual error in locating anatomical landmarks. The second method, traditional procedure, risks comparison error since movement was performed twice. We found that the angular differences between the two methods were minor (7.32° on average). Since the traditional procedure better reflects a standard clinical measurement, it was chosen for all subsequent analyses to evaluate the system's error.

The error is defined as the absolute angular difference between the system calculated measurement and the manual measurement using traditional procedure.

Figure 4 shows the error grouped by camera view. The differences between views were low, especially with 3D points. The average error of CAM1, CAM2, CAM3, and CAM4 were 26.02°, 27.54°, 30.74°, and 36.36° for 2D points and 24.76°, 20.86°, 23.96°, and 23.92° for 3D points.

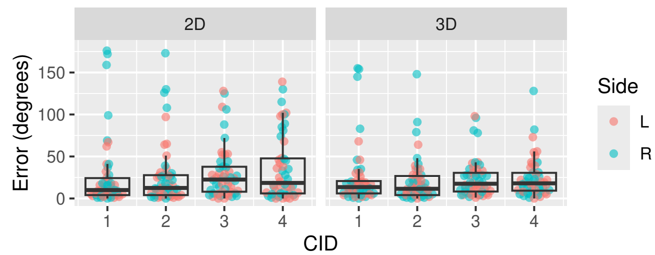


Fig. 4: Absolute angular differences using 2D points and 3D points. Errors are grouped by camera view and colored by the shoulder side (Left/Right).

Figure 5 shows the error grouped by movement video. The error were particularly high when the movement involved z-axis (i.e., external and internal rotation). The average error of movement 1 to 5 were 9.18°, 35.85°, 12.68°, 46.30°, and 46.83° for 2D points and 14.28°, 34.95°, 9.05°, 21.23°, and 37.38° for 3D points.

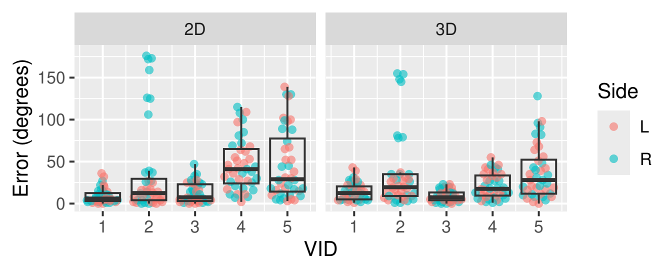


Fig. 5: Absolute angular differences using 2D points and 3D points. Errors are grouped by movement video and colored by the shoulder side (Left/Right).

Similarity Error. The error of the system's similarity score, which utilizes a 5-point grading scale, was assessed by calculating the absolute score difference between the system's output and the average of two manual expert grading. The average error was 0.16.

Usability. The system received the average system usability score of 78.5 (Good). The participants also rated the usefulness of the application in the following areas (1 - 5, where five means the participant strongly agrees that the application will be useful in the area):

- Diagnosis of frozen shoulder: 4.6
- Help in rehabilitation: 4.8
- Help in dance training: 4
- Help in practicing yoga: 4.4
- Prevent injury from exercise incorrectly: 4.2

Participant feedback centered on movement difficulty and interface improvements. P3 noted difficulty viewing movements with the body side facing the screen and suggested using fixation tools for the head/neck. P1 found sustained positions tiring and recommended posture assessment (e.g., walking/standing). P4 suggested relocating the virtual trainer (left side) for better viewing.

E. Discussion

Overall, the system shows potential for clinical use, supported by a good usability score (78.5, exceeding the reported mean of 68 for digital health applications [20]) and usefulness ratings. However, the measurement error was still too high. We therefore implemented preliminary enhancements and re-evaluated the system. Specifically, we upgraded the underlying pose estimation libraries (from Python v0.8.11/JS v0.4.16 to Python v0.10.11/JS v0.5.16) and refined the degree measurement rules by extending the pose assessment duration to three seconds.

A subsequent evaluation using data from CAM1 showed substantial error reductions: from 26.02° (SD = 41.82) to 13.00° (SD = 11.78) for 2D measurements, and from 24.76° (SD = 36.19) to 15.20° (SD = 11.47) for 3D measurements. The final errors align with those reported for pose estimation versus marker-based motion capture in athletic ($9.7° \pm 4.7°$) and sports contexts ($9.0° \pm 3.3°$) [21].

The system demonstrated excellent concurrent validity (ICC = 0.899 for 2D, 0.922 for 3D; $p < .001$), indicating high consistency in ranking movements relative to the human baseline. However, Bland-Altman analysis (Figure 6) revealed that while the mean biases were low, the 95% limits of agreement ($-55.64°$ to 51.72° for 2D; $-38.9°$ to 51.0° for 3D) exceeded the acceptable threshold for clinical utility. Furthermore, as the dataset involved multiple measurements per participant, the results may be influenced by within-subject correlation.

This study was restricted to a small cohort of healthy participants ($N = 5$) to identify necessary technical refinements before proceeding to a larger-scale trial. Beyond improving measurement accuracy, the following issues must be addressed to facilitate future clinical adoption.

Occlusion and 3D motion. Errors were particularly high for movements involving occluded body parts (VID2). Interestingly, the low error observed in VID3, where a similar occlusion was presented, suggests the system may infer the occluded position. Alternatively, mitigating occlusion

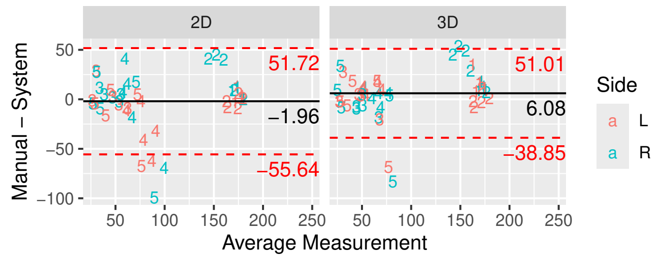


Fig. 6: Bland-Altman Analysis using 2D points and 3D points from the updated system. The number denotes movement VID and is colored by the shoulder side (Left/Right).

can be achieved by adjusting the camera or body angle. The average errors of VID2 were reduced from 59.30° (2D) and 58.90° (3D) with CAM1 (front view) to 5.80° (2D) and 10.30° (3D) with CAM4 (45° view).

Due to perspective distortion in 2D measurements, 3D measurements performed better for z-axis movements (VID4 and VID5), as indicated by fewer data points falling below the lower limits of agreement in Figure 6. However, while MediaPipe provides 3D pose estimates from 2D imagery, we observed occasional rotational inconsistencies, with reconstructed poses appearing tilted relative to the ground plane. Future integration of enhanced approaches, such as Improved MediaPipe [22], may help improve 3D point accuracy and overall usability.

Measurement method. The rule-based measurement approach is intuitive and easily tunable. One factor reducing the updated system's error was the extended measurement duration, increasing the likelihood of correct body part localization. Nevertheless, accurate 3D point estimation is essential. Our general-purpose angle method uses body parts (e.g., shoulder-hip line for the y-axis) as references, deviating from the traditional procedure. Higher 3D accuracy would enable closer adherence to traditional methods, aligning with the need for precise measurements highlighted by El-Rajab et al. [17].

Pose estimation preciseness. Human body keypoint precision from current pose estimation libraries might be inadequate for clinical usage. MediaPipe, for instance, detects the shoulder area but misses the crucial acromion process (Figure 7) used by experts. Developing detailed body landmarks, akin to MediaPipe's 468 3D face landmarks, is a promising direction for clinical integration.

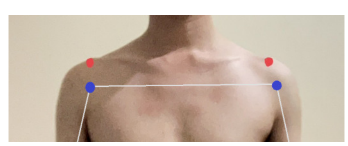


Fig. 7: The system could detect the shoulder area (the blue dots) but not the acromion process (the red dots).

Clothing. Clothing can impede participants from performing the posture correctly (e.g., avoiding maximum arm elevation to prevent exposing their midriff). Clear instructions on appropriate clothing should be provided, similar to the protocols in a traditional clinical diagnosis.

Wrong posture. Despite the apparent simplicity of the movements, participants often used compensatory motions to inflate their ROM (e.g., dropping the upper arm during internal rotation). The overlaying visual cues were insufficient to prompt self-correction, as users prioritized increasing ROM. Consequently, the current system is better suited for a clinical setting where experts can supervise and correct posture. With acceptable error, the system could aid screening and digital history tracking. For home use, a feature to detect and actively alert users to typical compensation is necessary.

Other factors. Clinical diagnosis relies on factors beyond body degree measurement, e.g., pain levels and patient profile. For example, using only the internal rotation measurement misrepresented healthy participants. Thus, the system should incorporate clinical factors and provide a probabilistic assessment instead of binary feedback. Establishing an ecosystem that engages human experts in evaluating user performance is also recommended.

V. Conclusion and Future Directions

We presented an online physical movement training system that allows trainers to upload videos for trainees to follow or practice. The system uses pose estimation technology to provide real-time visual comparison and feedback, and assesses performance based on posture similarity or trainer-defined rules. The user study (N = 5) demonstrated good usability and the measurement errors: 13.00° (SD = 11.78) for 2D points and 15.20° (SD = 11.47) for 3D points. Yet, several barriers to clinical adoption persist. Future research should prioritize improving 3D pose estimation accuracy and conducting long-term clinical validation with actual patients to confirm utility. Subsequently, developing vision-based usage guidelines (e.g., clothing, camera angle) and an expert-AI ecosystem will facilitate system adoption for physical movement training.

Acknowledgment

We thank Suwittaya Thienpratharn, Nithita Riansumrit, and Thanawat Srikaewsiew for their support. LLMs were used solely as a language assistant.

References

[1] G. M. Difini, M. G. Martins, and J. L. V. Barbosa, "Human Pose Estimation for Training Assistance: A Systematic Literature Review," in *ACM International Conference Proceeding Series*, 2021, pp. 189–196.

[2] A. Tharatipyakul, T. Srikaewsiew, and S. Pongnumkul, "Deep learning-based human body pose estimation in providing feedback for physical movement: A review," *Heliyon*, 2024.

[3] Y. Niu, J. She, and C. Xu, "A survey on imu-and-vision-based human pose estimation for rehabilitation," in *2022 41st Chinese Control Conference (CCC)*. IEEE, 2022, pp. 6410–6415.

[4] Y. Desmarais, D. Mottet, P. Slangen, and P. Montesinos, "A review of 3d human pose estimation algorithms for markerless motion capture," *Computer Vision and Image Understanding*, vol. 212, p. 103275, 2021.

[5] H. Chen, R. Feng, S. Wu, H. Xu, F. Zhou, and Z. Liu, "2d human pose estimation: A survey," *Multimedia Systems*, vol. 29, no. 5, pp. 3115–3138, 2023.

[6] C. Xu, A. Li, L. Chen, Y. Liu, R. Shi, H. Su, and M. Liu, "Sparp: Fast 3d object reconstruction and pose estimation from sparse views," in *European Conference on Computer Vision*. Springer, 2025, pp. 143–163.

[7] X. Xi, C. Zhang, W. Jia, and R. Jiang, "Enhancing human pose estimation in sports training: Integrating spatiotemporal transformer for improved accuracy and real-time performance," *Alexandria Engineering Journal*, vol. 109, pp. 144–156, 2024.

[8] CMU Perceptual Computing Lab, "OpenPose," accessed: 2023-02-20. [Online]. Available: https://github.com/CMU-Perceptual-Computing-Lab/openpose

[9] Google, "MediaPipe," accessed: 2023-02-20. [Online]. Available: https://google.github.io/mediapipe

[10] S. Saraswat and G. Malathi, "Pose estimation based fall detection system using mediapipe," in *2024 10th International Conference on Communication and Signal Processing (ICCSP)*. IEEE, 2024, pp. 1733–1738.

[11] E. Wessa, A. Ashraf, and A. Atia, "Can pose classification be used to teach Kickboxing?" in *International Conference on Electrical, Computer, and Energy Technologies, ICECET 2021*, Helwan University, HCI-LAB, Faculty of Computers and Artificial Intelligence, Egypt, 2021.

[12] S. Akiyama and N. Umezu, "Similarity-based Form Visualization for Supporting Sports Instructions," in *LifeTech 2022 - 2022 IEEE 4th Global Conference on Life Sciences and Technologies*, Ibaraki University, Dept. Mechanical System Engineering, Hitachi, Japan, 2022, pp. 480–484.

[13] A. Tharatipyakul, K. T. Choo, and S. T. Perrault, "Pose estimation for facilitating movement learning from online videos," in *Proceedings of the International Conference on Advanced Visual Interfaces*, 2020, pp. 1–5.

[14] A. I. T. Jaya, P. Puspitaningayu, A. P. Adiwangsa, and N. Funabiki, "Two-dimensional human pose estimation using key points' angular detection for basic strength training," *Journal of Intelligent System and Telecommunication*, vol. 1, no. 1, pp. 105–119, 2024.

[15] A. Rajšp and I. Fister, "A systematic literature review of intelligent data analysis methods for smart sport training," *Applied Sciences (Switzerland)*, vol. 10, no. 9, p. 3013, 2020.

[16] K. M. Tsiouris, V. D. Tsakanikas, D. Gatsios, and D. I. Fotiadis, "A Review of Virtual Coaching Systems in Healthcare: Closing the Loop With Real-Time Feedback," *Frontiers in Digital Health*, vol. 2, p. 567502, 2020.

[17] I. El-Rajab, T. J. Klotzbier, H. Korbus, and N. Schott, "Camera-based mobile applications for movement screening in healthy adults: a systematic review," *Frontiers in Sports and Active Living*, vol. 7, p. 1531050, 2025.

[18] F. E. Fontana, O. Furtado Jr, O. Mazzardo, and J. D. Gallagher, "Whole and part practice: A meta-analysis," *Perceptual and motor skills*, vol. 109, no. 2, pp. 517–530, 2009.

[19] T. Srikaewsiew, K. Khianchainat, A. Tharatipyakul, S. Pongnumkul, and S. Kanjanawattana, "A comparison of the instructor-trainee dance dataset using cosine similarity, euclidean distance, and angular difference," in *2022 26th International Computer Science and Engineering Conference (ICSEC)*. IEEE, 2022, pp. 235–240.

[20] M. Hyzy, R. Bond, M. Mulvenna, L. Bai, A. Dix, S. Leigh, S. Hunt *et al.*, "System usability scale benchmarking for digital health apps: meta-analysis," *JMIR mHealth and uHealth*, vol. 10, no. 8, p. e37290, 2022.

[21] T. Fukushima, P. Blauberger, T. Guedes Russomanno, and M. Lames, "The potential of human pose estimation for motion capture in sports: a validation study," *Sports Engineering*, vol. 27, no. 1, p. 19, 2024.

[22] Y. Lin, X. Jiao, and L. Zhao, "Detection of 3d human posture based on improved mediapipe," *Journal of Computer and Communications*, vol. 11, no. 2, pp. 102–121, 2023.