IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2024.0429000

Exercise Assessment based on Human Pose Estimation and Relative Phase for Real-Time Remote Exercise System

YEONGJU WOO[1], HIEYONG JEONG[1]
[1]Department of Artificial Intelligence Convergence, Chonnam National University, Buk-gu, Gwangju, Republic of Korea (e-mail: {dndudwn31, hang_208258, h.jeong}@jnu.ac.kr)

Corresponding author: Hieyong Jeong (e-mail: h.jeong@jnu.ac.kr).

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) under the Artificial Intelligence Convergence Innovation Human Resources Development (IITP-2023-RS-2023-00256629) grant funded by the Korea government (MSIT), and the Ministry of Science and ICT (MSIT), Korea, under the National Program for Excellence in Software (2021-0-01409), supervised by the Institute for Information & Communications Technology Planning & Evaluation (IITP).

ABSTRACT Human pose estimation (HPE) technology, a vital tool for assessing exercise posture by extracting the three-dimensional coordinates of each joint, has been applied in many studies. It is effective for determining static postures, such as yoga; however, it still presents challenges in assessing dynamic exercise that involves considering each joint's velocity. Although the HPE technology enables the derivation of position and velocity from each joint, assessing exercise posture using multiple time-series data requires time consumption and expert knowledge. Therefore, this study addressed this challenge by introducing a method for determining the velocity-based exercise posture, which combines the coordinates of significant-extracted joints using HPE with the relative-phase method. The relative phase angle ($\Delta\phi_{Angle}$) is valuable for assessing the combination of position and velocity. This study added the relative phase distance ($\Delta\phi_{Distance}$). An experiment was conducted to compare the exercise postures of experts and beginners during barbell back squats using a constructed dataset of time-series data for the positions and velocities of each joint and the relative phase Angle and Distance. Training and prediction were performed using a one-dimensional deep learning model. The results demonstrated the effectiveness of the proposed index in velocity-based exercise assessment with over 95% accuracy and confirmed the robustness of the method without requiring expert knowledge in real time. This study has significant implications for practical application in sports science and biomechanics. It has the potential to revolutionize the assessment and improvement of exercise posture.

INDEX TERMS relative phase, velocity-based exercise posture, human pose estimation, feature extraction, deep-learning-based personal training system

I. INTRODUCTION

Strength is an essential indicator not only for measuring individual performance in sports but also for optimizing the strength capacity of athletes [1], [2]. Coaches evaluate athletes' strength before and after prescribing training programs to assess the effectiveness of the programs [3].

This strength is commonly quantified through the one-repetition maximum (1RM) test, which measures the maximum strength capacity an athlete can lift with a proper form in a single repetition [4]. However, this method has a drawback in that it depends on the athlete's condition, resulting in up to a 36% variance, even for the same athlete [5].

Recently, velocity-based training (VBT) methods have been proposed as alternatives for evaluating strength. This method measures the angles and velocities of joints to assess the intensity of movements, offering a way to evaluate dynamic strength, similar to the 1RM test [6]. This method is evaluated using measurement data by attaching sensors or markers to specific joints that play a key role in the movement. O'Reilly et al. [7] proposed a method to identify differences in seven types of squat movements using IMU sensors, whereas Lee et al. [8] reported a method for classifying squat postures by training an AI model with IMU sensor data. Woo et al. [9] analyzed gait data collected using IMU sensors and discovered differences in gait symmetry between elderly patients with diabetes and healthy elderly individuals. Qi et al. [10] proposed a method to recognize exercise intensity by combining accelerometer and electrocardiogram data.

However, the VBT method requires further improvement. This method necessitates the direct attachment of wearable sensors to specific joints, which play key roles in movement, and it can only be tested in well-configured environments. Traditional sensor-based methods are often limited by their reliance on physical hardware, which can be intrusive and restrict natural movements. In contrast, deep learning models eliminate the need for physical sensors by leveraging human motion estimation technology, thereby enabling seamless data acquisition in various environments. Moreover, as the number of joints increases, the amount of time-series data to be analyzed also increases, making the analysis time-consuming. Deep learning models address this challenge by automating the feature extraction process, enabling the efficient processing of large-scale time-series data through parallel computations and hierarchical pattern recognition. From the coach's perspective, providing real-time feedback to athletes based on measured data requires specialized engineering knowledge and experience with sensor data. By utilizing deep learning models, this process is greatly simplified as these models automatically interpret the data and provide actionable insights, reducing the dependency on specialized expertise.

Therefore, this study proposes a method for evaluating velocity-based exercise postures by combining human motion estimation technology, which can estimate human joint coordinates, with a relative phase method, and aims to validate its effectiveness [11]. Deep learning models complement the relative phase method and motion estimation by providing robust data analysis capabilities and enhancing the user-friendliness and efficiency of the system. The relative phase method can derive a relative angle indicator that evaluates the combination of position and velocity, as well as a relative phase distance indicator that assesses the consistency of repeated postures. After training the deep learning model with time-series data from joints and these two indicators, we confirmed an accuracy rate of over 95% in the tests.

II. RELATED WORKS

A. MOTION-CAPTURE-SYSTEM-BASED HUMAN POSE ASSESSMENT

Motion capture, which is the process of recording the movement of an object or person by measuring its position and orientation in physical space, has seen rapid development in recent decades. Various approaches have been developed, for commercial motion capture commonly using inertial sensors. For example, the Xsens MVN [12] system employs 17 IMU sensors, such as accelerometers, gyroscopes, and magnetometers, to track the six degrees of freedom (DOF) at the joints of the body.

Compared with vision-based motion capture, sensor-based motion capture offers greater freedom of movement. However, the installation process requires a significant investment in time and money because of the need for many inertial sensors. Consequently, the current study focuses on achieving satisfactory results with fewer sensors despite the potential for performance decline. This underscores the need for proactive measures to maintain control and to ensure successful outcomes.

Motion capture can be achieved using an inertial sensor or an optical motion capture using an optical sensor or video. The prevalent optical motion capture device is OptiTrack [13], which utilizes multiple infrared cameras to track information from capture sensors affixed to the human body, and processes the captured two-dimensional(2D) position data from the sensors to generate three-dimensional(3D) data.

While this approach offers the advantages of unrestricted movements, the ability to track multiple individuals, and suitability for fast movements, it is important to note its drawbacks. These include the need for more data owing to markers being covered during operation and the limitation of capturing in a confined space where the camera is installed.

In recent years, single-camera-based Human Pose Estimation (HPE) has significantly surpassed the limitations of sensor-based and traditional motion capture systems. For example, OpenPose [14], introduced in 2017, enables the real-time extraction of feature points from videos or photos, regardless of the number of people, using deep learning. Furthermore, single-camera-based markerless systems have demonstrated strong performance in studies involving single-plane measurements, such as analyzing infant movements or spatiotemporal gait parameters, outperforming 3D marker-based systems in terms of specific clinical outcome measures.

Previously, markerless Motion Capture Systems (MCS) were less effective than marker-based systems for tasks requiring detailed 3D kinematics or precise movements, such as finger tracking. However, VideoPose3D [15], introduced in 2019, demonstrated effective 3D pose reconstruction using a dilated temporal convolution model to process 2D keypoints from images. Similarly, BlazePose [16], launched in 2020, facilitates real-time human pose inference on mobile devices by employing lightweight pose estimation through heatmap and regression techniques. While markerless MCS shows promise in expanding movement analysis beyond laboratory environments, achieving the level of accuracy required for detailed three-dimensional kinematics in clinical decision-making remains a challenge.

Recent advancements in HPE have significantly enhanced its potential for dynamic exercise assessment. Methods such as VideoPose3D and BlazePose have expanded the capabilities of marker-less motion capture, enabling precise 3D pose reconstruction. These advancements have been applied in studies, such as gait symmetry analysis and spatiotemporal parameter estimation for clinical applications. By integrating these cutting-edge technologies with the proposed relative phase method, this study bridges the gap between theoretical biomechanics and practical real-time exercise evaluations.

B. HUMAN-POSE-ESTIMATION-BASED EXERCISE ASSESSMENT

Human Pose Estimation (HPE) is a computer vision technology that predicts a person's posture by specifying a person's joints or essential body parts as key points. It is widely used in various fields such as autonomous driving, sports, medicine, and metaverse [17]–[21]. Recently, with the development of cameras and human estimation technology, methods for evaluating exercise performance have evolved from the existing sensors to optical methods. In particular, commercial motion capture uses many markers and multiple calibrated cameras, and several studies have attempted to overcome the drawbacks of popular approaches that use single or multiple cameras. Additionally, as HPE technology advances, the fitness technology market is filled with AI-powered personal trainer apps.

The Zurich-based VAY Sports VAY Fitness Coach [22] is a dedicated workout app that understands movements and provides real-time feedback during a workout. The VAY Fitness Coach evaluates poses and movements and compares them to the user's target execution through pose estimation. In addition, it measures the speed of specific body parts to calculate the speed of movement, count repetitions, collect critical angles of the body, and define thresholds for the analysis.

Munich-based Kaia Health's Kaia Personal Trainer [23] is the world's first full-body virtual personal training app that tracks workouts, creates personalized fitness plans, counts reps, and provides real-time audio feedback. It combines AI-powered motion tracking with customized training to deliver personalized full-body workouts featuring a variety of workouts, tracking body activity with a 16-point system that compares metrics for ideal movement and positioning of limbs, joints, and angles to track repetition. The fitness level was determined by counting and calculating the angle of the exercise posture.

ALFA-AI [24] monitors a user's exercise execution with real-time AI analysis and provides real-time visual and auditory improvement feedback. The key joints of the user were tracked through a two-dimensional coordinate system, and personal AI training was continuously adjusted according to the actual user's performance. In addition, ALFA-AI is a function of a golf coach, and the AI algorithm analyzes the user's swing movement and ball flight results to provide personal feedback. infiGro [25], created by Infivolve, is a fully automated, AI-powered, digital personal trainer app that guides, analyzes, corrects, and motivates in real-time via the phone's camera. It shows an example video of an expert and counts the number of repetitions by the user through pose estimation.

Based on our research surveys, existing applications primarily focus on simple exercise assessments, such as estimating angles using body coordinates or surpassing a specified threshold, and static exercise assessments, such as yoga. However, we did not encounter an exercise posture assessment method that considers the velocity of the individual joints. This unique aspect sets up the proposed exercise posture assessment method.

III. RELATIVE PHASE

A. RELATIVE PHASE ANGLE

In recent decades, numerous biomechanical and motor control studies have focused on quantifying coordination and coupling relationships within the human motor system. At the behavioral level, the continuous relative phase (CRP) method has not only emerged, but has also gained significant popularity as an approach for assessing intra- and inter-limb coordination. During a gait cycle, the CRP measures the phase angle difference between two degrees of freedom (DoFs), such as the femoral and tibial angles. The phase angle is obtained by constructing phase plane portraits for each DoF and then calculating the difference in the phase angle between them.

Consider the following representation of a sinusoidal signal:

$$x(t) = A\sin(\omega t + \theta_x). \tag{1}$$

In Equation 1, t, A, ω, and θ_x represent the time, magnitude, frequency, and phase shift, respectively. The standard method for determining the phase angle involves creating a phase plane portrait by plotting the signal on the horizontal axis and its time derivative on the vertical axis. Subsequently, the phase angle was calculated as the angle between the line connecting the origin and the point on the phase plane trajectory and the right horizontal line. Namely,

$$\begin{aligned}\phi_x(t) &= \arctan(\frac{\dot{x}(t)}{x(t)}) \\ &= \arctan(\frac{A\omega\cos(\omega_x t + \theta_x)}{A\sin(\omega_x t + \theta_x}).\end{aligned} \tag{2}$$

In Equation 2, ϕ_x represents the phase angle of the signal x, and $\dot{x}$ represents the time derivative of x. The transition from Equation 2 to Equation 3 involves a key step: the range normalization of ω. Range normalization ensures that the frequency term ω does not influence the phase angle calculation beyond its angular periodicity, which is critical for simplifying trigonometric relationships. After range-normalizing ω, Equation 2 can be expressed as

$$\begin{aligned}\phi_x(t) &= \arctan(\frac{\cos(\omega_x t + \theta_x)}{\sin(\omega_x t + \theta_x)}) \\ &= \arctan(\cot(\omega_x t + \theta_x)).\end{aligned} \tag{3}$$

Equation 3 can be transformed because the inverse tangent of the cotangent does not result in the associated angle:

$$\begin{aligned}\phi_x(t) &= \arctan(\tan(\frac{\pi}{2}) - (\omega_x t + \theta_x)) \\ &= \frac{\pi}{2} - (\omega_x t + \theta_x).\end{aligned} \tag{4}$$

Similarly, for the sinusoidal signal, $y(t)$, the phase angle is:

$$\phi_y(t) = \frac{\pi}{2} - (\omega_y t + \theta_y). \tag{5}$$

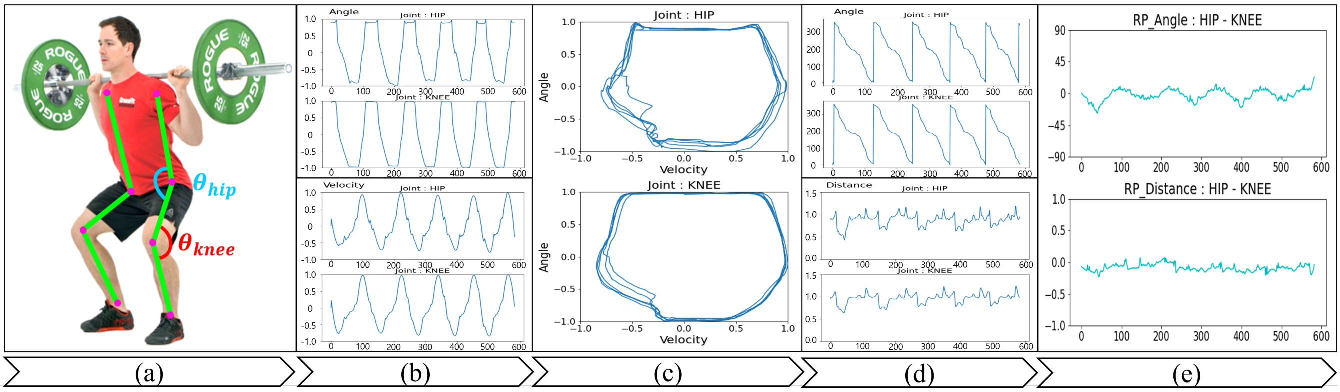


FIGURE 1. A descriptions of used angles and angular velocities and the method of calculating the proposed ERP(extended relative phase) indicator for the assessment of velocity-based exercise posture: (a) representation of two joints, (b) description of the results of normalized angles and angular velocities, (c) description of the results of the Phase Portrait of the hip-knee during exercise, (d) representation of the results of relative phase angle and relative phase distance, and (e) description of the results of the subtraction between two relative phase angles and between two relative phase distances, respectively.

The phase angle can be calculated while subtracting two equations of 4 and 5:

$$\begin{aligned}\Delta\phi_{x-y} &= \phi_x(t) - \phi_y(t)\\ &= \frac{\pi}{2} - (\omega_x t + \theta_x) - \frac{\pi}{2} - (\omega_y t + \theta_y)\\ &= \theta_x - \theta_y\\ &= \Delta\phi_{y-x} \perp . \end{aligned} \tag{6}$$

Here, $\perp$ represents a contradiction. Namely, while the intention is to determine the CRP in x-y order, the result is a y-x order or 180° phase shift. There are two solutions to rectify the order error: first, constructing phase-plane portraits by plotting position (x) over velocity ($\dot{x}$) and second, using the Hibert transform. The first method was adopted in this study. Thus, by plotting the position over the velocity, Equation 2 becomes

$$\begin{aligned}\phi_x &= \arctan(\frac{x(t)}{\dot{x}(t)})\\ &= \arctan(\tan(\omega t + \theta_x))\\ &= \omega t + \theta_x. \end{aligned} \tag{7}$$

Therefore, the 90° phase shift of the phase angle no longer exists and thus $\Delta\phi_{x-y} = \theta_x - \theta_y$.

B. RELATIVE PHASE ANGLE CALCULATION IN PRACTICAL SCENARIOS

The relative phase angle methodology is crucial to understanding movement coordination, and its practical implementation requires careful data handling. In real-world applications, after extracting joint coordinates using HPE, angular velocities and positions are computed using numerical differentiation techniques such as finite difference methods. These values are then normalized to account for intersubject variability. The phase angle is derived by plotting the angular displacement against velocity in a phase-plane representation, ensuring consistency across different movement cycles. By applying robust filtering techniques such as the Hampel and Savitzky-Golay filters, the noise from sensor drift and tracking errors is mitigated, improving the reliability of the computed phase angles.

C. CASE OF CALCULATING RELATIVE PHASE ANGLE

Figure 1 shows that the relative phase method analyzes the joint angles in the barbell back squat exercise posture. Because of the time-consuming nature of assessing exercise consistency in simple squat postures, we decided to use the barbell back squat as the exercise posture, as shown in Figure 1(a). Blaze Pose, a machine learning-based kit pose detection API, was utilized to detect the joint angles in this exercise posture. Blaze Pose is an algorithm for human pose estimation that can infer 33 3D landmarks and background segmentation masks for an entire body from RGB video frames. The performance of the BlazePose GHUM model was evaluated and demonstrated to have exceptionally high accuracy, as shown in Table 1, and was evaluated on three validation datasets: yoga, dance, and HIIT, confirming the model's strong performance.

One coordinate of the detected skeletal point indicates the coordinate position of the x, y, and z-axes viewed from the absolute coordinates. The 99 time-series data were acquired from the 33 3D skeletal coordinates detected. Because this study aims to verify the effectiveness of the proposed algorithm during barbell back squats, eight skeletal coordinate information points are required to assess this exercise posture. However, this method is adaptable and can be used for various exercises. Thus, 24 time-series data points need to be processed. however, only one plane is applied to reduce the data. Because image data are used to estimate the skeletal coordinate position, a preprocessing step is required to remove the

TABLE 1. The comparison results of human-pose-estimation quality of BlazePose GHUM model in MediaPipe Pose through three public data set.

Method	Yoga	Dance	HIIT
BlazePose-GHUM (Heavy)	96.4	97.2	97.5
BlazePose-GHUM (Full)	95.5	96.3	95.7
BlazePose-GHUM (Lite)	90.2	92.5	93.5
AlphaPose-Resnet50	96.0	95.5	96.0
Apple Vision	82.7	91.4	88.6

noise caused by movement between pixels and noise caused by calculating the hidden skeletal coordinates according to the exercise posture.

The preprocessing pipeline plays a pivotal role in ensuring the data quality for velocity-based posture evaluation. Initially, outliers in the estimated 3D landmark coordinates were removed using a Hampel filter [26]. This filter identifies anomalies by comparing each data point to its neighbors within a defined window and threshold based on the median absolute deviation (MAD). By focusing on the local median and MAD, the Hampel filter effectively eliminated spikes and outliers caused by sudden, erroneous pixel movements without distorting the overall data trend. Following the outlier removal, the Savitzky-Golay filter [27] was employed to smooth the coordinates. Unlike traditional moving average filters, the Savitzky-Golay filter fits successive subsets of data to a low-degree polynomial, preserving critical signal characteristics such as peaks and troughs. This approach ensures that the essential motion dynamics are retained while minimizing the noise from minor fluctuations. Figure 1(b) shows the normalized angles and angular velocities post-filtering, demonstrating the enhanced signal clarity achieved through this preprocessing pipeline. Together, these filters ensured robust and noise-reduced data suitable for subsequent analyses.

In the second step, the generation of a plot with the normalized angle on the x-axis and the angular velocity on the y-axis is a key part of visualizing joint movements. Figure 1(c) displays two plots depicting the normalized angles and angular velocities of each joint. The relative Phase (RP) provides a meticulous method for analyzing the relationship between the angles and angular velocities of two joints using phase angles. This thoroughness enables a meticulous assessment of whether specific postures are consistently executed with a constant combination of angles and angular velocities at a given exercise intensity. When the same posture is repeated with identical angles and angular velocities, the resulting trajectory on the phase portrait resembles a circle of radius r_1.

In the third step, it is necessary to obtain the phase angle. The following equations represent the normalization process through equation 7 transformed by equations 4 and 5:

$$\phi_i^{rb} = \frac{\phi_i^{raw} - \mathbf{Q}_1(\phi^{raw})}{\mathbf{Q}_3(\phi^{raw}) - \mathbf{Q}_1(\phi^{raw})}, \tag{8}$$

$$\phi_{i'} = \frac{2 \times (\phi_i^{rb} - \phi_{min}^{rb})}{\phi_{max}^{rb} - \phi_{min}^{rb}} - 1 \tag{9}$$

Here, $\phi_{i'}$ denotes the normalized angle, ϕ_i^{raw} the original angle, ϕ_i^{rb} the angle after the robust filter is applied, and i the cycle point. In the interquartile range, $\mathbf{Q}_1$ signifies a median below the median, while $\mathbf{Q}_3$ represents a median above the median. The first and third rows in Figure 1(d) show the results of the two phase angles using Equation 9.

In the final step, it is necessary to obtain the relative phase angle of $\Delta\phi_{Angle}$ with ϕ_{hip} and ϕ_{knee}:

$$\Delta\phi_{Angle} = \phi_{hip} - \phi_{knee} \tag{10}$$

Figure 1(e) shows the results of $\Delta\phi_{Angle}$. When exercises are consistently performed under constant intensity, the trajectory radius on the phase portrait may vary owing to the individual physical capacity. In such cases, the $\Delta\phi_{Angle}$, expressed solely by the phase angle, has limitations in representing scenarios in which the radius changes.

D. RELATIVE PHASE DISTANCE

The relationship between the normalized angle and angular velocity provides the relative phase angle. The normalized angle and angular velocity are affected when the shape of the circle remains similar but the radius changes. If a shape close to a circle can be maintained, the combination of the angle and angular velocity is suitable. However, a change in the radius of the circle alters the ratio of this combination. This change in the combination ratio must be fully captured by the relative phase angle, highlighting the limitations of the current evaluation method. To address this, we urgently need a new metric called the relative phase distance, which considers the change in radius when evaluating the motion posture. This new metric is crucial to overcome the limitations of the current method and provide a more accurate evaluation of motion posture.

The following equation can be used to derive this metric from the relative phase plane.

$$\phi_i^{PD} = \sqrt{(\phi_{i'})^2 + (\omega_{i'})^2}, \tag{11}$$

$$\Delta\phi_{Distance} = \phi_{hip}^{PD} - \phi_{knee}^{PD}, \tag{12}$$

where $\Delta\phi_{Distance}$ represents the Phase Distance, ω represents the normalized angular velocity, and ϕ represents the normalized angular position. The second and fourth rows in Figure 1(d) show the results of ϕ_{hip}^{PD} and ϕ_{knee}^{PD} through Equation 11. The results for $\Delta\phi_{Distance}$ are shown in the bottom row in Figure 1(e).

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3551834

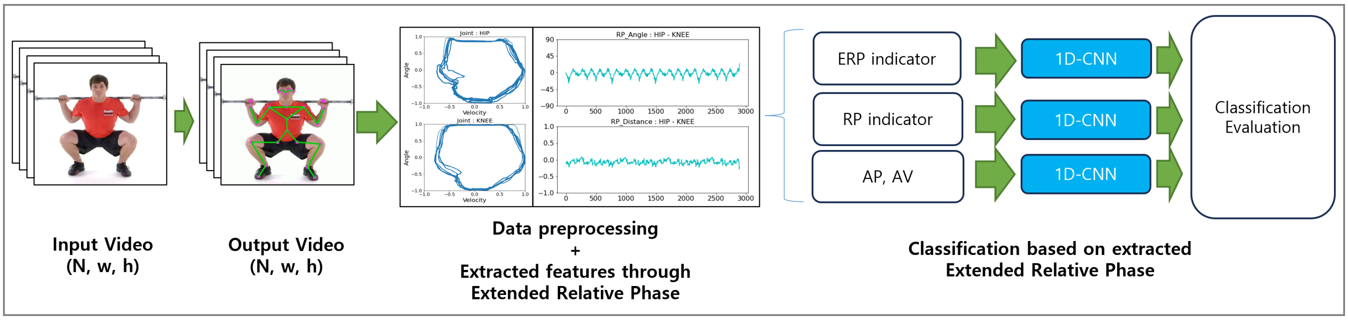


FIGURE 2. An overall process for experiment

TABLE 2. Personal and physical information for each participant

Subject	Gender	Age [years old]	BMI $[kg/m^2]$	Exercise Experience
Sub #1	M	27	28.37	3 years (sports)
Sub #2	M	27	22.60	4 years (free weight)
Sub #3	M	28	25.47	1 month (free weight)
Sub #4	M	30	22.79	NA
Sub #5	F	21	23.63	7 month (pilates)
Sub #6	F	25	24.56	6 month (pilates)
Sub #7	F	25	19.81	6 month (pilates)

IV. EXPERIMENTS

A. EXPERIMENTAL ENVIRONMENTS AND SYSTEMS

Figure 2 provides an overview of the data-analysis process. It describes the capture of user images, estimation of 3D coordinates for each joint, application of the proposed RP algorithm for feature extraction in velocity-based training, and classification of the differences between experts and users based on the extracted features.

Additionally, Table 2 summarizes the personal and physical information of seven non-professional participants who performed back-squat footage for data collection related to velocity-based training for novices. The footage was captured using an iPhone 12Pro camera with a resolution of 1920 $\times$ 1080 pixels and a frame rate of 30 fps in a configured environment. The distance between the camera and participants was set to 380($\pm$ 5) cm, and the camera height was fixed at 130 cm. The participants' heights varied slightly (mean $\pm$ standard deviation = 167.1 $\pm$ 8.7 cm).

During the study, the participants performed 20 squats in three sets with a 15 kg weight bar, taking 60-second breaks between each set and aiming to maintain a constant speed and movement.

The study complied with the Declaration of Helsinki and was approved by the Clinical Trial Center Ethics Committee, Department of Medical Innovation, Osaka University Hospital. All human participants provided informed consent before the experiment (no. 15408, 11 March 2016), and their athletic careers were also considered for experiment reliability.

B. EXTRACTED FEATURE THROUGH RELATIVE PHASE

The extended Relative Phase (RP) incorporates the $\Delta\phi_{Angle}$ and $\Delta\phi_{Distance}$. The $\Delta\phi_{Distance}$ represents the temporal relationship between the angular position and angular velocity, while the $\Delta\phi_{Distance}$ represents the consistency of the motion intensity. RP, extracted based on these two components, assesses the temporal coordination of a specific joint by introducing a temporal element to the interpretation of the two phases.

In our study, the manual analysis of RP involved categorizing the evaluation results of the participants' squat motions into several cases, as depicted in Figure 3. Upon reviewing videos of expert squats, it is considered that achieving a consistent graph form and range, such as in case #1 in Figure 3 in the RP plot, indicates good squat behavior.

Figure 3(a) shows the normalized angle and angular velocity. It is clear that discerning differences in participants' exercise posture based solely on these parameters is a formidable challenge. The task of analyzing time-series angle and angular velocity data, even for a small number of participants, is a daunting task for exercise coaches. As the number of participants increases, time-series data become overwhelming for human observation, and limited analysis methods are available. This underscores the urgent need for innovative solutions in this field. Our research is crucial for addressing this need and pushing the boundaries of exercise posture assessment methods.

Figure 3(b) shows the relative phase planes for each case. It is clear that the periodic shape of the phase plane varies depending on each subject's exercise posture, thereby influencing the relative phase results shown in Figure 3(d). The exercise posture of the expert can be evaluated using the relative phase plane formed by the joint angle and angular velocity. The radius and shape of the circle in the phase plane are crucial parameters for posture evaluation, underscoring the significance of our research in understanding and assessing exercise postures.

For instance, if the phase plane created by the angle and angular velocity of the expert is almost the same, the subject can be said to have an expert-level exercise posture. If the shape of the phase plane is the same but the radius is small,

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3551834

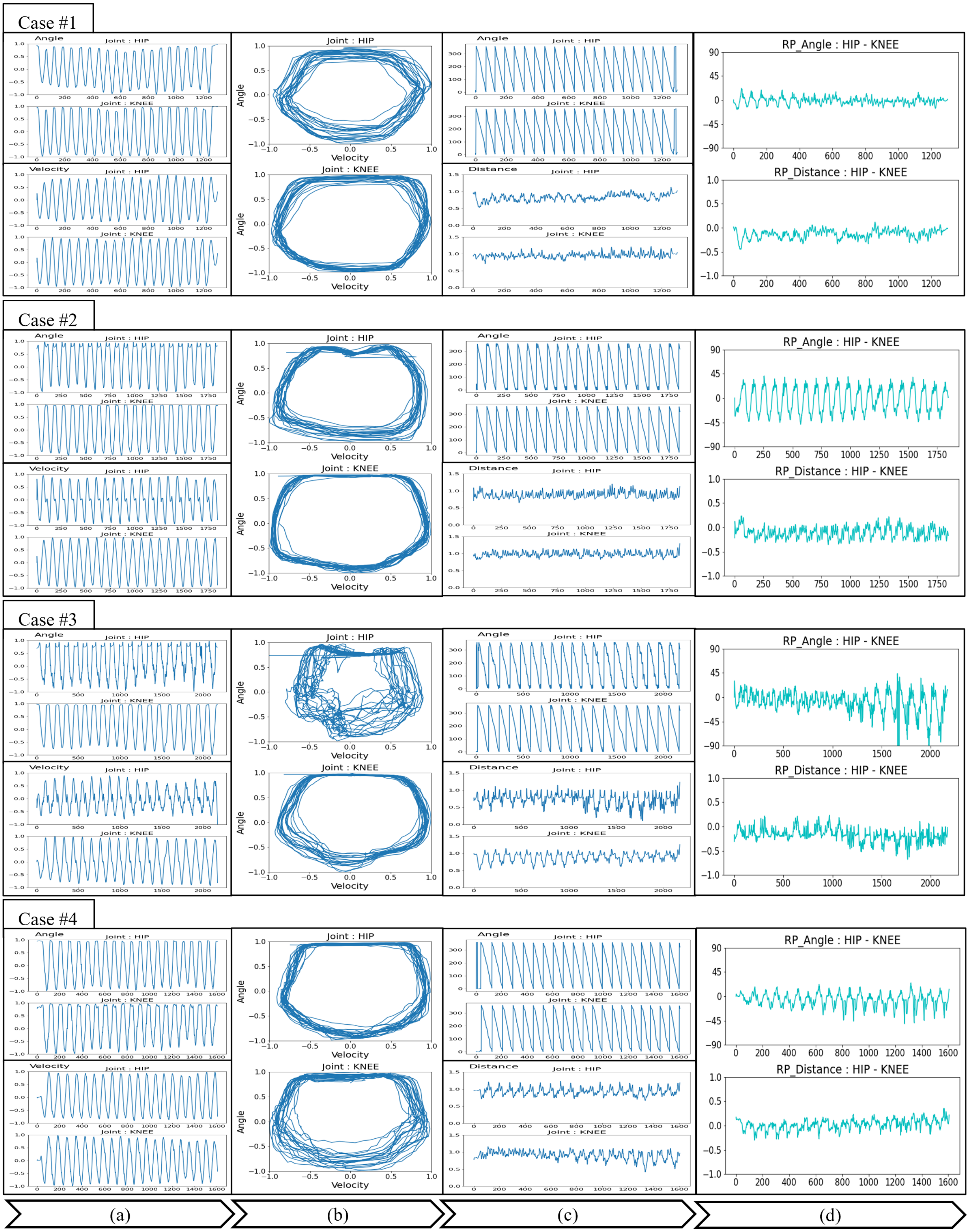


FIGURE 3. **Results of four different Cases: (a) represents normalized angular displacement and velocity, (b) represents the phase portrait, (c) represents the results of the phase angle (*PA*) and the phase distance (*PD*), (d) represents the results of the extended relative phase (ERP)**

it can be seen that the combination of the angle and angular velocity is appropriate. However, the angle and angular velocity size were slower than those of the expert. If the radius is similar but the shape of the circle is different, it can be seen that the combination of the angle and angular velocity is not appropriate. Thus, different phase planes, relative phase angles, and distances could be used to classify each class.

In Figure 3, Case #2 illustrates the outcomes of a non-expert participant who executed a similar exercise posture as the expert. The shape and radius of the circle closely resembled those in Case # 1. However, variations in the combination ratio of the joint angles and angular velocities resulted in a slightly different waveform for the relative phase angle.

In Case #3 in Figure 3, the results indicate a participant with low physical strength. The relative phase plane shows a non-constant shape and radius of the circle, which is particularly evident in the hip joint phase plane. The varying shape of the circle suggests a non-constant combination of the joint angle and angular velocity, whereas the changing radius indicates an inconsistent ratio of this combination. These changes in the circle shape and radius imply that the participants' joint dynamics were not consistent, which could lead to inefficient movement patterns and increased risk of injury. It is evident that although the participant initially matched an expert's angle and angular velocity ratio, as repetitions increased, the physical demands led to exercising with different angles and angular velocities.

In Figure 3, Case #4 illustrates the performance of a participant with limited physical strength. Inconsistency in the shape and radius of the circle in the relative phase plane are evident. Specifically, the phase plane of the ankle joint indicates non-constant angle and angular velocity combinations, leading to a non-continuous circle radius. Similar to Case #3, this participant initially matched the expert's angle and angular velocity ratio but gradually exhibited increased strength, exercising with different parameters as the repetitions progressed.

The proposed method facilitates velocity-based exercise posture evaluation and provides practical implications for sports science and biomechanics. In contrast to prior studies that relied on critical point extraction from visual information, our approach enabled the assessment of joint angles and angular velocities through the relative phase plane. Two RP features advance our comprehension of exercise posture and allow the evaluation of the correlation between exercise posture and intensity using the relative phase distance feature. This not only stimulates, but also encourages further research and application in this field, highlighting its potential for growth and development.

V. EFFECT OF RELATIVE-PHASED-BASED EXTRACTED FEATURES ON CLASSIFICATION

This study used relative phase analysis for classification, which yielded meaningful results. To validate these indicators in real-world scenarios and to enable automatic classification using artificial intelligence models, we employed a 1D-CNN-based deep learning classification model. The model was trained for four separate input difference categories and the results were subsequently analyzed. Leveraging a 1D Convolutional Neural Network (CNN) for sequence classification offers the advantage of learning directly from raw time-series data, allowing manual engineering of input features without domain-specific knowledge. If the model demonstrates satisfactory performance, it could be utilized for real-time evaluation of exercise posture, a potential breakthrough in the field of exercise science that could revolutionize the monitoring and improvement of exercise techniques.

A. DATA AUGMENTATION & PRE-PROCESSING

To enhance model robustness and address data scarcity, data augmentation was performed using strategic transformations while preserving the intrinsic characteristics of time-series data. Conventional augmentation techniques like jittering, rotation, and warping were avoided to maintain biomechanical integrity. Instead, each participant's angle (θ) and angular velocity (ω) were expanded by randomly sampling perturbations within physiologically reasonable limits, generating 1,000 augmented instances per participant. Gaussian noise was applied to simulate minor variations, and time-series stretching and contraction techniques were employed to model different movement speeds.

Subsequently, $\Delta\phi_{Angle}$ and $\Delta\phi_{Distance}$ were computed from the augmented data. Given their larger range compared to θ and ω, a MinMax filter was applied to normalize values between -1 and 1, mitigating data bias. Finally, the dataset was split into training and test sets in a 7:3 ratio, ensuring balanced class distribution for effective model generalization.

B. MODEL DESCRIPTION

The model used for training followed the architecture shown in Figure 4. This study customized the network framework by integrating Global Average Pooling (GAP) [28] into a 1D-CNN. The input data for the 1D CNN has a shape of (1300, f), where each sample consisted of 1300 time steps and f features. The number of features, f, varied across the experiments, with values of 8, 10, and 12. The model processes these data to classify participants into one of four target categories: professionals, pretty good, weak physical strength, and need more exercise experience.

The architecture comprises three Conv1D layers with filter sizes of 128, 256, and 128, and kernel sizes of 9, 6, and 3, respectively. Each Conv1D layer is followed by Batch Normalization to stabilize learning, ReLU activation to introduce non-linearity, and Dropout layers with rates of 0.3, 0.4, and 0.5 to prevent overfitting. A Global Average Pooling (GAP) layer reduces the temporal dimension by averaging feature maps and minimizing parameters while retaining critical information. Finally, a Dense layer with softmax activation outputs the class probabilities. This design effectively captures temporal patterns and high-level abstractions in the data, making it suitable for real-time time-series analysis. Combination of Conv1D layers, regularization techniques, and The GAP

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3551834

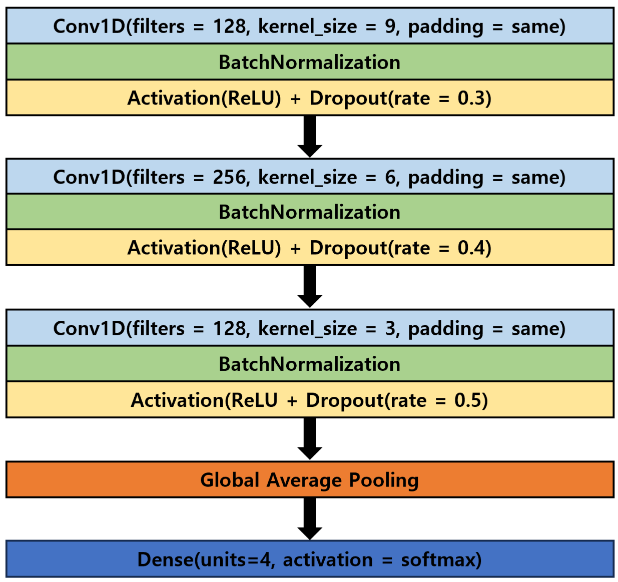
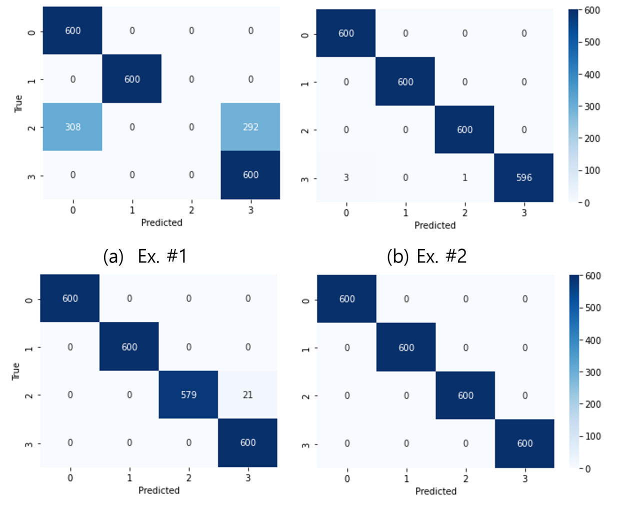


FIGURE 4. Model architecture



(a) Ex. #1 (b) Ex. #2 (c) Ex. #3 (d) Ex. #4

FIGURE 5. Confusion matrix of experimental results based on input features: The classes represent, in order, professionals, pretty good, weak physical strength, and need more exercise experience.

TABLE 3. Result of Experiments

Experiment	Features	Val auc & loss	test acc
Ex.#1 (Baseline)	θ, ω	[0.7500, 0.7500] [0.7620, 0.8072]	0.75
Ex.#2 (Comparable 1)	θ, ω, $\Delta\phi_{Angle}$	[0.9982, 0.9986] [0.4090, 0.4176]	0.9983
Ex.#3 (Comparable 2)	θ, ω, $\Delta\phi_{Distance}$	[0.9957, 0.9846] [0.5251, 0.5362]	0.9913
Ex.#4 (Best)	θ, ω $\Delta\phi_{Angle}$, $\Delta\phi_{Distance}$	[1.0, 1.0] [0.0016, 0.0004]	1.0

ensures that the model is lightweight, robust, and efficient for applications requiring accurate time-series classification.

To address data imbalances, we employed Stratified K-Fold cross-validation and dynamically adjusted the learning rate using the ReduceLROnPlateau strategy when validation loss plateaued. The initial learning rate was set to 0.001, and the learning rate was reduced by half if the loss did not improve over four consecutive epochs. Additionally, we incorporated an EarlyStopping mechanism to terminate training at an appropriate point, thereby preventing overfitting and unnecessary computation. A total of 20 epochs were allocated for the training.

Four experiments were conducted by varying the model input parameters. In Experiment 1, we utilized only eight features, excluding RP, using θ and ω for both the hips and knees. Experiment 2 added the existing $\Delta\phi_{Angle}$ to the inputs, resulting in 10 features. For Experiment 3, we introduced a new feature, $\Delta\phi_{Distance}$, to the existing eight features, resulting in 10 features used in the experiment. In the final experiment, we included $\Delta\phi_{Angle}$ and $\Delta\phi_{Distance}$, with a total of 12 features in the input. The results of four experiments were compared.

C. ANALYSIS RESULTS

The training and test results, which are of significant importance, are summarized in Table 3 based on the different input conditions. For the first condition, only angle and angular velocity were used. The relative phase angle features were then added to the angle and angular velocity. Following this, the relative phase distance features are included in addition to the angle and angular velocity. Finally, the relative phase angle and distance-derived features were used along with the angle and angular velocity. Table 3 presents the training, verification, and test results for the accuracy and loss for each input condition.

The results indicate that incorporating the proposed relative phase angle and distance features as input conditions ensures a high-accuracy performance. This is significant because it demonstrates the potential of these features to improve the accuracy of classification tasks in biomechanics and sports science. Furthermore, the loss value was closer to zero when both features were utilized as input conditions, indicating a more efficient and effective model.

The confusion matrix, which is a visual representation of the performance of the model, illustrates the test results in Figure 5. Figure 5(a) depicts the training utilizing only the angle and angular velocity as inputs, while Figure 5(b) shows the training incorporating relative phase angle features in addition to angle and angular velocity. Figure 5(c) illustrates the training integrating relative phase distance features alongside the angle and angular velocity, and Figure 5(d) demonstrates the training results using all input conditions. A confusion matrix is a valuable tool for evaluating a model's performance and understanding its strengths and weaknesses.

These findings indicate that classifying experts and proficient exercise participants can be effectively achieved using angles and angular velocities alone. However, for individuals who are not proficient in exercise and those exhibiting different postures from experts in the hip or ankle joints, it is essential to include the derived relative phase angle and distance as inputs, angles, and angular velocities to ensure high classification performance. These practical implications are crucial for the application of our research in real-world scenarios.

VI. CONCLUSION

This study introduced a novel Relative Phase (RP) indicator for evaluating velocity-based training, which effectively distinguished between expert and novice exercise postures. By integrating the $\Delta\phi_{Angle}$ and the newly proposed $\Delta\phi_{Distance}$ indices, the method demonstrated over 95% accuracy in assessing the performance and stability during dynamic exercises. The $\Delta\phi_{Distance}$ index adds a valuable dimension by capturing the consistency of periodic motion under varying intensities.

The proposed markerless approach offers a fast, scalable, and spatially unrestricted alternative to traditional sensor-based motion capture systems. This enables coaches and practitioners to assess joint coordination and exercise intensity without requiring specialized technical knowledge or complex equipment. Although RGB-based Human Pose Estimation (HPE) systems currently have limitations in terms of accuracy compared to sensor-based methods, their practical advantages make them suitable for real-world fitness, rehabilitation, and sports applications.

The findings of this study highlight the potential of integrating RP indicators into deep learning models to provide accurate real-time assessments. Future work will focus on scaling the proposed system for real-world applications by integrating it with existing fitness applications and developing real-time feedback mechanisms. Additionally, improvements in HPE models, such as leveraging transformer-based architectures, can enhance pose estimation accuracy, making the system more robust across diverse movement conditions. Further validation studies with larger and more diverse participant groups will be conducted to ensure broader applicability. Finally, developing a mobile-based version of the system could enable real-time assessment and feedback for users in non-laboratory settings, expanding its potential impact in fitness and rehabilitation domains.

REFERENCES

[1] W. B. Young, "Transfer of strength and power training to sports performance," *International journal of sports physiology and performance*, vol. 1, no. 2, pp. 74–83, 2006.
[2] M. R. McGuigan, G. A. Wright, and S. J. Fleck, "Strength training for athletes: does it really help sports performance?" *International journal of sports physiology and performance*, vol. 7, no. 1, pp. 2–5, 2012.
[3] B. Jidovtseff, N. K. Harris, J.-M. Crielaard, and J. B. Cronin, "Using the load-velocity relationship for 1rm prediction," *The Journal of Strength & Conditioning Research*, vol. 25, no. 1, pp. 267–270, 2011.
[4] P. Dohoney, J. A. Chromiak, D. Lemire, B. R. Abadie, and C. Kovacs, "Prediction of one repetition maximum (1-rm) strength from a 4-6 rm and a 7-10 rm submaximal strength test in healthy young adult males," *J Exerc Physiol*, vol. 5, no. 3, pp. 54–9, 2002.
[5] M. C. Zourdos, C. Dolan, J. M. Quiles, A. Klemp, E. Jo, J. P. Loenneke, R. Blanco, and M. Whitehurst, "Efficacy of daily one-repetition maximum training in well-trained powerlifters and weightlifters: a case series," *Nutricion hospitalaria*, vol. 33, no. 2, pp. 437–443, 2016.
[6] J. Weakley, B. Mann, H. Banyard, S. McLaren, T. Scott, and A. Garcia-Ramos, "Velocity-based training: From theory to application," *Strength & Conditioning Journal*, vol. 43, no. 2, pp. 31–49, 2021.
[7] M. O'Reilly, D. Whelan, C. Chanialidis, N. Friel, E. Delahunt, T. Ward, and B. Caulfield, "Evaluating squat performance with a single inertial measurement unit," in *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 2015, pp. 1–6.
[8] J. Lee, H. Joo, J. Lee, and Y. Chee, "Automatic classification of squat posture using inertial sensors: Deep learning approach," *Sensors*, vol. 20, no. 2, p. 361, 2020.
[9] Y. Woo, S. Ko, S. Ahn, H. T. P. Nguyen, C. Shin, H. Jeong, B. Noh, M. Lee, H. Park, and C. Youm, "Classification of diabetic walking for senior citizens and personal home training system using single rgb camera through machine learning," *Applied sciences*, vol. 11, no. 19, p. 9029, 2021.
[10] J. Qi, P. Yang, M. Hanneghan, A. Waraich, and S. Tang, "A hybrid hierarchical framework for free weight exercise recognition and intensity measurement with accelerometer and ecg data fusion," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 3800–3804.
[11] N. Seethapathi, S. Wang, R. Saluja, G. Blohm, and K. P. Kording, "Movement science needs different pose tracking algorithms," *arXiv preprint arXiv:1907.10226*, 2019.
[12] D. Roetenberg, H. Luinge, P. Slycke *et al.*, "Xsens mvn: Full 6dof human motion tracking using miniature inertial sensors," *Xsens Motion Technologies BV, Tech. Rep*, vol. 1, pp. 1–7, 2009.
[13] V. M. S. Ltd, "What is motion capture?: What can i use motion capture for?" https://www.vicon.com/about-us/what-is-motion-capture/, accessed: 2023-03-17.
[14] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7291–7299.
[15] D. Pavllo, C. Feichtenhofer, D. Grangier, and M. Auli, "3d human pose estimation in video with temporal convolutions and semi-supervised training," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 7753–7762.
[16] V. Bazarevsky, I. Grishchenko, K. Raveendran, T. Zhu, F. Zhang, and M. Grundmann, "Blazepose: On-device real-time body pose tracking," *arXiv preprint arXiv:2006.10204*, 2020.
[17] H. T. P. Nguyen, Y. Woo, N. N. Huynh, and H. Jeong, "Scoring of human body-balance ability on wobble board based on the geometric solution," *Applied Sciences*, vol. 12, no. 12, p. 5967, 2022.
[18] A. Lamas, S. Tabik, A. C. Montes, F. Pérez-Hernández, J. García, R. Olmos, and F. Herrera, "Human pose estimation for mitigating false negatives in weapon detection in video-surveillance," *Neurocomputing*, vol. 489, pp. 488–503, 2022.
[19] J. Liu, M. Shi, Q. Chen, H. Fu, and C.-L. Tai, "Normalized human pose features for human action video alignment," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 11 521–11 531.
[20] J. Chua, L.-Y. Ong, and M.-C. Leow, "Telehealth using posenet-based system for in-home rehabilitation," *Future Internet*, vol. 13, no. 7, p. 173, 2021.
[21] Y.-C. Li, C.-T. Chang, C.-C. Cheng, and Y.-L. Huang, "Baseball swing pose estimation using openpose," in *2021 IEEE International Conference on Robotics, Automation and Artificial Intelligence (RAAI)*. IEEE, 2021, pp. 6–9.
[22] VAY, "The 1 human motion analysis for fitness health: Vay: Switzerland," https://www.vay.ai/.
[23] K. H. US, "Democratizing healthcare: Kaia health us," https://kaiahealth.com/.
[24] A. AI, "Artificial intelligence fitness coach: Alfa ai," https://www.infivolve.com/.
[25] Infivolve, "infigro | infivolve," https://www.infivolve.com/.

[26] R. K. Pearson, Y. Neuvo, J. Astola, and M. Gabbouj, "Generalized hampel filters," *EURASIP Journal on Advances in Signal Processing*, vol. 2016, no. 1, pp. 1–18, 2016.

[27] R. W. Schafer, "What is a savitzky-golay filter?[lecture notes]," *IEEE Signal processing magazine*, vol. 28, no. 4, pp. 111–117, 2011.

[28] M. Lin, Q. Chen, and S. Yan, "Network in network," *arXiv preprint arXiv:1312.4400*, 2013.

YEONGJU WOO received the B.S. in Software Engineering from Chonnam National University, South Korea, in 2021, and the M.S. degree in Artificial Intelligence Convergence from Chonnam National University, South Korea, in 2023. Since March 2023, he has been a Ph.D. candidate in the Department of Artificial Intelligence Convergence, Chonnam National University, Gwangju, South Korea.

His research has focused on image processing, healthcare, time-series data analysis, and pattern recognition. His primary research involves the implementation of feature extraction and evaluation algorithms for sports assessment through image analysis.

HIEYONG JEONG received the M.S in Robotics from the Hiroshima University, Japan, in 2006 and the Ph.D. degree in Mechanical Engineering from Osaka University, Japan, in 2009.
From 2009 to 2013, he was a Senior Research Engineer (full-time) with SAMSUNG Heavy Industries, Co., Ltd, Daejeon, South Korea. From 2014 to August 2019, he was an Associate Professor (full-time) the Department of Robotics and Design for Innovative Healthcare, Graduate School of Medicine, Osaka University, Suita, Japan. Since September 2019, he has been a Professor (full-time) with the Department of Artificial Intelligence Convergence, Graduate School of Engineering, Chonnam National University, Gwangju, South Korea.
His research focuses on IoT sensors and robotic systems designed to capture and process large-scale data streams in real-time across various sectors, including manufacturing, healthcare, aerospace, defence, transportation, intelligent urban infrastructures, agriculture, and fisheries. Additionally, he investigates the application of artificial intelligence algorithms for pattern recognition and data mining, enabling users to derive actionable insights from complex datasets.