A webcam-based machine learning approach for the three-dimensional range of motion evaluation

Xiaoye Michael Wang[1]

Derek T. Smith[2]

Qin Zhu[2]

Draft version: November 14, 2022. This is an unpublished working paper. Contents may change prior to publication.

[1] Department of Kinesiology, University of Toronto, Toronto, ON, M5S 2W6, Canada
[2] Division of Kinesiology and Health, University of Wyoming, Laramie, WY, 82071, USA

[*]Corresponding author: X.M. Wang, michaelwxy.wang@utoronto.ca

Please cite as:
Wang, XM, Smith, DT, & Zhu, Q (2022). A machine learning-based approach for the measurement of three-dimensional full-body range of motion evaluation. SportRxiv.

Abstract

Joint range of motion (ROM) is an important quantitative measure for physical therapy. Commonly relying on a goniometer, accurate and reliable ROM measurement requires extensive training and practice. This, in turn, imposes a significant barrier for those who have limited in-person access to healthcare. The current study presents and evaluates an alternative machine learning-based ROM evaluation method that could be remotely accessed via a webcam. To evaluate its reliability, the ROM measurements for a diverse set of joints (neck, spine, and upper and lower extremities) derived using this method were compared to those obtained from a state-of-the-art marker-based, optical motion capture system. Results showed that the webcam-based solution provides high test-retest reliability and inter-rater reliability at a fraction of the cost of the marker-based system. More importantly, the machine-learning-based method has been shown to be more consistent in tracking joint positions during movements, making it more reliable than the optical motion capture system. The proposed webcam-based ROM evaluation method could be easily adapted for clinical practice and shows tremendous potential for the tele-implementation of physical therapy and rehabilitation.

Keywords: Range of motion, Physical therapy, Rehabilitation, Machine learning, Computer vision, Pose Estimation

1. Introduction

The recent COVID-19 Pandemic has accentuated the need of improving access to healthcare services for vulnerable populations. To improve accessibility, various medical fields have started transitioning from traditional healthcare to telehealth. Telehealth leverages the convenience of personal electronic devices to enable the remote delivery of healthcare services, such as through telephone consultation and videoconferencing. As a result of the booming personal electronic device market, telehealth's popularity has been consistently increasing over the past decade, especially during the COVID-19 Pandemic.[1] Compared to traditional in-person care, telehealth has been shown to be equivalent or even more effective in terms of clinical effectiveness.[2,3] Unfortunately, the main bottleneck for telehealth is devising objective outcome measures that are equally accessible.

In the context of physical therapy and rehabilitation, joint range of motion (ROM) is a widely used outcome measure. One of the most common ways to measure joint ROM is through a handheld goniometer.[4–8] ROM measurement using a goniometer requires certified physicians or physical therapists who had ample training and practice in using the device. This imposes challenges to members of rural and remote communities, where access to healthcare and trained professionals could be limited. Furthermore, although the goniometer has been treated as a gold standard for ROM evaluation,[9,10] its precision and inter-tester reliability could be low due to human errors and procedural inconsistency.[11–13] As a result, there is an exigency to devise an alternative ROM evaluation protocol that could address the accessibility, accuracy, and reliability issues related to the goniometer-based ROM evaluation method.

Various non-traditional solutions for ROM evaluation have been proposed over the years, such as digital photography,[9,14] photogrammetry,[10] convolutional neural networks (CNN),[15] and

even optical motion capture (MoCap).[16] MoCap systems, such as OptiTrack (NaturalPoint, Corvallis, OR, USA) and Vicon (Oxford metrics, UK), have been widely used in biomechanics and sports science studies as they demonstrate high accuracy (± 0.10 mm) and reliability at a high sampling frequency (up to 1000 FPS).[17,18] However, despite the advantages that MoCap systems offer, their hefty price tag and complex setup present a noticeable barrier of entry for most people. In recent years, the machine learning-based human pose estimation has been seen significant development.[19–21] These algorithms use various deep neural network architectures to identify, recover, and track the 2D or 3D locations of key joints of a human actor across multiple frames using a single image stream.[22] Capitalizing on the advances in computer vision, Google Research presented a framework called *MediaPipe* that offers developers and practitioners a free access to state-of-the-art machine learning solutions to various computer vision problems,[23] including the image-based real-time three-dimensional (3D) pose-estimation.

The pose-estimation component of the MediaPipe framework, called *BlazePose*,[24] uses a CNN architecture that combines a lightweight pose detection network with a pose prediction network. The pose detection network first detects any person presented in a single frame while the prediction network tracks the person across subsequent frames. This design allows the machine learning model to consistently track 33 body landmarks in real-time at over 30 frames per second using only a single stream RGB video, such as through a webcam. More importantly, the architecture behind BlazePose is relatively lightweight in terms of the computational resources that it requires. As a result, BlazePose could even be deployed to a web browser with images captured from a webcam, allowing anyone to access its features from a computer with a webcam. Due to its low computational load and high accessibility, BlazePose has been used by researchers to develop different applications, such as in the context of sports for exercise

abnormality detection,[25] yoga training,[26] postural disorder monitoring for Parkinson's patients,[27] and spinal dysfunction risk estimation.[28]

The current study proposed an alternative ROM evaluation method that leverages the power of computer vision algorithms. Using a single webcam and the output from BlazePose, 3D positions of various joints were recorded and used to estimate their corresponding ROM angles. Results obtained using this lightweight machine learning solution were compared to those produced by a state-of-the-art MoCap system (OptiTrack). Additionally, unlike previous studies[9,10,14–16] that only focus on a single joint, the current study evaluated the reliability of ROM measurements using a full-body multi-joint model. Results showed that the proposed solution not only offers an alternative to the traditional ROM evaluation methods with a noticeably lowered barrier of access, but also provides high intra- and inter-rater reliability necessary for practical usage.

2. Materials and Methods

2.1 Participants

Twenty-five adults (12 males, 13 females) from the University of Wyoming community volunteered in this study. This study was approved by the University of Wyoming Institutional Review Board (IRB). All participants provided their written informed consent prior to the study.

2.2 Data Acquisition

Data acquisition was performed using a desktop computer with an Intel Core i9 11th Gen CPU, Nvidia RTX 3090 graphics card, and 64 GB of RAM. To derive the range of motion (ROM) angles for different joints, joint trajectories of various ROM evaluation movements were simultaneously recorded using a webcam-based, real-time pose estimation algorithm and a marker-based infrared optical motion capture system. A custom Python program was implemented to stream and record trajectory data from both sources on a synchronized time scale.

For the webcam-based solution, the pose estimation component of Google's MediaPipe framework[23] (BlazePose[24]), was used. The video was streamed through a Logitech C922 Pro HD Stream Webcam mounted on a computer monitor. The video capture was controlled through the OpenCV Library[29] for Python. Figure 1 (left) shows the 33 joints obtained from the algorithm with selected labels. BlazePose can generate 3D coordinates of the 33 joints in the world coordinate system (the depth axis represents the relative distance to the camera), along with an estimated "visibility" index for each joint at each frame. The visibility index (ranges between 0 and 1) tracks the algorithm's confidence in the derived joint coordinates, which could be affected by factors such as occlusion and movement speed. A low visibility index indicates inaccuracy in

the derived coordinates. A threshold for visibility was set at 0.5 and any trajectory points with an index below the threshold were removed from the analysis. Figure 1 (middle) shows a frame from the video capture with the detected joints overlaid on top. For the optical motion capture system, OptiTrack Motion Capture System (NaturalPoint, Corvallis, OR, USA) with ten Prime^{x} 22 cameras was used. Data from OptiTrack were streamed to the Python interface through the Motive 2.0 Optical Motion Capture Software and OptiTrack's NatNet SDK. The conventional full-body biomechanical model with 39 markers (Figure 1, right) was used to capture joint trajectories (see Figure 2 in[30] for the placement of the entire marker set).

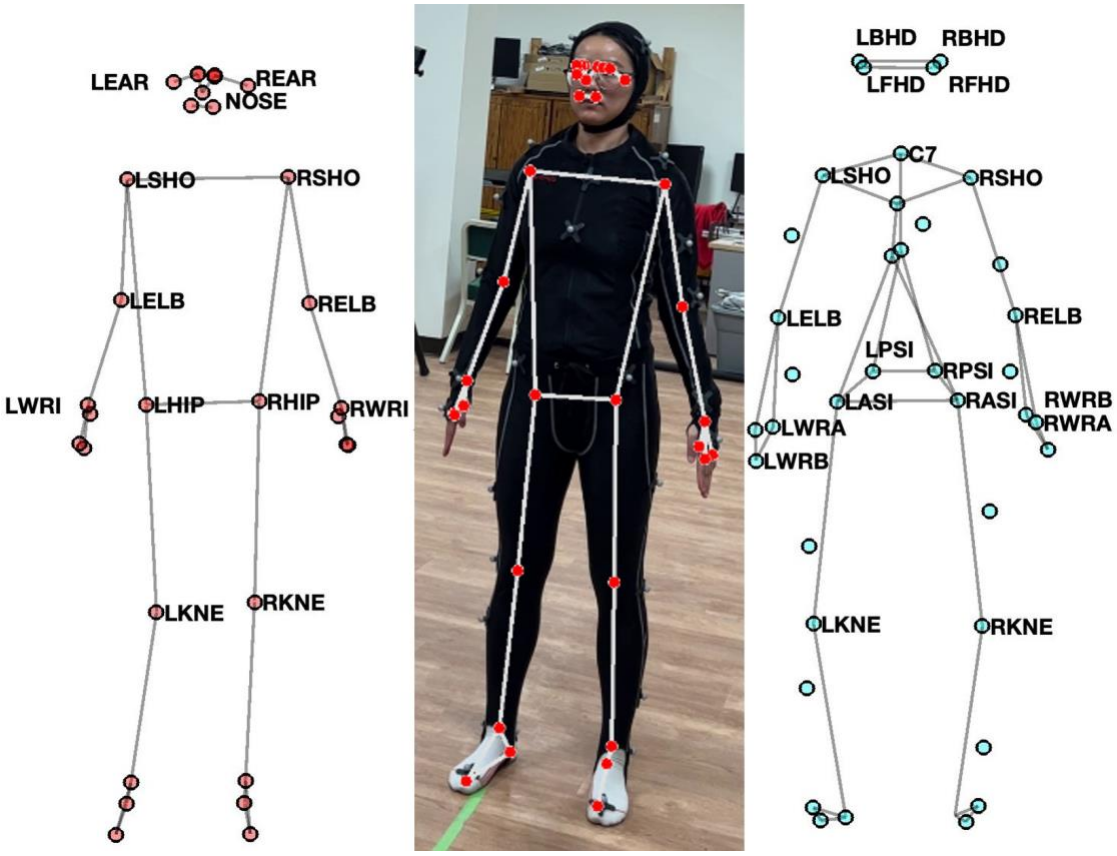

Figure 1: Two-dimensional landmarks for BlazePose (left) and the marker set used for Opti-Track's optical motion capture system (right). BlazePose produces 33 landmarks, whereas OptiTrack's marker set contains 39 markers. Markers relevant to the range of motion calculation are marked with their respective names. Figure in the middle is a screenshot of the landmarks

detected by MediaPipe's BlazePose overlay on top of a video image captured during data collection with the participant wearing a motion capture suit.

2.3 Procedures

After providing their informed consent, participants were instructed to put on a motion capture suit provided by OptiTrack. Experimenters then fitted the participants with 39 reflective markers based on the locations defined by the full-body model from OptiTrack's Motive software. Subsequently, an experimenter demonstrated each ROM evaluation movement using pre-recorded videos on a tablet computer and the participants were told to repeat the movement as practice. Table 1 shows the list of movements used to evaluate the ROM of their corresponding joints, including movements of the spine (back extension and flexion, back lateral flexion, and trunk rotation), the neck (neck extension and flexion, neck lateral bending, and neck rotation), the upper extremities (shoulder adduction and abduction, shoulder extension and flexion, and elbow extension and flexion), and the lower extremities (hip extension and flexion (knee extended), hip flexion (knee flexed), and hip adduction and abduction).

After being familiarized with all movements, participants were guided to stand in front of a computer monitor with a webcam. Participants' standing position coincided with the center of the capture volume of the OptiTrack camera system. All movements were performed while standing upright. To minimize occlusion for the webcam, participants were asked to change their orientation to the camera based on the movements they perform. For instance, for shoulder extension and flexion, participants presented the lateral view of their body to the webcam. Additionally, for movements of the lower extremities, participants were holding onto a stool while performing the movements to ensure stability. Each movement was repeated and recorded three times. For each recording, participants were instructed to perform the movement slowly

and deliberately. When the experimenter announced “Start!”, the recording would begin, and participants would start the movement. The experimenter would announce “Stop!” and then stop the recording once the participants completed the movement and returned to the upright position.

Table 1: The list of movements used to evaluate the range of motion (ROM) of various joints and their corresponding joints for BlazePose and OptiTrack marker sets. See Appendix for a breakdown of the acronyms.

Movement	BlazePose Joint 1 (Pivot)	BlazePose Joint 2 (End)	OptiTrack Joint 1 (Pivot)	OptiTrack Joint 2 (End)
Back Flexion and Extension	LHIP	LSHO	LPSI	C7
Truck Rotation	LSHO	RSHO	LSHO	RSHO
Neck Flexion and Extension	LSHO, RSHO	NOSE	LSHO, RSHO	LFHD, LBHD
Neck Lateral Bending	LSHO, RSHO	NOSE	LSHO, RSHO	LFHD, LBHD
Neck Rotation	LEAR	REAR	LFHD, LBHD	RFHD, RBHD
Shoulder Adduction and Abduction	LSHO/RSHO	LELB/RELB	LSHO/RSHO	LELB/RELB
Shoulder Flexion and Extension	LSHO/RSHO	LELB/RELB	LSHO/RSHO	LELB/RELB
Elbow Flexion	LELB/RELB	LWRI/RWRI	LELB/RELB	LWRA/RWRA, LWRB/RWRB
Hip Flexion and Extension	LHIP/RHIP	LKNE/RKNE	LASI/RASI	LKNE/RKNE
Hip Adduction and Abduction	LHIP/RHIP	LKNE/RKNE	LASI/RASI	LKNE/RKNE

2.4 Range of Motion Calculation

A novel ROM calculation method was developed and applied to calculate the joint's ROM angles based on the 3D joint trajectory data from OptiTrack and BlazePose.

2.4.1 Preprocessing

For trajectory data obtained from BlazePose, points with a low visibility index (< 0.5) were removed. Since the output from OptiTrack does not contain a visibility index, this step was not necessary to process OptiTrack's data. Then, a dual-pass Butterworth filter with a sampling frequency of 15 Hz (BlazePose) or 120 Hz (OptiTrack) and a cutoff frequency of 5 Hz (BlazePose) or 10 Hz (OptiTrack) was applied to the trajectory data. The joint trajectories were subsequently normalized using the B-spline method[31,32]. B-splines parameterize the trajectories and allow sub- or up-sampling points from the trajectory at a fixed time interval. A cubic (order 3) B-spline with a smoothing factor of 0 was applied to each dimension (x, y, and z) of the trajectory data and sampled 1000 points, equally spaced across time, from each parameterized trajectory.

2.4.2 Relevant Joints

Computationally determining a joint's ROM requires the identification of relevant joints based on the available pose information. To this end, the ROM movements were divided into two categories: rotational and non-rotational movements, as they entail slightly different relevant joint selection and ROM calculation procedures. Table 1 shows the corresponding joints for each movement. Non-rotational movements include movements such as flexion and extension, abduction, and adduction. For these movements, it is necessary to identify the relevant limb

segment formed between a “pivot” joint and an “end” joint. The “pivot” joint does not refer to the anatomical pivot movement; a pivot joint only refers to the joint around which the movement is performed, i.e., the joint of interest. For example, for shoulder adduction and abduction, the pivot joint is the shoulder whereas the end joint is the elbow. On the other hand, rotational movements include movements such as neck and trunk rotation. Compared to non-rotational movements, computing ROM for rotational movements does not require differentiating between the pivot and end joints. Instead, tracking the line formed between the two joints perpendicular to the rotational axis would provide relevant angle information.

2.4.3 Range of Motion Calculation

To derive the ROM, the angle formed by the relevant joints between the movement start and end should be computed. Because joint positions are not restricted to a two-dimensional (2D) plane, the resulting angles measured in a 3D space could be highly inconsistent due to the interaction between rotational and translational components of the movement (Figure 2 Original; note the movement of the pivot joint (green)). To address this issue, the dimensionality of joint positions was reduced by 1) identifying the best-fitting plane for each joint using singular value decomposition (SVD; Figure 2 Best-fitting Planes), and 2) projecting the joint trajectories onto their respective planes (Figure 2 Projection).

For non-rotational movements, the centroid of the pivot joint was derived by computing the mean of the pivot joint’s trajectory (Figure 2 Pivot Centroid). The line formed by the pivot joint’s centroid and the end joint was considered the limb segment. Subsequently, the ROM can be computed as the angle between the limb segment at the start and end (Figure 2 Maximum Movement Angle, top). For rotational movements, identifying the centroid is not necessary.

Instead, the two relevant joints are connected to form a line and the ROM angle is the angle of this line between the start and end positions (Figure 2 Maximum Movement Angle, bottom).

To calculate the angle, the vectors formed between the two joints at the start and end were used. At time t_0, let Joint 1 be $J_{1t_0}[x_{1t_0}, y_{1t_0}, z_{1t_0}]$ and Joint 2 be $J_{2t_0}[x_{2t_0}, y_{2t_0}, z_{2t_0}]$. The vector formed between joint 1 and 2 is $J_{1t_0} - J_{2t_0}$. Accordingly, at the end of the movement at time t, the vector formed between the two joints can be expressed as $J_{1t} - J_{2t}$. The ROM angle is the angle formed between the two vectors, α, which can be derived using dot product:

$$(J_{1t_0} - J_{2t_0}) \cdot (J_{1t} - J_{2t}) = \|(J_{1t_0} - J_{2t_0})\| \|(J_{1t} - J_{2t})\| \cos\alpha$$

Which yields:

$$\alpha = \arccos\left(\frac{(J_{1t_0} - J_{2t_0}) \cdot (J_{1t} - J_{2t})}{\|(J_{1t_0} - J_{2t_0})\| \|(J_{1t} - J_{2t})\|}\right)$$

2.5 Statistical Analysis

The three measurements from BlazePose and OptiTrack were used to evaluate their respective test-retest (intra-rater) reliability. Intra-class correlation coefficient (ICC) was computed for each movement with a two-way mixed-effect model for multiple measurements[33,34] using the ICC package in R.[35] The ICC values can be interpreted as: 0 - 0.2 (slight), 0.2 - 0.4 (fair), 0.4 - 0.6 (moderate), 0.6 - 0.8 (substantial), and 0.8 - 1.0 (almost perfect).[36] To further illustrate each measure's reliability, the standard error of measurement (SE_M) and minimal detectable change (MDC) were also computed. Different from ICC, SEM represents the measurement error in the same unit as the original measurement[37] and estimates the amount of deviation of repeated measures using the same measurement device from the "ground truth", which can be derived using its corresponding ICC:

$$SE_M = \sqrt{\sigma_T^2 \times (1 - ICC)} \quad \text{Equation 1}$$

where σ_T^2 represents the total variance. MDC represents the smallest change in value that can be detected beyond random error. MDC is based on SE_M and is expressed as

$$MDC = z_{95\%} \times \sqrt{2} SE_M \quad \text{Equation 2}$$

where $z_{95\%}$represents the z-score corresponding to the 95% confidence interval.

The inter-rater reliability between BlazePose and OptiTrack was also evaluated using two-way mixed-effect ICC for each movement. The average of three measurements from BlazePose and OptiTrack were used for the ICC computation, along with the resulting SE_M and MDC. Finally, to examine the correlation between measurements from BlazePose and OptiTrack on an individual level, a linear regression with all movements, where OptiTrack measurement was used as the predictor variable and BlazePose measurement was used as the response variable.

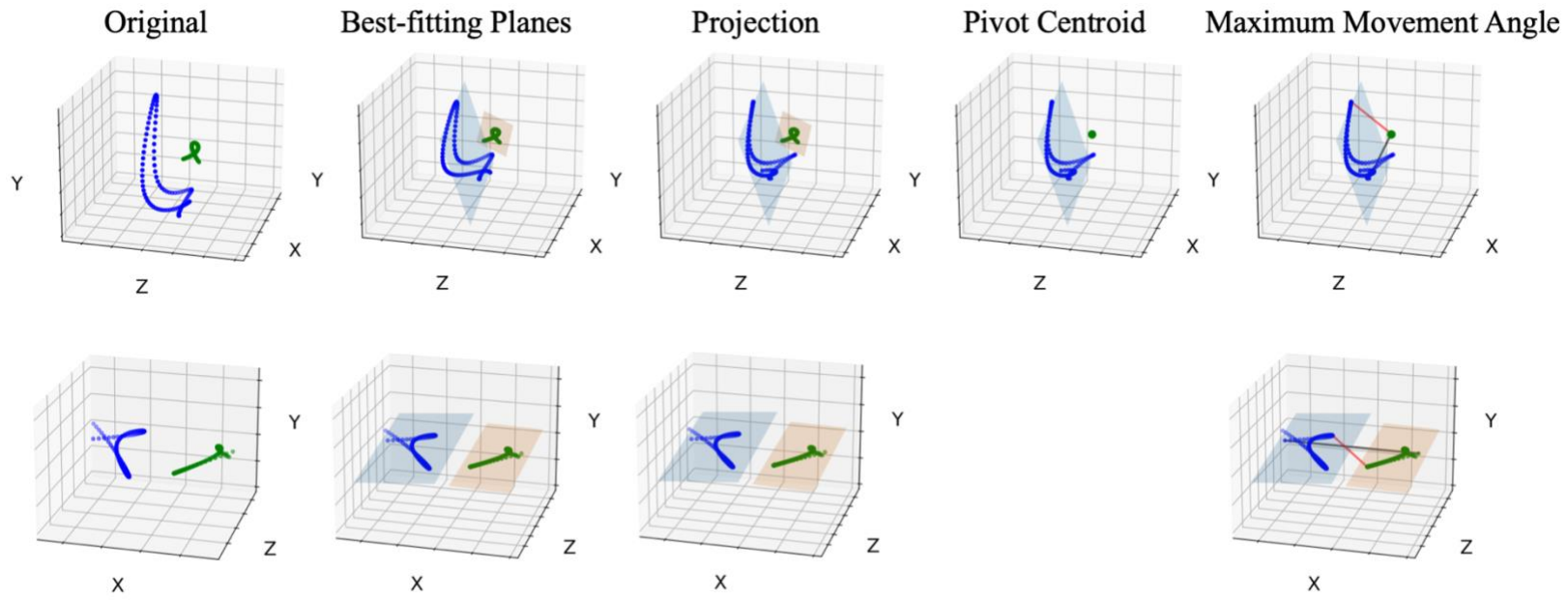

Figure 2: Joint trajectory processing procedure for the ROM calculation for non-rotational (shoulder flexion and extension; top) and rotational (trunk rotation, bottom) joints. **Original**: The original trajectories for the pivot (shoulder; green) and end (elbow; blue) joints (top), and the two shoulder joints (bottom). **Best-fitting Planes**: The original trajectories with each joint's corresponding best-fitting planes. **Projection**: The projected trajectories on the best-fitting planes. **Pivot Centroid**: The projected end joint trajectory with the pivot joint's centroid for non-rotational joints. **Maximum Movement Angle**: The angle between the initial (black) and end positions (red) that defines the range of motion.

Table 2: The intraclass correlation coefficient (ICC), standard error of measurement (SE_M), and minimal detectable change (MDC) for the range of motion angles derived using BlazePose and OptiTrack. ICC interpretations: 0 - 0.2 (slight), 0.2 - 0.4 (fair), 0.4 - 0.6 (moderate), 0.6 - 0.8 (substantial), and 0.8 - 1.0 (almost perfect).

Movement	BlazePose			OptiTrack			BlazePose and OptiTrack		
	ICC (95% CI)	SE_M	MDC	ICC (95% CI)	SE_M	MDC	ICC (95% CI)	SE_M	MDC
Spine									
Back Flexion	0.99 (0.99, 1.00)	2.55	7.08	0.99 (0.99, 1.00)	2.80	7.75	0.99 (0.99, 0.99)	3.31	9.19
Back Extension	0.82 (0.60, 0.93)	3.17	8.80	0.90 (0.80, 0.96)	3.82	10.59	0.68 (0.28, 0.86)	6.08	16.86
Trunk Rotation	0.95 (0.92, 0.97)	5.45	15.10	0.95 (0.93, 0.97)	5.24	14.54	0.90 (0.83, 0.94)	8.59	23.82
Neck									
Neck Flexion	0.80 (0.53, 0.92)	4.09	11.30	0.84 (0.68, 0.93)	6.83	18.92	0.69 (0.27, 0.87)	9.92	27.51
Neck Extension	0.89 (0.79, 0.95)	4.37	12.10	0.89 (0.79, 0.95)	5.00	13.85	0.77 (0.48, 0.90)	6.18	17.12
Neck Lateral Bending	0.95 (0.92, 0.97)	2.34	6.49	0.92 (0.87, 0.95)	3.84	10.64	0.73 (0.52, 0.85)	6.98	19.35
Neck Rotation	0.95 (0.92, 0.97)	3.22	8.92	0.97 (0.95, 0.98)	4.00	11.08	0.79 (0.63, 0.88)	10.27	28.48
Upper Extremity									
Shoulder Adduction	0.86 (0.76, 0.92)	3.90	10.80	0.89 (0.82, 0.94)	4.24	11.76	0.85 (0.72, 0.91)	4.70	13.03
Shoulder Abduction	0.95 (0.91, 0.97)	3.77	10.50	0.92 (0.88, 0.96)	4.41	12.22	0.71 (0.48, 0.83)	9.17	25.42
Shoulder Flexion	0.93 (0.88, 0.96)	3.38	9.36	0.78 (0.64, 0.87)	7.62	21.11	0.10 (-0.59, 0.49)	12.01	33.29
Shoulder Extension	0.88 (0.80, 0.93)	4.20	11.60	0.75 (0.60, 0.86)	5.97	16.54	0.84 (0.71, 0.91)	4.51	12.51
Elbow Flexion	0.83 (0.72, 0.9)	4.67	12.90	0.84 (0.74, 0.91)	5.59	15.49	0.55 (0.19, 0.75)	8.34	23.12
Lower Extremity									
Hip Flexion	0.94 (0.91, 0.97)	3.68	10.20	0.94 (0.90, 0.96)	3.54	9.80	0.76 (0.57, 0.86)	8.25	22.87
Hip Extension	0.90 (0.85, 0.94)	3.83	10.60	0.94 (0.90, 0.96)	2.75	7.63	0.85 (0.74, 0.92)	4.14	11.47
Hip Flexion (Knee Flexed)	0.85 (0.75, 0.91)	4.16	11.50	0.89 (0.81, 0.93)	3.42	9.48	0.44 (-0.01, 0.69)	7.97	22.10
Hip Adduction	0.87 (0.79, 0.93)	3.21	8.89	0.89 (0.83, 0.94)	3.14	8.71	0.84 (0.72, 0.91)	4.32	11.97
Hip Abduction	0.86 (0.78, 0.92)	3.56	9.87	0.93 (0.88, 0.96)	3.08	8.53	0.62 (0.31, 0.79)	6.81	18.87

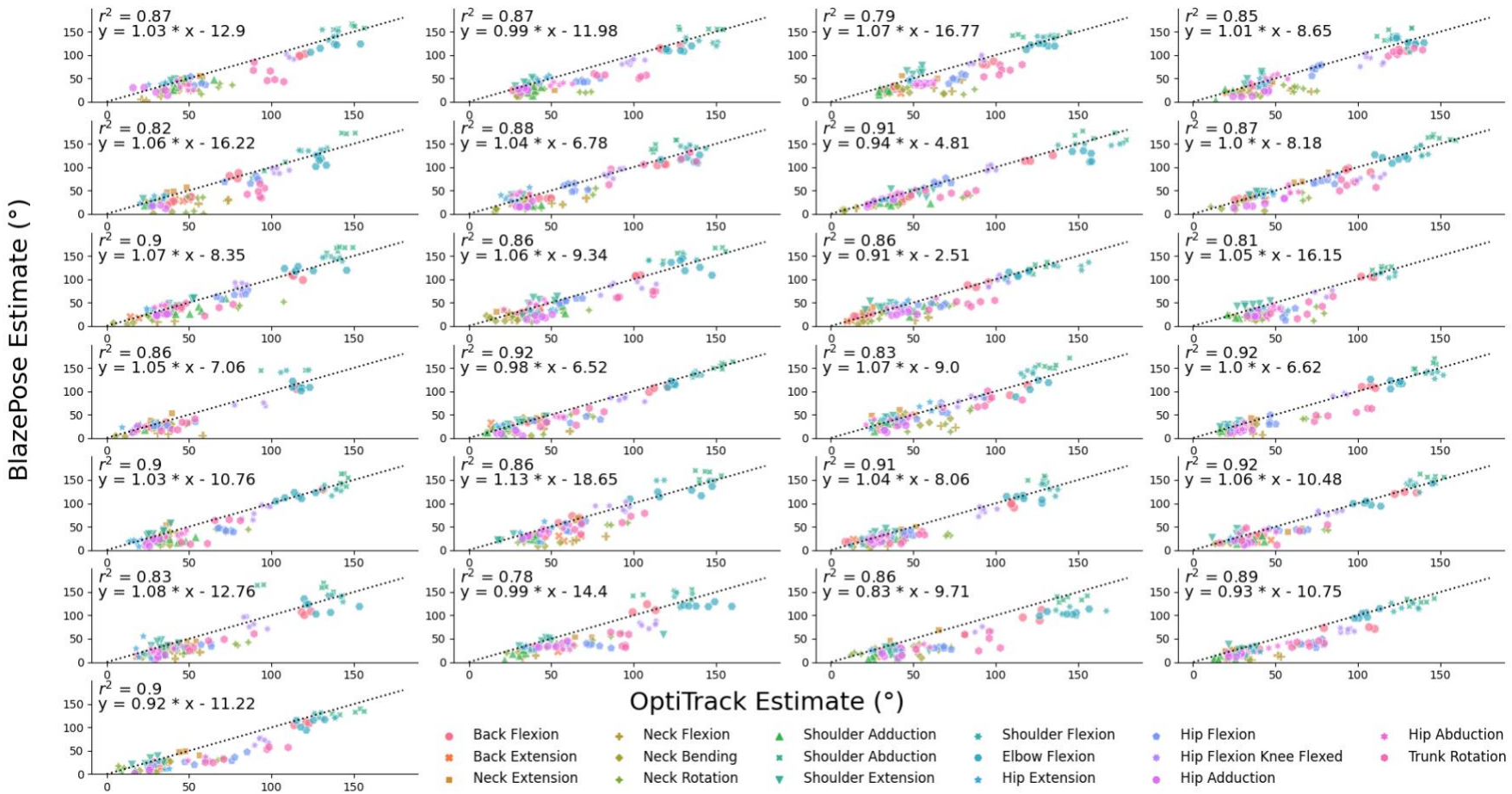

Figure 3: Regression between the range of motion (ROM) angles of different movements derived from OptiTrack (x-axis) and BlazePose (y-axis) for each participant (subplots). Different markers represent different joint movements. The dotted lines in each subplot are reference lines and have a slope of 1 and an intercept of 0.

3. Results

Table 2 shows the test-retest reliability using intraclass correlation coefficient (ICC), standard error of measurement (SE_M), and minimal detectable change (MDC) for BlazePose and OptiTrack, as well as their inter-rater reliability using the two-way mixed effect ICC, SE_M, and MDC. Overall, results indicate high test-retest and inter-rater reliability for both BlazePose and OptiTrack across different movements.

3.1 Intra-Rater Reliability

For spine movements (back flexion, back extension, and trunk rotation), the 95% confidence intervals (CI) for ICC indicate that the test-retest reliability was almost perfect for BlazePose (from 0.82 to 0.99) and OptiTrack (between 0.90 to 0.99). For BlazePose, the SEM was between 2.55° and 5.45° and the MDC was between 7.08° and 15.10°, whereas for OptiTrack, the SEM was between 2.80° and 5.24° and the MDC was between 7.75° and 14.54°. For both BlazePose and OptiTrack, trunk rotation has the highest SE_M and MDC values despite high ICCs. Because the calculations of SE_M and MDC rely on ICC and the total variability, σ_T^2 (Equation 1), the discrepancy between SE_M/MDC and ICC should be attributed to the high between-subject variability in the measurement. Furthermore, the lower bound of the 95% confidence interval for back extension (low bound = 0.60) derived using BlazePose data is relatively low compared to other joints. Because the ROM for this joint was computed based on the estimated hip and shoulder positions (Table 1), the low ICC could be attributed to the variability in pose estimation when performing the movement. Specifically, when performing the back extension movement, the camera's line of sight is perpendicular to the participants' sagittal plane, producing a lateral view of the participant. The limited view could affect the stability of

the pose estimation algorithm, which may have produced low ICC values.

For neck movements (neck extension and flexion, neck lateral bending, and neck rotation), the ICC's 95% CI indicates that BlazePose had substantial or almost perfect reliability for all movements except for neck flexion, which had moderate to substantial test-retest reliability (ICC = 0.80, CI = [0.53, 0.92]). The relatively low ICC for the BlazePose-based neck flexion estimate could be attributed to the way through which the ROM is calculated (Table 1). Given the available joints identified by BlazePose, the neck ROM was calculated based on the midpoint of the two shoulder joints (approximates the sternoclavicular joint) and the nose. During neck flexion, the webcam's view of the face and nose may become occluded as the participants tilt their head backward, which may introduce large variabilities in the estimated nose position, resulting in lower ICC values. For other measures, neck flexion and extension had relatively large SE_M (4.09° and 4.37°) and MDC (11.30° and 12.10°), whereas these values for other movements were below 4° for SE_M and 10° for MDC. For OptiTrack, all movements had almost perfect intra-rater reliability and comparable SE_M and MDC values as BlazePose.

For the upper extremity (shoulder adduction and abduction, shoulder extension and flexion, and elbow flexion), the test-retest reliability for the ROM measurements from BlazePose was generally high, ranging from moderate to almost perfect reliability. The SE_M and MDC for these movements were generally at around 4° and 10°, respectively. Reliability measures derived for the OptiTrack measurements were comparable to that from BlazePose. Based on the lower bound of the ICC's 95% confidence interval for BlazePose, two joints stood out – shoulder adduction (lower bound = 0.76) and elbow flexion (lower bound = 0.72). During data collection, participants had to wear an all-black MoCap suit (Figure 1 middle). Using a single RGB video stream, BlazePose largely relies on the contrast in the image to perform pose estimation. During

shoulder adduction and elbow flexion, there could be a significant amount of overlap between the participants' arms and torso, or between their forearm and upper arm. Combined with the suit that contains minimum contrast, the overlap presents a non-negligible challenge to the BlazePose algorithm in terms of segmenting and tracking different body parts. Therefore, it is possible that wearing everyday clothes with brighter colors during the collection process could significantly improve the reliability of these movements.

Finally, for the lower extremity (hip flexion and extension (knee extended), hip flexion (knee flexed), and hip adduction and abduction), the ICC values for BlazePose were generally high. The lower bounds of the 95% confidence intervals for hip flexion (knee flexed) (low bound = 0.75) and hip adduction (low bound = 0.79) and abduction (low bound = 0.78) could be considered relatively low. This could again be attributed to the issue of occlusion combined with a lack of contrast in the clothing as in the case of shoulder adduction and elbow flexion. All the other measures were comparable to those of other joints. Measurements derived using OptiTrack data produced similar reliability measures as those from BlazePose.

In summary, the intra-rater reliability of BlazePose's ROM calculation is relatively high and comparable to that of OptiTrack. Although the intra-class ICC values for some joints are relatively low, it could be an artifact of the experimental setup, such as the low contrast MoCap suits and inappropriate orientation when performing the movements. These issues could be easily addressed in practice to further improve the reliability of the ROM measurement.

3.2 Inter-Rater Reliability

Despite of their respective high intra-class ICC, the mixed-effect ICC between BlazePose and OptiTrack varies notably from joint to joint. On the one hand, some movements have

relatively higher ICC values, such as back flexion (ICC = 0.99, CI = [0.99, 0.99]), trunk rotation (ICC = 0.90, CI = [0.83, 0.94]), shoulder adduction (ICC = 0.85, CI = [0.72, 0.91]), shoulder extension (ICC = 0.84, CI = [0.71, 0.91]), hip extension (ICC = 0.85, CI = [0.74, 0.92]), and hip adduction (ICC = 0.84, CI = [0.72, 0.91]). On the other hand, some movements have extremely low ICC values, such as shoulder flexion (ICC = 0.10, CI = [-0.59, 0.49]) and hip flexion (knee flexed) (ICC = 0.44, CI = [-0.01, 0.69]).

This discrepancy could be attributed to the different joint/marker positions based on which the ROM angles were derived between OptiTrack and BlazePose. For instance, for shoulder flexion, although both OptiTrack and BlazePose use shoulder and elbow positions to derive the ROM angles, the shoulder locations from the two methods vary (Figure 4 left). For OptiTrack, the marker has to be attached to the MoCap suit at around the participants' acromion. When performing the shoulder flexion movement, the compression at the shoulder joint creates wrinkles on the MoCap suit and perturbs the location of the OptoTrack marker throughout the movement. For BlazePose, because joint locations were derived using images instead of a physical marker, the estimated shoulder joint location is around the humeral head and remains relatively stable throughout the movement. In contrast, although focusing on the same joint, shoulder extension had a much higher inter-rater reliability (ICC = 0.84, CI = [0.71, 0.91]). During the shoulder extension movement, the MoCap suit remains relatively stable (Figure 4 right), which circumvents the potential artifact in the shoulder marker's position during the movement. Therefore, the low inter-rater reliability between OptiTrack and BlazePose could be attributed to the intrinsic limitations of OptiTrack.

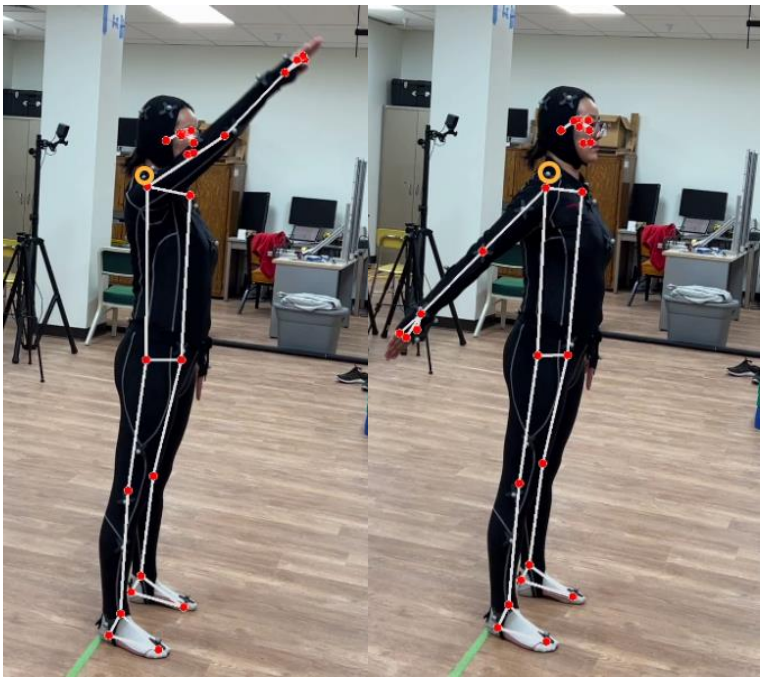

Figure 4. A participant performing the shoulder flexion (left) and extension (right) movements. The orange circle highlights the locations of the OptiTrack markers.

3.3 Individual-Level Reliability

Finally, the individual-level correlations between OptiTrack and BlazePose were evaluated using linear regressions (Figure 3). The r^2 values are generally high (min = 0.78, max = 0.92, mean = 0.87, SE = 0.008), suggesting good fits of the linear models. The regression slopes (mean = 1.01, SE = 0.01) indicate a one-to-one mapping between the two measurements. The combination of high r^2 values and regression slopes of 1 demonstrate the high consistency of ROM measurement between OptiTrack and BlazePose on an individual level across all joints.

What is noticeable, however, is the negative regression intercept (mean = -10.34, SE = 0.77). As Figure 3 shows, the ROM angles derived from BlazePose were smaller than those from OptiTrack, as the data points generally fall below the reference line. This pattern could be attributed to the artifacts associated with OptiTrack's marker placements. While the predicted joint location from BlazePose is around the anatomical joint, OptiTrack's markers are placed

topically and attached to the MoCap suit (Figure 1 middle). As a result, the marker position is not an accurate reflection of the actual joint location. For instance, Figure 5 shows sample poses of a participant performing the shoulder abduction movement derived from BlazePose and OptiTrack at the same time stamp during data collection. Although the shoulder positions are comparable between the two samples, the elbow position is much higher in the OptiTrack sample than that in the BlazePose sample. This is because the marker for OptiTrack was attached to the participant's lateral epicondyle, which was spatially higher than the elbow joint for this pose. Consequently, the final derived ROM angle from OptiTrack was larger as compared to that from BlazePose. This issue also applies to the computation of other movements, such as those of the lower extremity (note the difference in the knee positions).

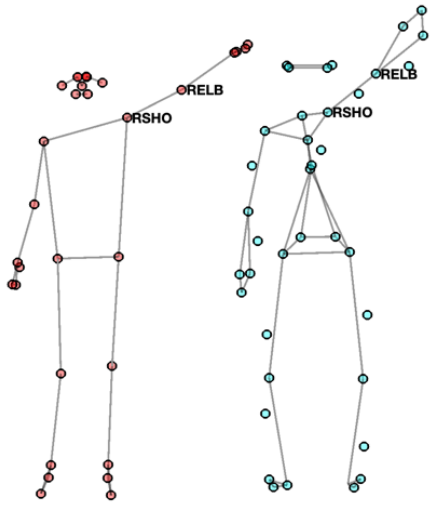

Figure 5. Joint position comparison between BlazePose (left) and OptiTrack (right) during the shoulder abduction movement at a synchronized time. The relevant joints (RSHO or Right Shoulder, and RELB or Right Elbow) are annotated. See text for details.

4. Discussion

The current study presented a webcam-based machine learning solution using BlazePose for range of motion (ROM) measurement and evaluated its intra- and inter-rater reliability as compared to a marker-based motion capture system, OptiTrack. Unlike previous studies, the present study examined a diverse set of joints using a full-body biomechanical model, including those of the spine, the neck, and the upper and lower extremities.

Results revealed high intra-rater reliability for BlazePose and OptiTrack. For almost all movements, both measurement methods had substantial or almost perfect intra-rater reliability, with their corresponding measurement errors (SE_M) below 5° and the minimal detectable change (MDC) at around 10°. This indicates that the webcam-based solution could reliably differentiate changes in ROM from measurement errors when the amount of change is greater than 10°. From a practitioner's perspective, this would be sufficient to evaluate the effect of an intervention.[38] Furthermore, careful examination of movements with lower inter-rater reliability showed that changing the participants' orientation to the camera's line of sight and the color scheme of their cloth may further improve BlazePose's inter-rater reliability.

Despite their respective high test-retest reliability, the inter-rater reliability between BlazePose and OptiTrack is variable across different joints, with some joint movements having relatively low inter-rater reliability. Scrutinizing these movements uncovered that the variable inter-rater reliability could be attributed to the intrinsic limitations of the OptiTrack MoCap system. Because OptiTrack's markers must be attached to the participants' cloth, certain movements may create deformation in the cloth and, consequently, false marker movements that do not reflect the movement of the limbs. In contrast, because the detected joint positions from BlazePose are around the participants' anatomical joints, ROM measurement obtained from

BlazePose should contain fewer artifacts. Similarly, the limitation associated with the marker placement also affects the correlation between BlazePose and OptiTrack on an individual level. Linear regressions showed that the individual-level reliability was high and there was a one-to-one mapping between the two measurement methods. However, there was also an approximate 10° in average difference between measurements from OptiTrack and BlazePose. Analysis of single poses during the movements showed that this difference could be considered an overestimation in OptiTrack's ROM measurement because of the offset between the marker and the anatomical joint.

From a practitioner's perspective, the effectiveness of a ROM measurement method should be evaluated from a functional standpoint. In this context, the measurement method should not solely focus on achieving high anatomical accuracy and precision. Instead, the method's intra-rater reliability and accessibility are also equally important. Specifically, ROM is commonly used to evaluate and quantify the improvement in a patient's mobility at a certain joint as a result of an intervention. Therefore, an ideal ROM measurement method should reliably evaluate the ROM across multiple sessions. Additionally, accessibility is also a critical consideration. The COVID-19 Pandemic raises awareness of the importance of telehealth and its effectiveness in lowering the barrier of access for disadvantaged populations. In the context of physical therapy, ROM evaluation has always relied on a goniometer,[4,5] which has long been known for its lack of reliability despite requiring extensive training and practice.[11,12,39] Merely relying on a goniometer for physical therapy would disproportionately affect patients of different racial and socioeconomic backgrounds.

As an alternative, the webcam-based ROM evaluation tool presented in the current study is affordable and accessible with high intra-rater reliability. Since BlazePose is lightweight and

compatible with multiple platforms, it is possible to adapt the current data collection and processing pipeline to a browser-based web application and mobile applications on iOS and Android devices. Future studies should focus on refining the applied aspect of this tool, such as identifying the ideal type of cloth and camera angles for different movements to maximize tracking reliability.

5. Conclusion

In conclusion, the current study presents an alternative way of measuring joint range of motion (ROM) that merely relies on a webcam setup. Compared to a state-of-the-art motion capture system, the webcam-based machine learning approach demonstrated high intra- and inter-rater, as well as individual-level reliability in quantifying and assessing joint ROM. Adapting and adopting this tool for tele-implementation of physical therapy and rehabilitation could significantly reduce the barrier of access to healthcare.

References

1. Garfan S, Alamoodi AH, Zaidan B, et al. Telehealth utilization during the Covid-19 pandemic: A systematic review. *Comput Biol Med*. 2021;138:104878.

2. Snoswell CL, Chelberg G, De Guzman KR, et al. The clinical effectiveness of telehealth: a systematic review of meta-analyses from 2010 to 2019. *J Telemed Telecare*. Published online 2021:1-16.

3. Tsou C, Robinson S, Boyd J, et al. Effectiveness of Telehealth in Rural and Remote Emergency Departments: Systematic Review. *J Med Internet Res*. 2021;23(11):e30632. doi:10.2196/30632

4. Muir SW, Corea CL, Beaupre L. Evaluating change in clinical status: reliability and measures of agreement for the assessment of glenohumeral range of motion. *North Am J Sports Phys Ther NAJSPT*. 2010;5(3):98.

5. Mullaney MJ, McHugh MP, Johnson CP, Tyler TF. Reliability of shoulder range of motion comparing a goniometer to a digital level. *Physiother Theory Pract*. 2010;26(5):327-333.

6. Riddle DL, Rothstein JM, Lamb RL. Goniometric reliability in a clinical setting: shoulder measurements. *Phys Ther*. 1987;67(5):668-673.

7. Rothstein JM, Miller PJ, Roettger RF. Goniometric reliability in a clinical setting: elbow and knee measurements. *Phys Ther*. 1983;63(10):1611-1615.

8. van de Pol RJ, van Trijffel E, Lucas C. Inter-rater reliability for measurement of passive physiological range of motion of upper extremity joints is better if instruments are used: a systematic review. *J Physiother*. 2010;56(1):7-17.

9. Meislin MA, Wagner ER, Shin AY. A comparison of elbow range of motion measurements: smartphone-based digital photography versus goniometric measurements. *J Hand Surg*. 2016;41(4):510-515.

10. Li R, Jiang Q. A photogrammetric method for the measurement of three-dimensional cervical range of motion. *IEEE J Biomed Health Inform*. Published online 2021.

11. Reese NB, Bandy WD. *Joint Range of Motion and Muscle Length Testing-E-Book*. Elsevier Health Sciences; 2016.

12. Kolber MJ, Hanney WJ. The reliability and concurrent validity of shoulder mobility measurements using a digital inclinometer and goniometer: a technical report. *Int J Sports Phys Ther*. 2012;7(3):306-313.

13. Pérez-de la Cruz S, de León ÓA, Mallada NP, Rodríguez AV. Validity and intra-examiner reliability of the Hawk goniometer versus the universal goniometer for the measurement of range of motion of the glenohumeral joint. *Med Eng Phys*. 2021;89:7-11.

14. Dent Jr PA, Wilke B, Terkonda S, Luther I, Shi GG. Validation of teleconference-based goniometry for measuring elbow joint range of motion. *Cureus*. 2020;12(2).

15. Cejnog LWX, Cesar RM, de Campos TE, Elui VMC. Hand range of motion evaluation for Rheumatoid Arthritis patients. In: *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*. IEEE; 2019:1-5.

16. Feng M, Liang L, Sun W, et al. Measurements of cervical range of motion using an optical motion capture system: Repeatability and validity. *Exp Ther Med*. 2019;18(6):4193-4202.

17. Nagymáté G, Kiss RM. Application of OptiTrack motion capture systems in human movement analysis: A systematic literature review. *Recent Innov Mechatron*. 2018;5(1.):1-9.

18. Furtado JS, Liu HH, Lai G, Lacheray H, Desouza-Coelho J. Comparative analysis of optitrack motion capture systems. In: *Advances in Motion Sensing and Control for Robotic Applications*. Springer; 2019:15-31.

19. Toshev A, Szegedy C. Deeppose: Human pose estimation via deep neural networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. ; 2014:1653-1660.

20. Cao Z, Simon T, Wei SE, Sheikh Y. Realtime multi-person 2d pose estimation using part affinity fields. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. ; 2017:7291-7299.

21. Mehta D, Sridhar S, Sotnychenko O, et al. Vnect: Real-time 3d human pose estimation with a single rgb camera. *ACM Trans Graph TOG*. 2017;36(4):1-14.

22. Wang J, Tan S, Zhen X, et al. Deep 3D human pose estimation: A review. *Comput Vis Image Underst*. 2021;210:103225. doi:10.1016/j.cviu.2021.103225

23. Lugaresi C, Tang J, Nash H, et al. Mediapipe: A framework for building perception pipelines. *ArXiv Prepr ArXiv190608172*. Published online 2019.

24. Bazarevsky V, Grishchenko I, Raveendran K, Zhu T, Zhang F, Grundmann M. BlazePose: On-device Real-time Body Pose tracking. *CoRR*. 2020;abs/2006.10204. https://arxiv.org/abs/2006.10204

25. Kulikajevas A, Maskeliūnas R, Damaševičius R, et al. Exercise Abnormality Detection Using BlazePose Skeleton Reconstruction. In: *International Conference on Computational Science and Its Applications*. Springer; 2021:90-104.

26. Mohammed SW, Garrapally V, Manchala S, Reddy SN, Naligenti SK. Recognition of Yoga Asana from Real-Time Videos using Blaze-pose. *Int J Comput Digit Syst*. 2022;5(3).

27. Rojas-Arce JL, Jimenez-Angeles L, Marmolejo-Saucedo JA. Telerehabilitation Prototype for Postural Disorder Monitoring in Parkinson Disease. In: *International Conference on Computer Science and Health Engineering*. Springer; 2021:129-142.

28. Singha RG, Lad M, Shipurkar GM, Rohekar A, Chauhan C, Rathod N. Dynamic Pose Diagnosis with BlazePose and LSTM for Spinal Dysfunction Risk Estimation. In: *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*. IEEE; 2022:1547-1552.

29. Bradski G. The OpenCV Library. *Dr Dobbs J Softw Tools*. Published online 2000.

30. Kluwak K, Klempous R, Chaczko Z, Rozenblit JW, Kulbacki M. People Lifting Patterns—A Reference Dataset for Practitioners. *Sensors*. 2021;21(9):3142. doi:10.3390/s21093142

31. Gallivan JP, Chapman CS. Three-dimensional reach trajectories as a probe of real-time decision-making between multiple competing targets. *Front Neurosci*. 2014;8. Accessed July 13, 2022. https://www.frontiersin.org/articles/10.3389/fnins.2014.00215

32. Ramsay JO, Silverman BW. Functional Data Analysis. *İnternet Adresi Http*. Published online 2008.

33. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods*. 1996;1(1):30.

34. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155-163.

35. Wolak ME, Fairbairn DJ, Paulsen YR. Guidelines for Estimating Repeatability. *Methods Ecol Evol*. 2012;3(1):129-137.

36. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. Published online 1977:159-174.

37. Stratford PW, Goldsmith CH. Use of the standard error as a reliability index of interest: an applied example using elbow flexor strength data. *Phys Ther*. 1997;77(7):745-750.

38. Hall T, Briffa K, Hopper D, Robinson K. Long-term stability and minimal detectable change of the cervical flexion-rotation test. *J Orthop Sports Phys Ther*. 2010;40(4):225-229.

39. Perez-Grau F, Ragel R, Caballero F, Viguria A, Ollero A. Semi-autonomous teleoperation of UAVs in search and rescue scenarios. In: *2017 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE; 2017:1066-1074.

Appendix
Joint Acronym Lookup

Head

LFHD/LFHD: Left/Right Front of the Head

LBHD/LBHD: Left/Right Back of the Head

LEAR/REAR: Left/Right Ear

NOSE: Nose

Torso

LSHO/RSHO: Left/Right Shoulder

LASI/RASI: Left/Right Anterior Sacroiliac

LPSI/RPSI: Left/Right Posterior Sacroiliac

Upper Limbs

LELB/RELB: Left/Right Elbow

LWRI/RWRI: Left/Right Wrist

LWRA/RWRA: Left/Right Wrist A (Anterior)

LWRB/RWRB: Left/Right Wrist B (Posterior)

Lower Limbs

LHIP/RHIP: Left/Right Hip

LKNE/RKNE: Left/Right Knee