

Original Paper

Evaluation of Smartphone Camera Positioning on Artificial Intelligence Pose Estimation Accuracy for Exercise Detection: Observational Study

Eduarda Oliosí^{1,2,3}, BS, MS; Soraia Ferreira^{1,4,5}, MS, PhD; Ana Paula Giordano^{1,6}, MS, PhD; Guilherme Viveiros⁷, MS; José Parraca^{4,5}, PhD; Paulo Pereira^{4,5}, MS; Federico Guede-Fernández^{1,2}, PhD; Salomé Azevedo^{1,7,8,9}, PhD

¹Value for Health CoLAB, Lisbon, Portugal

²Laboratory for Instrumentation, Biomedical Engineering and Radiation Physics (LIBPhys-UNL), Physics Department, Nova School of Science and Technology, Nova University of Lisbon, Caparica, Portugal

³Research Centre for Physical Activity, Health and Leisure (CIAFEL), Faculty of Sport, University of Porto, Porto, Portugal

⁴Departamento de Desporto e Saúde, Escola de Saúde e Desenvolvimento Humano, Universidade de Évora, Évora, Portugal

⁵Comprehensive Health Research Centre (CHRC), Escola de Saúde e Desenvolvimento Humano, Departamento de Ciências Médicas, Universidade de Évora, Évora, Portugal

⁶Research Unit in Business and Economics, Católica Lisbon School of Business and Economics (CUBE), Catholic University of Portugal, Lisbon, Portugal

⁷Department of Machine Learning, Dotmoovs, Braga, Portugal

⁸Department of Engineering and Management, Instituto Superior Técnico (CEG-IST), University of Lisbon, Lisbon, Portugal

⁹Comprehensive Health Research Center (CHRC), Nova Medical School, Nova University of Lisbon, Lisbon, Portugal

Corresponding Author:

Eduarda Oliosí, BS, MS

Value for Health CoLAB

15 Fontes Pereira de Melo Ave, 2nd Fl, Right

Lisbon 1050-115

Portugal

Phone: 351 937091767

Email: eduarda.oliosí@vohcolab.org

Abstract

Background: Artificial intelligence (AI)-driven pose estimation (PE) offers a scalable and cost-effective solution to track exercises in mobile health apps. However, occlusion, influenced by camera angle and distance, can reduce detection accuracy and repetition counting precision. The influence of smartphone positioning on these performance metrics remains underexplored in controlled studies.

Objective: The study aimed to examine how smartphone camera angle (front, side, and diagonal) and distance (90 cm, 180 cm, 200 cm, and 360 cm) affect detection performance and repetition counting accuracy during push-ups and squats using AI-based PE.

Methods: In this cross-sectional, within-subject study, 44 healthy university students (9 [20.5%] female participants; mean age 20.3 y, SD 0.4 y; mean BMI 23.2, SD 0.6 kg/m²) were assigned to perform either squats or push-ups. Each participant completed their assigned exercise across 12 predefined smartphone camera configurations, yielding approximately 264 squat trials (n=22) and 264 push-up trials (n=22). Each trial consisted of an average of 5 repetitions, totaling approximately 1320 repetitions per exercise. PE performance was assessed using binary classification accuracy, detection rate, and mean absolute error (MAE) for repetition counting. Generalized linear mixed-effects models evaluated classification odds, linear mixed-effects models analyzed MAE, and Tukey-adjusted post hoc tests followed significant effects.

Results: The mean detection rate was 61.1% (SD 48.8%) for push-ups and 61.5% (SD 48.7%) for squats, with MAEs of 1.08 (SD 1.78) and 1.11 (SD 1.82) repetitions, respectively. Push-ups were most accurately detected from diagonal views at 90 to 180 cm (up to 85.7% detection; MAE=0.28) and least accurately from the front at 360 cm (20%; MAE=2.70). Squats performed best from a diagonal view at 200 cm (95.5%; MAE=0.05) and worst from the side at 90 cm (0%; MAE=5). Generalized linear mixed models showed that for push-ups, the front 90 cm and diagonal 360 cm views significantly reduced classification odds compared to the side 90 cm view ($P=.03$ and $P=.04$, respectively), whereas for squats, diagonal and front

views significantly outperformed side views across all distances ($P < .001$). Post hoc tests confirmed that for push-ups, diagonal close or mid-range views had significantly lower MAEs than far front views, and for squats, diagonal and front views at 180 to 200 cm achieved the highest accuracy and lowest MAEs ($P < .05$).

Conclusions: AI-based PE effectiveness for exercise tracking is significantly affected by smartphone positioning. Diagonal and frontal views at mid-range distances (180-200 cm) provided the highest detection accuracy and counting precision. These findings offer actionable guidance for developers, clinicians, coaches, and users optimizing mobile health exercise monitoring.

JMIR Mhealth Uhealth 2026;14:e82412; doi: [10.2196/82412](https://doi.org/10.2196/82412)

Keywords: computer vision; digital health; human pose estimation; human activity recognition; mobile health; mhealth; mobile apps; physical activity

Introduction

Physical inactivity is a major global public health concern, contributing to higher rates of chronic diseases, reduced quality of life, and increased health care costs [1-3]. While the health benefits of regular physical activity are well documented, ranging from improved cardiovascular fitness and muscle strength to enhanced mental well-being, participation levels remain insufficient worldwide [4]. Current estimates indicate that over 31% of adults and 81% of adolescents fail to meet the World Health Organization's recommended physical activity guidelines [5]. The consequences of physical inactivity are particularly severe for populations managing chronic or neurological conditions, for whom movement-based interventions can serve as both preventive and therapeutic measures [3,6,7].

In response to the demand for scalable and accessible physical activity tools, mobile health (mHealth) apps have emerged as a cost-effective means of promoting exercise participation [3]. A notable innovation in this field is pose estimation (PE), or markerless motion capture, which uses computer vision to detect joint positions from standard video footage without the need for specialized equipment [8]. Open-source PE models, such as OpenPose (Carnegie Mellon University), BlazePose (Google), and MoveNet (Google), facilitate real-time PE via smartphones or webcams. This enables the provision of remote, personalized feedback [9-13]. In biomedical contexts, PE can identify movement impairment patterns and support rehabilitation or neurological diagnoses [8]. It is also widely applied in fitness, telerehabilitation, occupational health, and sports to track joint angles, velocity, form, and repetitions [14-18]. Validation studies have demonstrated strong agreement with the gold standard motion capture [19-21], and recent studies integrating deep learning have reported repetition-counting accuracies above 90% [22-25].

Despite recent advances, most PE systems have only been evaluated in controlled laboratory environments. However, real-world mHealth apps occur in uncontrolled settings where factors such as handheld devices held at different distances and angles, and partial obstructions (eg, from pets, furniture, or body position) can significantly degrade performance [10,12,26]. These challenges are particularly pronounced for dynamic, whole-body movements such as squats and push-ups, where tracking accuracy can vary depending on body position and elevation. Although a few commercial

solutions have begun to experiment with camera-based self-assessment (eg, Halo Movement [Amazon] [27] and Kaia Health GmbH [28]), there is still a lack of empirical studies on how smartphone positioning affects the reliability of PE in practical settings [10,14,27]. Improving accuracy in these contexts will likely require advances not only in 2D keypoint detection but also in our understanding of optimal camera placement and how to transform 2D keypoint data into accurate spatial representations.

This study systematically evaluates the impact of smartphone camera angle and distance on the performance of artificial intelligence (AI)-based PE when counting 2 foundational bodyweight exercises: squats and push-ups. These exercises are commonly used for fitness and rehabilitation tracking. They involve multijoint movements and present common occlusion challenges. Repetition counts are obtained from 2D PE across 12 camera configurations (3 angles \times 4 distances) and benchmarked against expert-labeled ground truth. The aim is to examine how smartphone recording position, specifically camera angle and distance, affects detection performance and repetition counting accuracy, and to inform best practices for using mHealth systems in real-world environments. This work is part of the Blockchain.PT initiative (Project No. 51), which supports sustainable blockchain-based digital innovation in the health and sports sectors.

Methods

Study Design

This cross-sectional, within-subject study was designed to assess the influence of smartphone camera configurations, specifically angle and distance, on the detection of PE during squats and push-ups. These exercises were selected due to their complex multijoint biomechanics and typical self-occlusions (eg, torso obscuring limbs in push-ups).

Setting

All data were collected in a controlled laboratory environment at the School of Health and Human Development, University of Évora, Portugal, in May 2025. Environmental conditions, including lighting and temperature, were standardized across trials. Smartphone placement was controlled at predefined angles and distances to ensure reproducible video capture. The laboratory provided enough space for participants to safely perform push-ups and squats.

Participants

Participants were recruited from undergraduate courses at the University of Évora in Portugal through a targeted study announcement that specified the eligibility criteria, study procedures, and recommended attire for physical exercise during classes. Eligible volunteers were identified through convenience sampling with eligibility screening.

A total of 44 healthy university students participated and self-reported their status. Inclusion criteria were: (1) aged 18 years or over; (2) able to safely perform squats and push-ups; (3) not having used sedative or balance-impairing medications within the last 24 hours; and (4) absence of neurological or musculoskeletal conditions that could impair movement.

Ethical Considerations

The study protocol was reviewed and approved by the Ethics Committee of the University of Évora (GD/27378/2024), in accordance with institutional and international standards for human participant research. All procedures complied with the Declaration of Helsinki and the General Data Protection Regulation. Written informed consent was obtained from all participants prior to enrollment. Participants were informed about the study objectives, procedures, and potential risks and were explicitly advised of their right to withdraw at any time without penalty. To ensure privacy and confidentiality, all collected data were anonymized and coded to prevent participant identification. Data were securely stored on password-protected institutional servers accessible only to the research team. No monetary or material compensation was provided for participation in this study.

Exercise Protocol and Dataset

Participants were first briefed on the study protocol and completed a demographic questionnaire to record baseline characteristics. For the squat task, participants stood with feet shoulder-width apart and arms extended forward, performing repetitions by lowering their hips until their thighs were parallel to the floor. Male participants performed push-ups in a standard plank position, while female participants were allowed a modified knee-supported variation. All exercise procedures adhered to the 11th edition of the

American College of Sports Medicine guidelines (2021) [29], which recommend these variations to accommodate differences in strength and fitness levels. To reduce fatigue-related performance decline and ensure high-quality movement data, participants were assigned to perform either squats or push-ups based on personal preference. Of the 44 participants, 22 (16 male participants and 6 female participants) completed the squat protocol, and 22 (19 male participants and 3 female participants) completed the push-up protocol. Each participant performed their assigned exercise across all 12 predefined camera configurations, yielding 264 squat trials (22×12) and 264 push-up trials (22×12). Each trial consisted of a continuous movement sequence, averaging approximately 5 repetitions per video [16], totaling approximately 1320 repetitions per exercise type (264 trials × ~5 reps).

Experimental Setup

To capture a representative sample of commonly used consumer devices, exercise sessions were recorded using three smartphone models, iPhone 11 (Apple Inc.), iPhone 13 (Apple Inc.), and Samsung Galaxy A52 (Samsung Electronics), each capturing video at 1080 p resolution and 30 frames per second. Smartphones were positioned horizontally on the floor at 3 fixed angles relative to the participant's body: frontal (0°), diagonal (45°), and lateral (90°). Recordings were conducted at 4 distances, including 180 cm based on prior literature [30], and 3 additional distances (90 cm, 200 cm, and 360 cm) defined by the research team to test a wider range of realistic recording conditions. This configuration resulted in 12 distinct camera setups (3 angles × 4 distances). All smartphone positions were marked and standardized to ensure consistency across trials.

This experimental setup is contextualized within existing research by Table 1, which summarizes previous studies on push-up and squat detection, detailing camera perspectives, distances, PE methods, and reported performance metrics. The comparison shows that most prior studies rely on limited viewpoints and fixed distances, restricting generalizability. In contrast, our multiangle, multidistance dataset addresses these constraints, providing a more diverse and representative resource for evaluating PE performance.

Table 1. Summary of state-of-the-art push-up and squat pose estimation datasets and methods.

Reference	Exercise	Camera perspectives	Camera distances	PE ⁿ method	Instrument	Key evaluation metrics	Dataset
Park et al [31]	Push-up	Frontal, side	Full body visible, N/R ^b value	OpenPose	2 cameras	ACC ^c =90%	Custom dataset of n=12
Youssef et al [32]	Squat	Frontal, side	N/R	BlazePose	Mobile devices + inertial sensors	ACC=94%	EJUST-SQUAT-21 ^o , single individual, MM-Fit ^e datasets
Hande et al [33]	Squat and others	Frontal, side	Full body visible, N/R value	OpenPose, MobileNet (Google), InceptionV3 (Google)	Single camera	ACC = ~98% (MobileNet), ~96% (InceptionV3)	Penn Action Dataset

Reference	Exercise	Camera perspectives	Camera distances	PE ⁿ method	Instrument	Key evaluation metrics	Dataset
Chae et al [34]	Squat	Frontal, diagonal	250 cm Kinect (Microsoft), 380 cm webcam	OpenPose, Temporal Conv1D ^f , BiLSTM ^g	Kinect + webcam	ACC=85%	Custom dataset of n=52
Chariar et al [35]	Squat	Frontal	Distance N/R; 120 cm height	MediaPipe (Google), Bi-GRU ^h	2 depth cameras	ACC=94%	Custom dataset of n~50
Zhang et al [36]	Push-up	Frontal, side	N/R	MoveNet; angle-heuristic; Optical flow	N/R	Average F_1 -score: angle-heuristic=0.85 (side > front); pose classification=0.94 (side > front); optical flow=0.79 (front > side)	Kaggle "Push-up Exercise" dataset
Japhne et al [24]	Push-up, squat, and others	N/R, full body visible	200 cm	OpenPose, LSTM ⁱ	Mobile devices	Push-up: ACC=~99%; Squat: ACC=~99%	Custom dataset of n=3
Mercadal-Baudart et al [21]	Squat and others	Frontal; multiangle validation	~300 cm radius; 150 cm height	Detectron2 (Meta Platforms, Inc), Strided Trans-former	Mobile devices	RMSE ^j of joint angles versus VICON ^k (Vicon Motion Systems): <10° for most joints (shin, knee, hip, trunk, and spine), <15° for shoulder and ASIS ^l (notably front squats)	Custom dataset of n=8-12

^aPE: pose estimation.

^bN/R: not reported.

^cACC: accuracy.

^dEJUST-SQUAT-21: Egypt-Japan University squat dataset 2021.

^eMM-Fit: multimodal fitness dataset.

^fConv1D: one-dimensional convolutional neural network.

^gBiLSTM: bidirectional long short-term memory.

^hBi-GRU: bidirectional gated recurrent unit.

ⁱLSTM: long short-term memory.

^jRMSE: root mean square error.

^kVICON: Vicon motion capture system.

^lASIS: anterior superior iliac spine.

Pose Estimation and Processing

Video data were processed using a multistage PE and repetition detection pipeline designed for real-world deployment on the Dotmoovs mobile platform [37]. The pipeline was optimized for low-latency, on-device inference while maintaining sufficient accuracy for exercise recognition.

Model Selection

A lightweight 2D PE model was prioritized to ensure real-time performance on consumer smartphones and without requiring external sensors. Among candidate architectures (eg, PoseNet [Google], BlazePose, and MoveNet), MoveNet was selected due to its computational efficiency, suitability for edge inference, and enhanced generalization to fitness-related movements. MoveNet was trained on the COCO (Common Objects in Context) dataset (Microsoft) and Google's internal Active dataset, which includes annotated yoga, fitness, and dance poses exhibiting substantial motion variability [38-40]. For cloud-based processing, EvoPose2D (Huawei Technologies Co, Ltd) and ViTPose (Microsoft Research Asia) were implemented to achieve higher precision

at the cost of increased computational load. This hybrid configuration balances latency, accuracy, and hardware constraints, supporting both on-device responsiveness and scalable cloud inference [39]. The models were trained on approximately 1 million samples, reserving 10,000 samples each for validation and testing, ensuring robustness and generalizability across diverse movement types.

Training Datasets and Domain Adaptation

To improve robustness for fitness-specific postures, the training pipeline incorporated multiple publicly available datasets: SMART (Sports Motion and Recognition Tasks) [41], LSP (Leeds Sports Pose) Extended [42], Penn Action [43], and MPII Human Pose [44]. Despite these datasets, complex movements, particularly floor-based exercises (eg, push-ups), remained challenging for accurate PE. This motivated the creation of DotPose, a custom internal dataset designed to complement existing datasets and improve the detection of occluded limbs and challenging postures. The combination of DotPose with public datasets further mitigates potential biases related to body types, cultural differences, exercise contexts, and environmental scenarios, thereby

enhancing generalizability across diverse users and real-world conditions.

Pipeline Implementation

The PE outputs (17 keypoints per frame) were processed through a proprietary deep learning module performing two primary tasks:

1. State classification: A lightweight neural network classifies each frame into discrete movement states (eg, squat-down, squat-up, or other).
2. Repetition counting: A Markov-chain-based algorithm tracks temporal transitions between states, triggering a repetition event when a predefined sequence (eg, up-down-up) is detected.

This architecture ensures temporal stability and robustness against intermittent keypoint noise. The pipeline was benchmarked for real-time inference, with MoveNet demonstrating subsecond latency, suitable for interactive mobile feedback. Finally, all pose outputs were exported in comma-separated values format to enable statistical analysis of accuracy and classification metrics across varying camera angles and distances.

Manual Annotation

All video samples were independently annotated by 2 trained raters (an exercise physiologist and a physical therapist), who followed a predefined schema based on standardized exercise movement criteria. Each annotator labeled the exercise type and manually counted the number of valid repetitions per trial. No discrepancies were observed between raters. These annotations were then used to evaluate the accuracy of detection and repetition counting.

Statistical Analysis

All analyses were conducted in RStudio (version 4.5.0; R Foundation for Statistical Computing) using established packages including lme4, lmerTest, emmeans, and performance. Performance metrics were reported in alignment with prior benchmark studies in human activity recognition and AI-based motion analysis [26,45]. The primary outcome was a binary accuracy indicator, coded as 1 when the system's predicted repetition count exactly matched the ground truth, and as 0 otherwise. This was modeled using generalized linear mixed-effects models (GLMMs) with a logit link. The secondary outcome was the mean absolute error (MAE), defined as the absolute difference between predicted and

ground-truth repetition counts per video, modeled using linear mixed effects models (LMMs) with a Gaussian distribution. For descriptive reporting purposes, we also calculated a detection rate for each video, defined as (prediction or ground truth) x100, representing the percentage of repetitions counted correctly. Both models included camera angle and distance as fixed effects, with participant ID modeled as a random intercept to account for intraindividual variability and repeated measures. When the fixed effects were statistically significant ($P<.05$), Tukey-adjusted post hoc contrasts were conducted to compare condition pairs. Odds ratios (ORs) were computed by exponentiating the log odds from the GLMM to facilitate interpretation.

Results

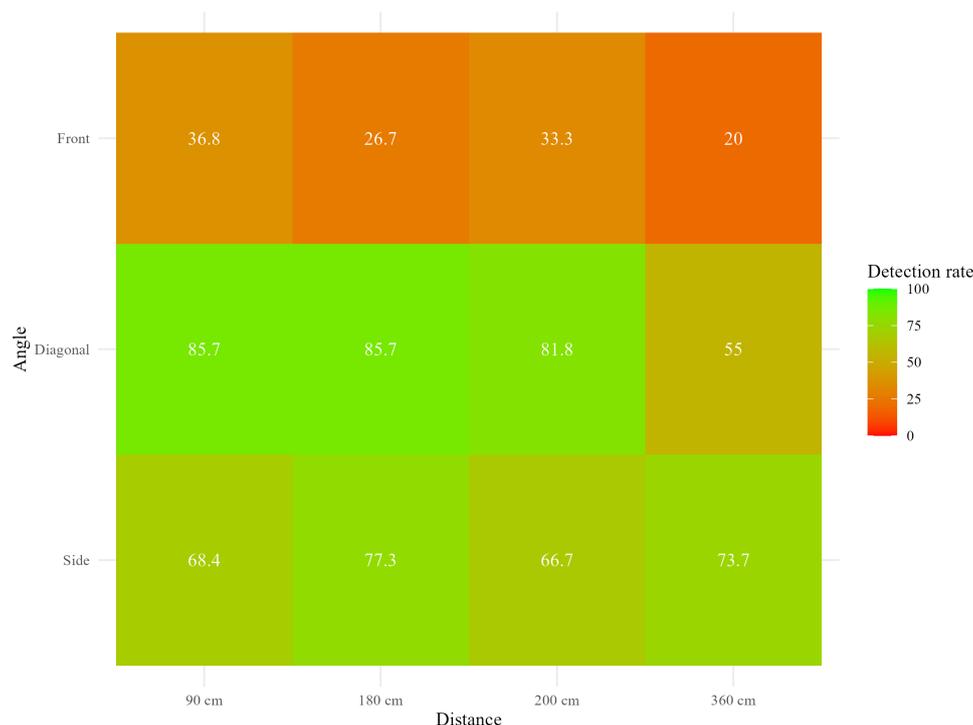
Participant Characteristics

Our dataset comprised 44 healthy university students from the University of Évora, Portugal, including 9 (20.5%) females and 35 (79.5%) males. The participants had a mean age of 20.31 (SD 0.40) years, a mean height of 1.74 (SD 0.09) meters, and a mean body mass of 70 kg (SD 14.72), resulting in a mean BMI of 23.16 kg/m² (SD 0.61).

Based on previous studies assessing AI-based PE for push-up and squat detection [24], we assumed a medium-to-large expected effect size (Cohen $d=0.65$) for power estimation. A post hoc analysis conducted in G*Power 3.1 (Heinrich Heine University Düsseldorf) indicated that the current sample size ($n=44$) provides approximately 0.74 statistical power to detect this effect at a 2-tailed significance level of $\alpha=.05$, suggesting adequate sensitivity for detecting meaningful differences in PE performance across camera conditions.

Push-Up Performance

Table 2 shows detection rates and MAEs for push-ups across camera angles and distances. Overall, the mean detection rate was 61.1% (SD 48.8%), with an average MAE of 1.08 (SD 1.78) repetitions. Detection rates peaked for diagonal views at 90 cm and 180 cm (both 85.7%, SD 35.9%), as illustrated in Figure 1. The corresponding MAEs at these distances were low, at 0.29 (SD 0.90) repetitions. In contrast, the front view at 360 cm showed the lowest detection rate of 20% (SD 41.0%) and the highest MAE of 2.70 repetitions (SD 2.05).

Figure 1. Detection rate of push-ups across distances and angles.**Table 2.** Detection rate and MAE^a for push-ups across camera angles and distances.

Camera angle	Distance (cm)	Detection rate (%), mean (SD)	MAE, mean (SD)
Side	90	68.4 (47.8)	0.842 (1.537)
Side	180	77.3 (42.9)	0.409 (0.796)
Side	200	66.7 (48.3)	1.095 (1.972)
Side	360	73.7 (45.2)	0.789 (1.653)
Diagonal	90	85.7 (35.9)	0.286 (0.902)
Diagonal	180	85.7 (35.9)	0.286 (0.902)
Diagonal	200	81.8 (39.5)	0.227 (0.528)
Diagonal	360	55.0 (51)	1.500 (2.482)
Front	90	36.8 (49.6)	1.684 (1.974)
Front	180	26.7 (45.8)	2.067 (2.017)
Front	200	33.3 (48.8)	1.733 (1.981)
Front	360	20 (41)	2.700 (2.055)

^aMAE: mean absolute error.

The GLMM (Table S1 in [Multimedia Appendix 1](#)) revealed a significant main effect of camera angle on accuracy. Specifically, front views had significantly lower odds of detection (OR=0.17, $P=.03$). However, no significant effects of distance alone were found. Conversely, a statistically significant interaction was found between the diagonal view and the 360 cm distance, which reduced the ability to correctly identify push-ups ($P=.04$).

Post hoc pairwise contrasts were performed to examine differences between the specific camera angles and distances (Table S2 in [Multimedia Appendix 1](#)). These contrasts confirmed that diagonal camera views at 90 cm significantly outperformed frontal views at all distances, including 90 cm (estimate=3.41, $P=.019$), 180 cm (estimate=3.67, $P=.02$), 200 cm (estimate=3.52, $P=.02$), and 360 cm (estimate=4.40, $P<.001$). There were also notable differences in the multiple

mid-range diagonal and frontal configurations: the diagonal 180 cm view outperformed the frontal 180 cm view (estimate=3.64, $P=.02$), 200 cm (estimate=3.50, $P=.03$), as well as 360 cm (estimate=4.35, $P=.002$).

Likewise, the diagonal 200 cm view surpassed the frontal 360 cm view (estimate=3.96, $P=.003$). Additionally, the side views at 180 cm and 360 cm were significantly better than the frontal 360 cm view (estimate=3.586, $P=.008$, and estimate=3.174, $P=.03$, respectively). These results demonstrate that diagonal and mid-range frontal camera placements substantially increase push-up detection accuracy compared to frontal or side views at extreme distances.

Analyses of push-up repetition MAE using LMM were conducted to evaluate the error comparisons in the respective configurations. Although the main effects of angle and distance were not individually significant beyond the

intercept (Table S3 in [Multimedia Appendix 1](#)), Tukey’s post hoc tests (Table S4 in [Multimedia Appendix 1](#)) revealed meaningful differences between specific camera setups. Diagonal angles at mid-range distances consistently reduced counting errors compared with frontal and side angles at extreme distances. Diagonal views at 90 cm and 180 cm, for example, significantly outperformed frontal views at 180 cm and 360 cm (estimates=-1.686 to -2.395, $P=.04$). Diagonal views at 200 cm were also superior to frontal views at 360 cm (estimate=-2.428, $P<.001$). Side views at extreme distances exhibited higher MAE than frontal views at 360 cm (estimate=-1.800, $P<.01$). Overall, these results suggest that diagonal positioning at mid-range distances (approximately 180R-200 cm) minimizes counting errors in push-ups.

Squat Performance

Table 3 shows detection rates and MAEs for squats across camera angles and distances. Overall, the mean detection rate was 61.5% (SD 48.7%), with an average MAE of 1.11 (SD 1.82) repetitions. Detection rates peaked for the diagonal view at 200 cm (95.5%, SD 21.3%), as illustrated in Figure 2. The corresponding MAE at this distance was minimal, at 0.05 (SD 0.21) repetitions. In contrast, the side view at 90 cm showed the lowest detection rate of 0% (SD 0%) as well as the highest MAE of 5 (SD 0) repetitions.

Figure 2. Detection rate of squats across distances and angles.



Table 3. Detection rate and MAE^a for squats across camera angles and distances.

Camera angle	Distance (cm)	Detection rate (%), mean (SD)	MAE, mean (SD)
Side	90	0 (0)	5 (0)
Side	180	28.6 (46.3)	0.714 (0.463)
Side	200	36.4 (49.2)	0.636 (0.492)
Side	360	23.8 (43.6)	3.095 (2.3)
Diagonal	90	45.5 (51)	1.318 (1.673)
Diagonal	180	90.9 (29.4)	0.091 (0.294)
Diagonal	200	95.5 (21.3)	0.045 (0.213)
Diagonal	360	90.9 (29.4)	0.091 (0.294)
Front	90	86.4 (35.1)	0.318 (0.945)
Front	180	81 (40.2)	0.571 (1.248)
Front	200	90 (30.8)	0.4 (1.273)
Front	360	70 (47)	1.1 (1.889)

^aMAE: mean absolute error.

Analyzing the binary accuracy through the GLMM (Table S5 in [Multimedia Appendix 1](#)) demonstrated that the diagonal and frontal angles significantly outperformed the side angle at 90 cm (OR=20.43, $P<.001$, and OR=23.01, $P<.001$, respectively). Detection also significantly improved at mid-range distances of 180 cm (OR=4.36, $P=.002$), 200 cm (OR=7.62, $P<.001$), and 360 cm (OR=3.054, $P=.01$) compared to 90 cm. Post hoc pairwise comparisons (Table S6 in [Multimedia Appendix 1](#)) revealed that the diagonal view at 180 cm significantly outperformed the front view at 90 cm (estimate=2.810, $P=.02$) and the front view at 180 cm (estimate=3.058, $P=.009$). However, it underperformed compared to the front view at 200 cm (estimate=-3.271, $P=.01$) and the front view at 360 cm (estimate=-4.016, $P=.02$). Side views at 180 cm and 360 cm outperformed the front view at 360 cm in correctly detecting push-ups (estimates=2.907 to 3.239, both $P=.04$). The diagonal view at 200 cm had significantly better detection accuracy than the front view at 360 cm (estimate=3.174, $P=.03$). These findings suggest that diagonal mid-range placements generally perform well, but that frontal views at longer distances (200 and 360 cm) can sometimes outperform the diagonal view at 180 cm for squat detection.

For squats, LMM analysis of MAE (Table S7 in [Multimedia Appendix 1](#)) revealed significantly lower errors for diagonal and frontal angles compared with the side (estimates were -3.682 and -4.682, respectively; both $P<.001$). All tested distances (180 cm, 200 cm, and 360 cm) were superior to the 90 cm reference (estimates ranged from -4.364 to -1.910; P values<.001). Several significant interactions were observed. For instance, combining a diagonal angle with 180 cm or 200 cm decreased MAE by 3.06 and 3.09 repetitions, respectively (both $P<.001$), which moderated the improvements to the main effects expected from angle and distance independently. The front \times 180 cm interaction (estimate=4.547, $P<.001$) and the front \times 200 cm interaction (estimate=4.460, $P<.001$) exhibited similar patterns, indicating smaller improvements than the sum of the main effects predicted. In the model, only the diagonal \times 360 cm interaction was nonsignificant ($P=.16$). Several significant post hoc pairwise comparisons were identified (Table S8 in [Multimedia Appendix 1](#)). These confirmed that diagonal and frontal mid-range placements consistently minimized error. Diagonal at 200 cm produced the largest reduction (estimate=-4.96 repetitions, $P<.001$), similar to diagonal at 180 cm (estimate=-4.91 repetitions, $P<.001$). Most configurations substantially reduced MAE compared to the 90 cm side, but the magnitude of improvement depended on the specific combination of angle and distance. This highlights notable interaction effects.

Discussion

Main Findings

This study demonstrates that smartphone camera angle and distance critically affect the accuracy of PE models for detecting and counting push-ups and squats. Consistent with our results, camera placements at intermediate distances

(180-200 cm) combined with oblique (diagonal) or frontal views generally yielded the highest detection rates and lowest counting errors. In contrast, very close-range setups (90 cm) and long-distance frontal views (360 cm) often showed reduced performance. For push-ups, diagonal views between 90 and 200 cm outperformed frontal angles, with the diagonal view at 200 cm position achieving the lowest MAE. For squats, diagonal and frontal views at 180 to 200 cm significantly outperformed side views, with diagonal 180 cm, diagonal 200 cm, and front 200 cm producing the smallest MAEs and the largest error reductions in post hoc tests. Although front 360 cm occasionally approached the accuracy of diagonal 180 cm, most mid-range configurations substantially outperformed both close-range and long-distance side views. These findings provide direct evidence that mid-range diagonal and frontal camera configurations optimize PE performance, informing best practices for smartphone-based exercise monitoring.

Comparison to Prior Work

Unlike prior lab-based, multicamera studies [11,31], our study provides practical, configuration-specific guidance for monocular smartphone setups. This approach enhances ecological validity by closely mimicking typical home or gym environments where single-camera devices are commonly used. While [17] demonstrated that full-body visibility at 300 cm supports accurate gait tracking, we show that diagonal views at 180 to 200 cm consistently optimize visibility and accuracy for dynamic strength exercises. Views at 360 cm, however, yield variable and often low detection rates (push-ups: 20%-73.7%; squats: 23%-90.9%), depending on the angle [46], evaluated smartphone-based distance estimation between 100 and 300 cm, the trade-offs between spatial accuracy, and usability. Our findings confirm that mid-range diagonal placements are optimal for PE performance, likely due to differences in exercise posture (horizontal for push-ups vs upright for squats).

Prior studies have evidenced smartphone apps using TensorFlow Lite and COCO-trained models for exercise counting [10,47], but have not comprehensively examined spatial configuration effects. Our results address this gap, complementing [48] who emphasized viewpoint in PE data acquisition [9], advocating for accessible PE tools beyond labs, and [12] who discussed challenges of motion tracking in naturalistic settings. By offering concrete camera setup recommendations, this work contributes actionable insights for mHealth, home fitness, telerehabilitation, sports performance, and clinical decision-making contexts.

Practical Implications for mHealth

Positioning the smartphone camera diagonally ($\sim 45^\circ$) at 180 to 200 cm significantly enhances PE accuracy without additional hardware. Mobile apps can integrate augmented reality or setup guides to assist users in achieving optimal device placement. Accurate repetition counting supports load monitoring, fatigue tracking, and muscular endurance assessment in unsupervised environments [29,49,50]. Combining spatial optimization with adaptive feedback and personalized experiences may further improve tracking

reliability and user engagement, consistent with evidence from behavior change and human-computer interaction research [51,52]. This integrated approach presents a promising path for scalable, user-friendly mHealth exercise platforms.

Limitations

This study has several limitations. First, PE performance was evaluated using manual annotations rather than a gold-standard motion capture system, which may introduce variability. Second, the sample comprised healthy young adults in a controlled laboratory environment, limiting generalizability to clinical populations, older adults, or real-world contexts with variable lighting and backgrounds. Third, while only 1 PE pipeline was tested, integrating multiple public datasets and the custom DotPose dataset helped mitigate biases related to scene composition, body type, exercise variety, and environmental scenarios. Furthermore, comparisons with other open-source PE models (eg, OpenPose, HRNet [Microsoft Research Asia], and BlazePose) are limited by variations in datasets, computational demands, and architectural design [12,35,53]. Finally, lightweight 2D models, such as MoveNet, enable near-real-time, on-device inference. These models balance speed, portability, and user engagement with the fine-grained accuracy achievable by heavier, cloud-reliant models [13]. This hybrid mobile-cloud configuration supports the robust evaluation of push-ups and squats across multiple camera setups. Nevertheless, findings may not generalize to other PE architectures, populations, or unstructured environments.

Future Directions

Future work should explore multimodal pipelines, diverse participant groups, and variable environmental conditions (eg,

low-light and high-contrast conditions) to enhance robustness, applicability, and real-world relevance of mobile PE systems. While repetition counting is a fundamental first step, future mHealth systems should also evaluate movement quality. This should encompass compensatory strategies and fatigue-related adaptations [54-56]. Emerging real-time PE technologies that optimize sensor or camera placement, combined with adaptive feedback responsive to user performance, offer opportunities to enhance movement accuracy and user experience in unsupervised settings [28,48,57,58]. Longitudinal studies are needed to assess the integration of spatial setup guidance, personalization, and real-time feedback for improved usability, engagement, and long-term clinical or fitness outcomes.

Conclusions

Camera angle and distance significantly affect the accuracy of PE systems for exercise detection. For optimal performance, smartphone cameras should be positioned at mid-range distances (180-200 cm) with diagonal or frontal views. For push-ups, diagonal views are preferred, while for squats, both diagonal and frontal views perform well. Conversely, close-range setups (90 cm) and long-distance frontal views (360 cm) substantially reduce detection and counting accuracy. These findings provide actionable guidance for developing scalable, accurate, and user-friendly mHealth exercise tracking platforms.

Acknowledgments

The authors would like to thank the students who participated in this study for their time and commitment. They would also like to acknowledge the University of Évora for its support in facilitating the research. During the preparation of the manuscript, ChatGPT was used to suggest grammatical edits. The authors reviewed and edited all content and took full responsibility for the final version.

Funding

This work was financially supported by Project Blockchain.PT—Decentralize Portugal with Blockchain Agenda (Project 51), WP4 Sports, Leisure, and Culture, Call number 02/C05-i01.01/2022, funded by the Portuguese Recovery and Resilience Program, the Portuguese Republic, and the European Union under the framework of the Next Generation European Union Program. The funders had no involvement in the study design, data collection, analysis, interpretation of results, or the writing of the manuscript.

Data Availability

The study leveraged established benchmark datasets for human pose estimation, including COCO (Common Objects in Context) [38], SMART (Sports Motion and Recognition Tasks) [41], LSP (Leeds Sports Pose) Extended [42], Penn Action [43], and MPII Human Pose [44]. Additionally, the data collected from the 44 participants are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: EO, SF, APG

Data curation: EO, GV

Formal analysis: EO, FG-F

Investigation: EO, SF, JP, PP

Methodology: EO, GV
Project administration: SA
Supervision: APG
Writing – original draft: EO
Writing – review & editing: SF, APG, GV, JP, PP, FG-F, SA
All authors have reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Detailed results of generalized linear mixed-effects models, linear mixed-effects models, and post-hoc pairwise comparisons for push-ups and squats.

[\[DOC File \(Microsoft Word File\), 6016 KB-Multimedia Appendix 1\]](#)

References

1. Dallinga J, Janssen M, van der Werf J, Walravens R, Vos S, Deutekom M. Analysis of the features important for the effectiveness of physical activity-related apps for recreational sports: expert panel approach. *JMIR mHealth uHealth*. Jun 18, 2018;6(6):e143. [doi: [10.2196/mhealth.9459](https://doi.org/10.2196/mhealth.9459)] [Medline: [29914863](https://pubmed.ncbi.nlm.nih.gov/29914863/)]
2. Kohl HW 3rd, Craig CL, Lambert EV, et al. The pandemic of physical inactivity: global action for public health. *Lancet*. Jul 2012;380(9838):294-305. [doi: [10.1016/S0140-6736\(12\)60898-8](https://doi.org/10.1016/S0140-6736(12)60898-8)] [Medline: [23041199](https://pubmed.ncbi.nlm.nih.gov/23041199/)]
3. Kardan M, Jung A, Iqbal M, Keshtkar S, Geidl W, Pfeifer K. Efficacy of digital interventions on physical activity promotion in individuals with noncommunicable diseases: an overview of systematic reviews. *BMC Digit Health*. 2024;2(1):40. [doi: [10.1186/s44247-024-00097-6](https://doi.org/10.1186/s44247-024-00097-6)]
4. Physical activity. World Health Organization. URL: <https://www.who.int/news-room/fact-sheets/detail/physical-activity> [Accessed 2026-02-10]
5. Hallal PC, Andersen LB, Bull FC, et al. Global physical activity levels: surveillance progress, pitfalls, and prospects. *Lancet*. Jul 21, 2012;380(9838):247-257. [doi: [10.1016/S0140-6736\(12\)60646-1](https://doi.org/10.1016/S0140-6736(12)60646-1)] [Medline: [22818937](https://pubmed.ncbi.nlm.nih.gov/22818937/)]
6. Buecker S, Simacek T, Ingwersen B, Terwiel S, Simonsmeier BA. Physical activity and subjective well-being in healthy individuals: a meta-analytic review. *Health Psychol Rev*. Dec 2021;15(4):574-592. [doi: [10.1080/17437199.2020.1760728](https://doi.org/10.1080/17437199.2020.1760728)] [Medline: [32452716](https://pubmed.ncbi.nlm.nih.gov/32452716/)]
7. Núñez-Cortés R, Salazar-Méndez J, Nijs J. Physical activity as a central pillar of lifestyle modification in the management of chronic musculoskeletal pain: a narrative review. *J Funct Morphol Kinesiol*. May 20, 2025;10(2):183. [doi: [10.3390/jfmk10020183](https://doi.org/10.3390/jfmk10020183)] [Medline: [40407467](https://pubmed.ncbi.nlm.nih.gov/40407467/)]
8. Avogaro A, Cunico F, Rosenhahn B, Setti F. Markerless human pose estimation for biomedical applications: a survey. *Front Comput Sci*. 2023;5:1153160. [doi: [10.3389/fcomp.2023.1153160](https://doi.org/10.3389/fcomp.2023.1153160)]
9. Stenum J, Cherry-Allen KM, Pyles CO, Reetzke RD, Vignos MF, Roemmich RT. Applications of pose estimation in human health and performance across the lifespan. *Sensors (Basel)*. Nov 3, 2021;21(21):7315. [doi: [10.3390/s21217315](https://doi.org/10.3390/s21217315)] [Medline: [34770620](https://pubmed.ncbi.nlm.nih.gov/34770620/)]
10. El-Rajab I, Klotzbier TJ, Korbus H, Schott N. Camera-based mobile applications for movement screening in healthy adults: a systematic review. *Front Sports Act Living*. 2025;7:1531050. [doi: [10.3389/fspor.2025.1531050](https://doi.org/10.3389/fspor.2025.1531050)] [Medline: [40416048](https://pubmed.ncbi.nlm.nih.gov/40416048/)]
11. Luangaphirom T, Lueprasert S, Kaewwichit P, Boonphotosiri S, Burapasikarin T, Siriborvornratanakul T. Real-time weight training counting and correction using MediaPipe. *Adv in Comp Int*. Jun 2024;4(2):3. [doi: [10.1007/s43674-024-00070-w](https://doi.org/10.1007/s43674-024-00070-w)]
12. Roggio F, Trovato B, Sortino M, Musumeci G. A comprehensive analysis of the machine learning pose estimation models used in human movement and posture analyses: a narrative review. *Heliyon*. 2024;10(21):e39977. [doi: [10.1016/j.heliyon.2024.e39977](https://doi.org/10.1016/j.heliyon.2024.e39977)] [Medline: [39553598](https://pubmed.ncbi.nlm.nih.gov/39553598/)]
13. Phosanarack M, Wallard L, Lepreux S, Kolski C, Avril E. Smartphone exergames with real-time markerless motion capture: challenges and trade-offs. arXiv. Preprint posted online on Jul 9, 2025. [doi: [10.48550/ARXIV.2507.06669](https://doi.org/10.48550/ARXIV.2507.06669)]
14. Badiola-Bengoia A, Mendez-Zorrilla A. A systematic review of the application of camera-based human pose estimation in the field of sport and physical exercise. *Sensors (Basel)*. Sep 7, 2021;21(18):5996. [doi: [10.3390/s21185996](https://doi.org/10.3390/s21185996)] [Medline: [34577204](https://pubmed.ncbi.nlm.nih.gov/34577204/)]
15. Cornman HL, Stenum J, Roemmich RT. Video-based quantification of human movement frequency using pose estimation: a pilot study. *PLoS One*. 2021;16(12):e0261450. [doi: [10.1371/journal.pone.0261450](https://doi.org/10.1371/journal.pone.0261450)] [Medline: [34929012](https://pubmed.ncbi.nlm.nih.gov/34929012/)]

16. Jeon H, Kim D, Kim J. Human motion assessment on mobile devices. Presented at: 2021 International Conference on Information and Communication Technology Convergence (ICTC); Oct 20-22, 2021; Jeju Island, Republic of Korea. [doi: [10.1109/ICTC52510.2021.9621114](https://doi.org/10.1109/ICTC52510.2021.9621114)]
17. Yang J, Park K. Improving gait analysis techniques with markerless pose estimation based on smartphone location. *Bioengineering (Basel)*. Jan 30, 2024;11(2):141. [doi: [10.3390/bioengineering11020141](https://doi.org/10.3390/bioengineering11020141)] [Medline: [38391625](https://pubmed.ncbi.nlm.nih.gov/38391625/)]
18. Chen Z, Liao K, Bishop C, Bian C, Li Y. Validity of three commercial devices for recording movement velocity in the Bulgarian split squat. *J Hum Kinet*. 2024;95:161-171. [doi: [10.5114/jhk/189365](https://doi.org/10.5114/jhk/189365)] [Medline: [39944971](https://pubmed.ncbi.nlm.nih.gov/39944971/)]
19. Stamm O, Heimann-Steinert A. Accuracy of monocular two-dimensional pose estimation compared with a reference standard for kinematic multiview analysis: validation study. *JMIR mHealth uHealth*. Dec 21, 2020;8(12):e19608. [doi: [10.2196/19608](https://doi.org/10.2196/19608)] [Medline: [33346739](https://pubmed.ncbi.nlm.nih.gov/33346739/)]
20. van den Hoorn W, Lavaill M, Cutbush K, Gupta A, Kerr G. Comparison of shoulder range of motion quantified with mobile phone video-based skeletal tracking and 3D motion capture-preliminary study. *Sensors (Basel)*. Jan 15, 2024;24(2):534. [doi: [10.3390/s24020534](https://doi.org/10.3390/s24020534)] [Medline: [38257626](https://pubmed.ncbi.nlm.nih.gov/38257626/)]
21. Mercadal-Baudart C, Liu CJ, Farrell G, et al. Exercise quantification from single camera view markerless 3D pose estimation. *Heliyon*. Mar 30, 2024;10(6):e27596. [doi: [10.1016/j.heliyon.2024.e27596](https://doi.org/10.1016/j.heliyon.2024.e27596)] [Medline: [38510055](https://pubmed.ncbi.nlm.nih.gov/38510055/)]
22. Ferreira B, Ferreira PM, Pinheiro G, et al. Deep learning approaches for workout repetition counting and validation. *Pattern Recognit Lett*. Nov 2021;151:259-266. [doi: [10.1016/j.patrec.2021.09.006](https://doi.org/10.1016/j.patrec.2021.09.006)]
23. Rao P, Asha CS, Rao PR. Real-time posture correction of squat exercise: a deep learning approach for performance analysis and error correction. *IEEE Access*. 2025;13:39557-39571. [doi: [10.1109/ACCESS.2025.3545207](https://doi.org/10.1109/ACCESS.2025.3545207)]
24. Japhne F, Janada K, Theodorus A, Chowanda A. Fitcam: detecting and counting repetitive exercises with deep learning. *J Big Data*. 2024;11(1):101. [doi: [10.1186/s40537-024-00915-8](https://doi.org/10.1186/s40537-024-00915-8)]
25. Sinclair A, Kautai K, Shahamiri SR. Pūioio: on-device real-time smartphone-based automated exercise repetition counting system. arXiv. Preprint posted online on Jul 22, 2023. [doi: [10.48550/ARXIV.2308.02420](https://doi.org/10.48550/ARXIV.2308.02420)]
26. Fisal NR, Fathalla A, Elmanakhly DA, Salah A. Reported challenges in deep learning-based human pose estimation: a systematic review. *IEEE Access*. 2025;13:80520-80539. [doi: [10.1109/ACCESS.2025.3567337](https://doi.org/10.1109/ACCESS.2025.3567337)]
27. Fanton M, Harari Y, Giffhorn M, et al. Validation of Amazon Halo Movement: a smartphone camera-based assessment of movement health. *NPJ Digit Med*. Sep 6, 2022;5(1):134. [doi: [10.1038/s41746-022-00684-9](https://doi.org/10.1038/s41746-022-00684-9)] [Medline: [36065060](https://pubmed.ncbi.nlm.nih.gov/36065060/)]
28. Biebl JT, Rykala M, Strobel M, et al. App-based feedback for rehabilitation exercise correction in patients with knee or hip osteoarthritis: prospective cohort study. *J Med Internet Res*. Jul 13, 2021;23(7):e26658. [doi: [10.2196/26658](https://doi.org/10.2196/26658)] [Medline: [34255677](https://pubmed.ncbi.nlm.nih.gov/34255677/)]
29. American College of Sports Medicine. Liguori G, Feito Y, Fountaine C, Roy BA, editors. *ACSM's Guidelines for Exercise Testing and Prescription*. 11th ed. Wolters Kluwer; 2021. URL: <https://books.google.co.in/books?id=yjibzQEACAAJ> [Accessed 2026-02-10] ISBN: 9781975150198
30. Miranda GHL. *Biofotogrametria Para Fisioterapeutas* [Book in Portuguese]. 1st ed. Andreoli; 2014. URL: <https://www.indicalivros.com/livros/biofotogrametria-para-fisioterapeutas-geraldo-henrique-lobes-miranda> [Accessed 2026-02-10] ISBN: 9788560416394
31. Park HJ, Baek JW, Kim JH. Imagery based parametric classification of correct and incorrect motion for push-up counter using OpenPose. Presented at: 2020 IEEE 16th International Conference on Automation Science and Engineering (CASE); Aug 20-21, 2020; Hong Kong. [doi: [10.1109/CASE48305.2020.9216833](https://doi.org/10.1109/CASE48305.2020.9216833)]
32. Youssef F, Zaky AB, Gomaa W. Analysis of the squat exercise from visual data. Presented at: 19th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2022); Jul 14-16, 2022; Lisbon, Portugal. [doi: [10.5220/0011347900003271](https://doi.org/10.5220/0011347900003271)]
33. Hande T, Kakirwar B, Bharadwaja AV, Kshirsagar P, Gupta A, Vijayakumar P. Correction and estimation of workout postures with pose estimation using AI. Presented at: 2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE); Jan 27-28, 2023; Bengaluru, India. [doi: [10.1109/IITCEE57236.2023.10090463](https://doi.org/10.1109/IITCEE57236.2023.10090463)]
34. Chae HJ, Kim JB, Park G, O'Sullivan DM, Seo J, Park JJ. An artificial intelligence exercise coaching mobile app: development and randomized controlled trial to verify its effectiveness in posture correction. *Interact J Med Res*. Sep 12, 2023;12:e37604. [doi: [10.2196/37604](https://doi.org/10.2196/37604)] [Medline: [37698913](https://pubmed.ncbi.nlm.nih.gov/37698913/)]
35. Chariar M, Rao S, Irani A, Suresh S, Asha CS. AI Trainer: autoencoder based approach for squat analysis and correction. *IEEE Access*. 2023;11:107135-107149. [doi: [10.1109/ACCESS.2023.3316009](https://doi.org/10.1109/ACCESS.2023.3316009)]
36. Zhang X, Han SZH, Lim KYT. Designing and prototyping of AI-based real-time mobile detectors for calisthenic push-up exercise. Presented at: International Conference on Health and Social Care Information Systems and Technologies (HCist 2023); Nov 8-10, 2023; Porto, Portugal. [doi: [10.1016/j.procs.2024.06.192](https://doi.org/10.1016/j.procs.2024.06.192)]
37. Dotmoovs. URL: <https://www.dotmoovs.com/> [Accessed 2026-02-10]

38. Lin TY, Maire M, Belongie S, et al. Microsoft COCO: common objects in context. Presented at: 13th European Conference on Computer Vision (ECCV 2014); Sep 6-12, 2014; Zurich, Switzerland. [doi: [10.1007/978-3-319-10602-1_48](https://doi.org/10.1007/978-3-319-10602-1_48)]
39. Votel R, Li N. Next-generation pose detection with MoveNet and TensorFlow.js. TensorFlow Blog. May 17, 2021. URL: <https://blog.tensorflow.org/2021/05/next-generation-pose-detection-with-movenet-and-tensorflowjs.html> [Accessed 2025-10-28]
40. Kumar P, Chauhan S, Awasthi LK. Human pose estimation using deep learning: review, methodologies, progress and future research directions. *Int J Multimed Info Retr.* 2022;11(4):489-521. [doi: [10.1007/s13735-022-00261-6](https://doi.org/10.1007/s13735-022-00261-6)]
41. Chen X, Pang A, Yang W, Ma Y, Xu L, Yu J. SportsCap: monocular 3D human motion capture and fine-grained understanding in challenging sports videos. *Int J Comput Vis.* 2021;129(10):2846-2864. [doi: [10.1007/s11263-021-01486-4](https://doi.org/10.1007/s11263-021-01486-4)]
42. Johnson S, Everingham M. Learning effective human pose estimation from inaccurate annotation. Presented at: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jun 20-25, 2011; Colorado Springs, CO, USA. [doi: [10.1109/CVPR.2011.5995318](https://doi.org/10.1109/CVPR.2011.5995318)]
43. Zhang W, Zhu M, Derpanis KG. From actemes to action: a strongly-supervised representation for detailed action understanding. Presented at: Proceedings of the IEEE International Conference on Computer Vision (ICCV) 2013; Dec 1-8, 2013; Sydney, Australia. [doi: [10.1109/ICCV.2013.280](https://doi.org/10.1109/ICCV.2013.280)]
44. Andriluka M, Pishchulin L, Gehler P, Schiele B. 2D human pose estimation: new benchmark and state of the art analysis. Presented at: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jun 23-28, 2014; Columbus, Ohio, USA. [doi: [10.1109/CVPR.2014.471](https://doi.org/10.1109/CVPR.2014.471)]
45. Dibenedetto G, Sotiropoulos S, Polignano M, Cavallo G, Lops P. Comparing human pose estimation through deep learning approaches: an overview. *Comput Vis Image Underst.* Feb 2025;252:104297. [doi: [10.1016/j.cviu.2025.104297](https://doi.org/10.1016/j.cviu.2025.104297)]
46. Hamilton-Fletcher G, Liu M, Sheng D, et al. Accuracy and usability of smartphone-based distance estimation approaches for visual assistive technology development. *IEEE Open J Eng Med Biol.* 2024;5:54-58. [doi: [10.1109/OJEMB.2024.3358562](https://doi.org/10.1109/OJEMB.2024.3358562)] [Medline: [38487094](https://pubmed.ncbi.nlm.nih.gov/38487094/)]
47. Jeon H, Yoon Y, Kim D. Lightweight 2D human pose estimation for fitness coaching system. Presented at: 2021 36th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC); Jun 27-30, 2021; Jeju, Korea (South). [doi: [10.1109/ITC-CSCC52171.2021.9501458](https://doi.org/10.1109/ITC-CSCC52171.2021.9501458)]
48. Konak O, Wischmann A, van De Water R, Arnrich B. A real-time human pose estimation approach for optimal sensor placement in sensor-based human activity recognition. Presented at: 8th International Workshop on Sensor-Based Activity Recognition and Artificial Intelligence (iWOAR 2023); Sep 21-22, 2023; Lübeck, Germany. [doi: [10.1145/3615834.3615848](https://doi.org/10.1145/3615834.3615848)]
49. Ebben WP, Feldmann CR, Dayne A, et al. Using squat testing to predict training loads for the deadlift, lunge, step-up, and leg extension exercises. *J Strength Cond Res.* Nov 2008;22(6):1947-1949. [doi: [10.1519/JSC.0b013e31818747c9](https://doi.org/10.1519/JSC.0b013e31818747c9)] [Medline: [18978614](https://pubmed.ncbi.nlm.nih.gov/18978614/)]
50. Lin II, Chen YL, Chuang LL. Test-retest reliability of home-based fitness assessments using a mobile app (R Plus Health) in healthy adults: prospective quantitative study. *JMIR Form Res.* Dec 8, 2021;5(12):e28040. [doi: [10.2196/28040](https://doi.org/10.2196/28040)] [Medline: [34657835](https://pubmed.ncbi.nlm.nih.gov/34657835/)]
51. Baretta D, Perski O, Steca P. Exploring users' experiences of the uptake and adoption of physical activity apps: longitudinal qualitative study. *JMIR mhealth uHealth.* Feb 8, 2019;7(2):e11636. [doi: [10.2196/11636](https://doi.org/10.2196/11636)] [Medline: [30735143](https://pubmed.ncbi.nlm.nih.gov/30735143/)]
52. Li X, Li Y. Sports training strategies and interactive control methods based on neural network models. *Comput Intell Neurosci.* 2022;2022:7624578. [doi: [10.1155/2022/7624578](https://doi.org/10.1155/2022/7624578)] [Medline: [35295278](https://pubmed.ncbi.nlm.nih.gov/35295278/)]
53. Edriss S, Romagnoli C, Caprioli L, Bonaiuto V, Padua E, Annino G. Commercial vision sensors and AI-based pose estimation frameworks for markerless motion analysis in sports and exercises: a mini review. *Front Physiol.* 2025;16:1649330. [doi: [10.3389/fphys.2025.1649330](https://doi.org/10.3389/fphys.2025.1649330)] [Medline: [40873758](https://pubmed.ncbi.nlm.nih.gov/40873758/)]
54. Liao Y, Vakanski A, Xian M. A deep learning framework for assessing physical rehabilitation exercises. *IEEE Trans Neural Syst Rehabil Eng.* Feb 2020;28(2):468-477. [doi: [10.1109/TNSRE.2020.2966249](https://doi.org/10.1109/TNSRE.2020.2966249)] [Medline: [31940544](https://pubmed.ncbi.nlm.nih.gov/31940544/)]
55. Bauer CM, Kankaanpää MJ, Meichtry A, Rissanen SM, Suni JH. Efficacy of six months neuromuscular exercise on lumbar movement variability - a randomized controlled trial. *J Electromyogr Kinesiol.* Oct 2019;48:84-93. [doi: [10.1016/j.jelekin.2019.06.008](https://doi.org/10.1016/j.jelekin.2019.06.008)] [Medline: [31252284](https://pubmed.ncbi.nlm.nih.gov/31252284/)]
56. Woo Y, Jeong H. Exercise assessment based on human pose estimation and relative phase for real-time remote exercise system. *IEEE Access.* 2025;13:53203-53213. [doi: [10.1109/ACCESS.2025.3551834](https://doi.org/10.1109/ACCESS.2025.3551834)]
57. Tharatipyakul A, Srikaewsiew T, Pongnumkul S. Deep learning-based human body pose estimation in providing feedback for physical movement: a review. *Heliyon.* Aug 26, 2024;10(17):e36589. [doi: [10.1016/j.heliyon.2024.e36589](https://doi.org/10.1016/j.heliyon.2024.e36589)] [Medline: [39281455](https://pubmed.ncbi.nlm.nih.gov/39281455/)]

58. Elsayed H, Hoffmann P, Günther S, et al. CameraReady: assessing the influence of display types and visualizations on posture guidance. Presented at: DIS '21: Proceedings of the 2021 ACM Designing Interactive Systems Conference; Jun 28 to Jul 2, 2021. [doi: [10.1145/3461778.3462026](https://doi.org/10.1145/3461778.3462026)]

Abbreviations

AI: artificial intelligence
COCO: Common Objects in Context
GLMM: generalized linear mixed model
LMM: linear mixed effects model
LSP: Leeds Sports Pose
MAE: mean absolute error
mHealth: mobile health
OR: odds ratio
PE: pose estimation
SMART: Sports Motion and Recognition Tasks

Edited by Lorraine Buis; peer-reviewed by Luis Alberto Morales-Rosales, Takuya Sasatani; submitted 14.Aug.2025; final revised version received 04.Nov.2025; accepted 27.Jan.2026; published 05.Mar.2026

Please cite as:

*Oliosí E, Ferreira S, Giordano AP, Viveiros G, Parraca J, Pereira P, Guede-Fernández F, Azevedo S
Evaluation of Smartphone Camera Positioning on Artificial Intelligence Pose Estimation Accuracy for Exercise Detection:
Observational Study
JMIR Mhealth Uhealth 2026;14:e82412
URL: <https://mhealth.jmir.org/2026/1/e82412>
doi: [10.2196/82412](https://doi.org/10.2196/82412)*

© Eduarda Oliosí, Soraia Ferreira, Ana Paula Giordano, Guilherme Viveiros, José Parraca, Paulo Pereira, Federico Guede-Fernández, Salomé Azevedo. Originally published in JMIR mHealth and uHealth (<https://mhealth.jmir.org>), 05.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <https://mhealth.jmir.org/>, as well as this copyright and license information must be included.