UDC 004.8:004.93:796
IRSTI 28.23.15
https://doi.org/10.47526/2025-4/2524-0080.45

Y.U. SERDALIYEV[1], N.T.TOILYBAYEVA[2], B.O. BAYMURZAEV[3], H. CAKIR[4]
Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkestan, Kazakhstan
[1]E-mail: erlan.serdaliev@ayu.edu.kz,*
[2]E-mail: nurgul.toilybayeva@ayu.edu.kz,
[3]E-mail: bboranbek@ayu.edu.kz
[4]Gazi University, Department of Computer and Instructional Technology Education, Ankara, Turkey
[4]E-mail: hcakir@gmail.com

CONCEPTUAL FOUNDATIONS OF REAL-TIME MONITORING OF PHYSICAL EXERCISES BASED ON COMPUTER VISION AND DEEP LEARNING MODELS

Annotation. The article presents a challenge in automatically monitoring and evaluating physical exercises in real time. The goal of this study is a comparative analysis of the effectiveness of modern deep learning models for determining human posture (BlazePose, YOLOv8 - Pose, MoveNet and HRNet) and to identify the possibilities of their inclusion in the training quality assessment system. The system is built on an RGB video stream and consists of pre-processing, posture points determination, joint angles calculation and training classification (correct/incorrect) stages. From the experimental results, it can be found that HRNet model has the highest accuracy index, but as a result of its high calculation complexity, delay value increases. It had been shown that BlazePose provides high frame rates at low-resource devices. The YOLOv8-Pose model showed the best compromise between accuracy and performance and was suggested as an optimal option for Real-Time Fitness Systems. The method of biomechanical analysis based on calculating joint angles allows to quantify the quality of exercise performing has been proved. The systematic alignment score may be applied in the scope of fitness, sports training and medical rehabilitation and can further be enhanced by adding on 3D body pose models as well as spatio-temporal neural networks.

Keywords: human posture determination; computer vision; deep learning; physical exercise monitoring; real-time systems; calculation of joint angles; YOLOv8-Pose; BlazePose; biomechanical analysis.

Е.У. СЕРДАЛИЕВ[1], Н.Т.ТОЙЛЫБАЕВА[2], Б.О. БАЙМУРЗАЕВ[3], Х. ЧАКЫР[4]
Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан қ., Қазақстан
*[1]E-mail: erlan.serdaliev@ayu.edu.kz**
[2]E-mail: nurgul.toilybayeva@ayu.edu.kz
[3]E-mail: bboranbek@ayu.edu.kz
[4]Гази университеті, Компьютер және оқыту технологиялары білім беру кафедрасы, Анкара, Турция
[4]E-mail: hcakir@gmail.com

КОМПЬЮТЕРЛІК КӨРУ ЖӘНЕ ТЕРЕҢ ОҚЫТУ МОДЕЛЬДЕРІ НЕГІЗІНДЕ ФИЗИКАЛЫҚ ЖАТТЫҒУЛАРДЫ НАҚТЫ УАҚЫТ РЕЖИМІНДЕ БАҚЫЛАУДЫҢ КОНЦЕПТУАЛДЫҚ НЕГІЗДЕРІ

Андатпа. Мақалада нақты уақыт режимінде физикалық жаттығуларды автоматты түрде бақылау және бағалаудағы қиындық қарастырылған. Бұл зерттеудің мақсаты - адамның дене бітімін анықтауға арналған заманауи терең оқыту модельдерінің (BlazePose, YOLOv8 - Pose, MoveNet және HRNet) тиімділігін салыстырмалы талдау және оларды жаттығу сапасын бағалау жүйесіне қосу мүмкіндіктерін анықтау. Жүйе RGB бейне ағынына негізделген және алдын ала өңдеу, дене бітімінің нүктелерін анықтау, буын бұрыштарын есептеу және жаттығуды жіктеу (дұрыс/дұрыс емес) кезеңдерінен тұрады. Тәжірибелік нәтижелерден HRNet моделінің ең жоғары дәлдік индексі бар екені анықталды, бірақ оның жоғары есептеу күрделілігінің нәтижесінде кідіріс мәні артады. BlazePose төмен ресурстарды қажет ететін құрылғыларда жоғары кадр жиілігін қамтамасыз ететіні көрсетілді. YOLOv8-Pose моделі дәлдік пен өнімділік арасындағы ең жақсы ымыраны көрсетті және нақты уақыт режиміндегі фитнес жүйелері үшін оңтайлы нұсқа ретінде ұсынылды. Буын бұрыштарын есептеуге негізделген биомеханикалық талдау әдісі жаттығуды орындау сапасын сандық бағалауға мүмкіндік беретіні дәлелденді. Жүйелі туралау ұпайы фитнес, спорттық жаттығулар және медициналық оңалту саласында қолданылуы мүмкін және 3D дене позасы модельдерін, сондай-ақ кеңістіктік-уақыттық нейрондық желілерді қосу арқылы одан әрі жақсартылуы мүмкін.

Кілт сөздер: адамның позасын анықтау; компьютерлік көру; терең оқыту; дене жаттығуларын бақылау; нақты уақыт жүйелері; буын бұрыштарын есептеу; YOLOv8-Pose; BlazePose; биомеханикалық талдау.

Е.У. СЕРДАЛИЕВ[1], Н.Т.ТОЙЛЫБАЕВА[2], Б.О. БАЙМУРЗАЕВ[3], Х. ЧАКЫР[4]

[1,2,3]Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, Туркестан, Казахстан
*[1]E-mail: erlan.serdaliev@ayu.edu.kz**
[2]E-mail: nurgul.toilybayeva@ayu.edu.kz
[3]E-mail: bboranbek@ayu.edu.kz
[4]Университет Гази, факультет компьютерных и образовательных технологий, Анкара, Турция
[3]E-mail: hcakir@gmail.com

КОНЦЕПТУАЛЬНЫЕ ОСНОВЫ МОНИТОРИНГА ФИЗИЧЕСКИХ УПРАЖНЕНИЙ В РЕАЛЬНОМ ВРЕМЕНИ НА ОСНОВЕ МОДЕЛЕЙ КОМПЬЮТЕРНОГО ЗРЕНИЯ И ГЛУБОКОГО ОБУЧЕНИЯ

Аннотация. В статье рассматривается задача автоматического мониторинга и оценки физических упражнений в реальном времени. Целью данного исследования является сравнительный анализ эффективности современных моделей глубокого обучения для определения осанки человека (BlazePose, YOLOv8-Pose, MoveNet и HRNet) и выявление возможностей их включения в систему оценки качества тренировки. Система построена на основе RGB-видеопотока и включает этапы предварительной обработки, определения точек осанки, расчета углов суставов и классификации тренировки (правильно/неправильно). Из экспериментальных результатов можно установить, что модель HRNet имеет самый высокий показатель точности, но из-за высокой вычислительной сложности увеличивается значение задержки. Было показано, что BlazePose обеспечивает высокую частоту кадров на устройствах с ограниченными ресурсами. Модель YOLOv8-Pose показала наилучший компромисс между точностью и производительностью и была предложена в качестве оптимального варианта для фитнес-систем реального времени. Доказан метод биомеханического анализа, основанный на расчете углов суставов, позволяющий количественно оценить качество выполнения упражнений. Систематическая оценка выравнивания может применяться в сфере фитнеса, спортивной подготовки и медицинской реабилитации, а также может быть дополнительно улучшена за счет добавления 3D-моделей позы тела и пространственно-временных нейронных сетей.

Ключевые слова: определение позы человека; компьютерное зрение; глубокое обучение; мониторинг физических упражнений; системы реального времени; расчет углов суставов; YOLOv8-Pose; BlazePose; биомеханический анализ.

Introduction

Indeed, the topic of monitoring physical activity and health via digital technologies has taken on particular significance in recent years. The need for monitoring motion correctness in fitness, sport training and medical rehabilitation is critical due to its high importance regarding injury prevention, movement quality improvement or at least ensuring the effectiveness of the rehab process. The conventional monitoring is often on the basis of a trainer or specialist visually assessing performance, which results in subjectivity, time consumption and lower repeatability (Jo & Kim, 2023).

In particular, the evolution of computer vision technologies and deep learning has enabled automatic analysis of techniques for physical exercises. Human pose estimation techniques enable us to spot a person in an image and find the key joint points of his body, thereby depicting movement in the form of a skeletal model. According to Huang et al. (2025), this approach is suitable for solving tasks such as joint angles, trajectory analysis of movement and whether the movement of performing an exercise is correct or incorrect.

There are a number of models available for identifying pose in humans today. Mobile and light-weight models with processing time in real time like BlazePose or MoveNet in MediaPipe can be used on mobile or other resource-poor devices (Opiña Jr. & Fajardo, 2024). Detection-based YOLOv8-Pose architectures enable arising multiple pose tracking with high efficacy in real-time systems (Dong & Du, 2024; Cai et al., 2025). Moreover, HRNet and transformer-based models have very high accuracy in detecting pose point but much higher computation costs (Li et al., 2023).

Numerous analyses regarding exercise assessment, action recognition and motion analysis through pose detection models are available in the scientific literature. Specifically, spatio-temporal model based on skeletal data has achieved effective results in the assessment of exercise performance quality (Xie et al., 2024). Furthermore, basic properties of real-time systems and their pros and cons for low-resource devices have been studied in review papers (Huang & Li, 2024; Zheng et al., 2023).

However, there is still no systematic conceptual framework for embedding state-of-the-art posture recognition architectures into motion control databases. Models with different accuracy, speed and computational resource requirements must be systematically compared and adapted to application scenarios. This article introduces a conceptual approach for realtime physical exercise analysis, based on posture recognition models, and discusses its scientific and practical implications.

Materials and methods

Dataset Description

This study used a specially collected video dataset to monitor and evaluate physical exercise in real time. The dataset consists of a total of 200 videos and includes exercises performed by 10 participants. The participants included 5 men and 5 women. This balanced structure was aimed at reducing the dependence of the model on gender.

Five different physical exercises were considered in the study: squat, push-up, deadlift, jumping jack and bent-over row. These exercises involve different parts of the body and differ in movement dynamics. Therefore, their selection allows us to assess the ability of the model to distinguish between different movement patterns.

Each video clip was approximately 6–12 seconds long, which was enough to capture one or more complete movement cycles. The videos were recorded in three different environments: a gym, a laboratory room and a home environment. In addition, the footage was shot from two different angles: a front view and a side view. Such diversity allows us to assess the generalization ability of the model despite changes in lighting, background, camera angle, and spatial conditions.

Thus, the dataset used had sufficient diversity for the tasks of recognizing physical exercises and assessing the quality of their execution.

Data partitioning principle

In the experimental evaluation, the dataset was divided into training, validation, and testing sets according to the classical supervised learning approach. The partitioning ratio was chosen as 70%-training set, 15%-validation set, and 15%-testing set, respectively.

The training set was used to train the model parameters, the validation set was used to tune hyperparameters and control overfitting, and the testing set was kept for independent evaluation of the final model performance.

During data partitioning, the distribution of different training types, shooting environments, and camera angles was kept as equal as possible. This approach prevents the overpredominance of a single category, ensuring the fairness and stability of the evaluation results. In addition, the results obtained during the testing period characterize the model's ability to generalize to new data.

Futhermore, This study investigated several contemporary deep learning models of human pose detection for monitoring physical exercise in real time. The main criteria used to select models were real-time performance (FPS), computational complexity, detection accuracy and the possibility of running on low-resource devices (smartphone, webcam, low-end laptop). Objective: The objective of this study is to compare the strengths and the weaknesses of different architectures in order to identify which group of models can be used as a formula for building physical exercises monitoring systems. To use computer vision for pose detection, we primarily adapted BlazePose from the MediaPipe platform since it is lightweight and very efficient on mobile devices. The BlazePose Model can run on stable real time and detect more than 30 key anatomical points of a human body. The key benefits of this model are low computational load and high frame rate on mobile devices. Due to this, BlazePose is useful for home fitness training and mobile rehabilitation systems.

We used the YOLOv8-Pose architecture to operate in a crowded space and in complex scenes. This system integrates object detection and pose regression tasks into a neural network. High real-time performance allowing multi-person tracking as seen in gyms or group exercises. Moreover, this architecture is also highly scalable and hardware platform agnostic. Also, the MoveNet and HRNet architectures have been considered for comparison purposes and methodology analysis as they are classified into high-precision models. The MoveNet is available in easily two different models (Lightning and Thunder), both providing excellent accuracy and robustness. Despite the great accuracy achieved for pose points detection with HRNet architecture, it has a high computational complexity (Table 1) limiting its usage in real-time mode on low-resource devices.

Table 1. Comparative characteristics of the considered pose detection models

Model	Main Feature	FPS (approx.)	Accuracy	Resource Requirement	Application
BlazePose (MediaPipe)	Lightweight, mobile-friendly	25–30	Medium	Low	Mobile fitness
YOLOv8-Pose	Detection + Pose estimation	20–30	High	Medium	Multi-person scenarios
MoveNet	High stability	15–25	High	Medium	Individual exercise
HRNet	High precision	5–10	Very high	High	Laboratory analysis

The BlazePose model is aimed at real-time operation on mobile devices using a lightweight encoder-decoder architecture. The HRNet model allows for accurate pose point detection by storing high-resolution features, but its computational complexity is high (Figure 1).

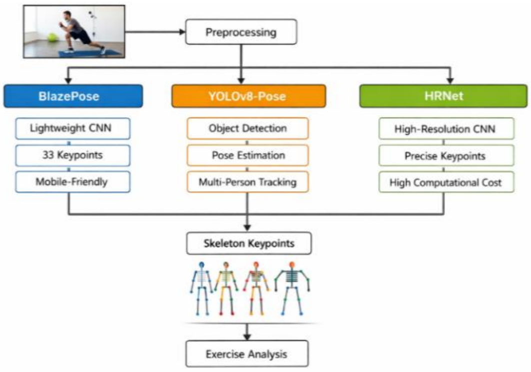

Figure 1. General architecture of pose detection models used in a physical exercise tracking system.

BlazePose model runs at high frame rate but with a low accuracy. The YOLOv8-Pose architecture provides a good balance of performance and accuracy, and does well in a high number of people. Although the MoveNet model offers a high accuracy with stable performance, HRNet architecture provides the top-line accuracy score, but runs at low frame rate because of its computation complexity. As depicted (Fig. 2), this means there is a clear trade-off between performance and accuracy in real-time system model selection.

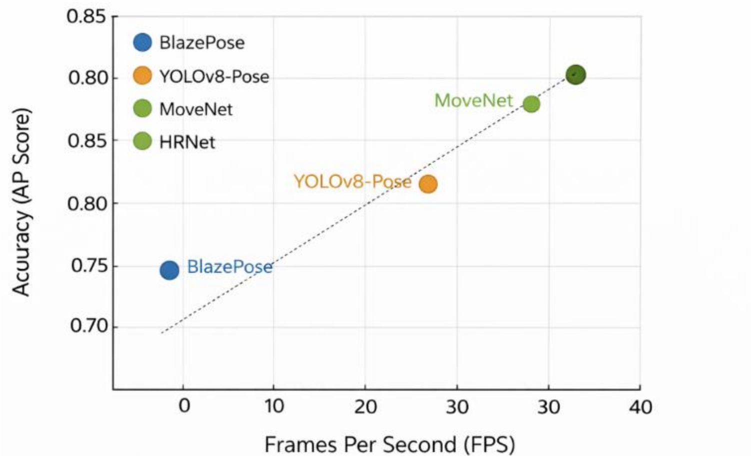

Figure 2. Comparative chart of performance and accuracy (FPS vs Accuracy) for different models.

This study retrieves data for physical exercise analysis as video streams in RGB format. The videos were taken with a webcam or smart-phone camera and modified for immediate processing. The camera was set up to fully capture the individual from the side or front thus, ensuring accurate and stable pose points detection. Wir haben Videodaten mit 25-30 FPS, 720p oder 1080p aufgezeichnet.

The obtained video streams passed through multiple preprocessing steps before getting as input to the pose detection algorithms. All these stages were performed to make the system more stable and the models more computationally efficient.

The video stream was received sliced into single frames first:

$$V = \{F_1, F_2, \dots, F_T\}, \quad (1)$$

where: F_t – t- time-lapse video, and T – number of frames.

In the next step, each frame was resized according to the model requirements. For example, for the BlazePose and YOLOv8-Pose models, the input dimensions were converted to 256×256 and

640×640 pixels, respectively. This operation reduces the computational load and ensures real-time processing.

Pixel values were normalized to the range[0,1]:

$$I_{norm} = \frac{I}{255}, \tag{2}$$

where: I – original pixel value.

The frames were then subjected to light filtering and contrast correction methods in order to dampen noise and mitigate differences from changing lighting conditions. Also, temporal smoothing techniques were explored to smooth abrupt changes between frames in the video stream.

In last step of preprocess, frames were inputed to the model and human pose's key points are extracted. Therefore, this preprocessing phase is a crucial step that directly impacts the whole accuracy and stability of the entire system (see Figure 3).

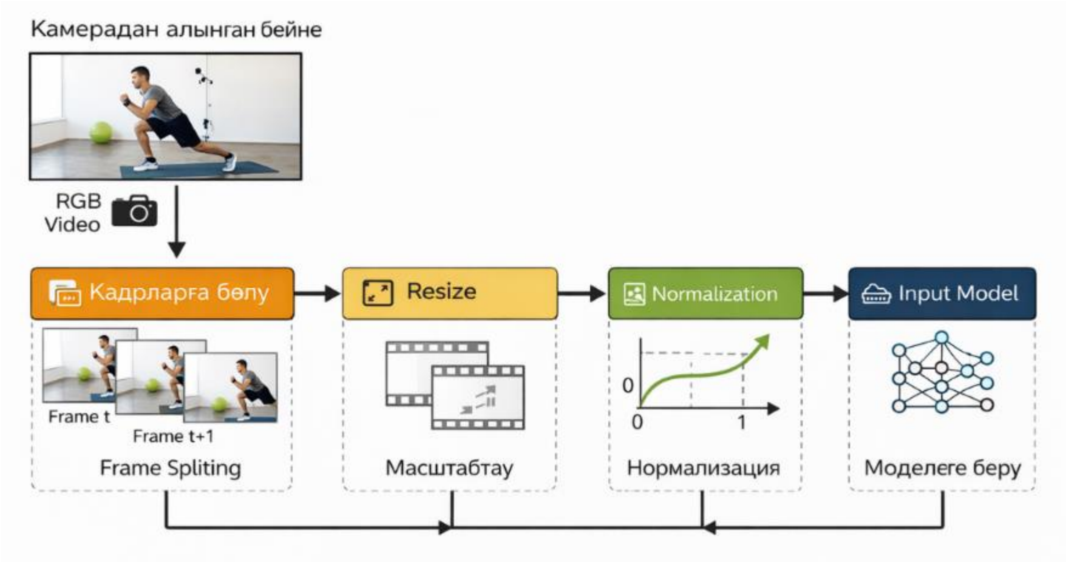

Figure 3. Data acquisition and preprocessing stages in a pose detection system.

The proposed data acquisition and preprocessing methods are adapted to real-time physical exercise monitoring systems. This approach ensures uniformity of input data when using different pose detection models and increases the overall performance of the system.

Pose Point Detection Method

In this study, the human pose detection method is based on finding keypoints of a human body from video frames. Pose points are considered as a set of coordinates describing the structure of the human skeleton and are used for further biomechanical and kinematic analysis. The considered models (BlazePose, YOLOv8-Pose, MoveNet and HRNet) detect pose points in 2D space.

Formal description of pose points

For each video frame, the model produces a set of pose points in the following form:

$$P^t = \{(x_1^t, y_1^t), (x_2^t, y_2^t), \dots, (x_N^t, y_N^t)\}, \tag{3}$$

where:
t– time period (frame index),
N– the number of pose points to be determined,
(x_i^t, y_i^t)– i-2D coordinates of the joint point.

For example, The BlazePose model detects N=33 points, while the YOLOv8-Pose and HRNet models typically detect N=17 points.

The principle of pose point regression

Pose detection models solve the problem of regressing the P^t points from the input frame using a neural network. This process is described as follows:

$$P^t = f_\theta(F_t), \tag{4}$$

where:

F_t – pre-processed input frame,

$f_\theta(\cdot)$ – a deep neural network with parameters θ.

In the YOLOv8-Pose architecture, this process simultaneously performs person detection and pose point detection, while the BlazePose and HRNet models use a direct pose regression approach.

Confidence indicator

Many models provide a confidence coefficient (c_i) for each pose point:

$$P^t = \{(x_i^t, y_i^t, c_i^t)\}_{i=1}^N, c_i^t \in [0,1] \tag{5}$$

This indicator allows us to assess the reliability of point detection. Points with low reliability can be excluded from the analysis or corrected by temporal smoothing.

Representation as a time series

Since physical exercises are dynamic, pose points are considered as a time series, not individual frames:

$$S = \{P^1, P^2, \dots, P^T\}. \tag{6}$$

This sequence describes the continuity of movement and is further used in algorithms for calculating joint angles, determining movement phases, and evaluating exercise.

Skeleton graph

Position points are represented as a skeleton graph using predefined anatomical connections:

$$G = (V, E), \tag{7}$$

where:

V – points of pose (vertices),

E – connections between joints (edges).

This representation approach is effective in spatiotemporal analysis and the use of graph neural networks (ST-GCN) (Figure 4).

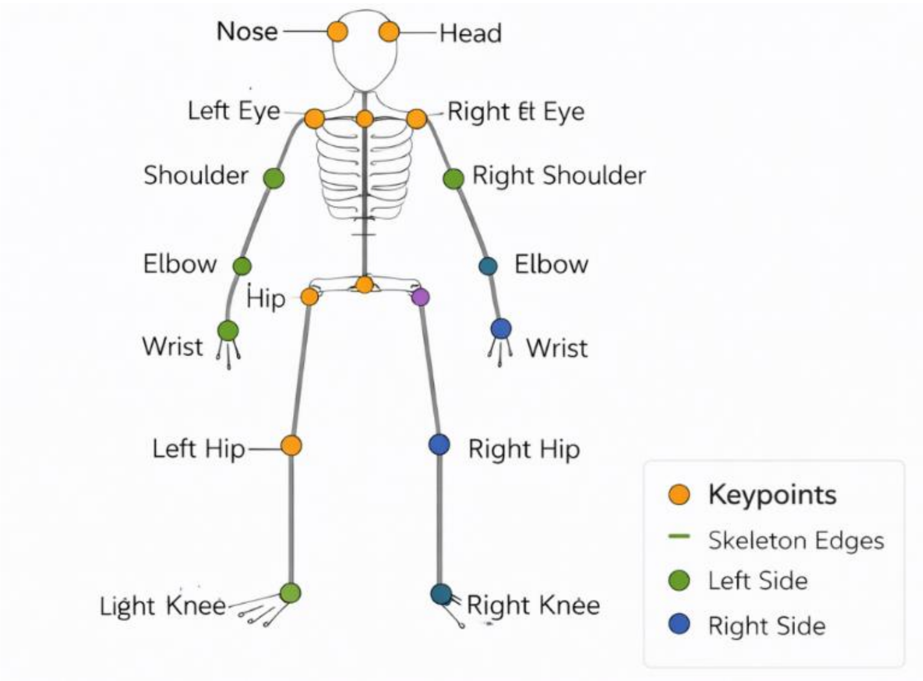

Figure 4. Skeleton view of pose points (keypoints).

Implementation details

The proposed system was implemented in the Python programming language. The OpenCV library was used for video data processing and pre-analysis, and various tools were used to detect human poses, depending on the corresponding models.

The BlazePose model was implemented using the MediaPipe platform. This model is adapted to work in real-time at high speed and shows effective results on CPU-based systems. The YOLOv8-Pose model was implemented based on the Ultralytics YOLOv8 library and implemented human detection and pose point detection in a single architecture. In addition, the MoveNet model was implemented based on TensorFlow Lite, and the HRNet model was implemented using the PyTorch framework.

The experiments were conducted in a CPU-based computing environment. This approach was aimed at evaluating the ability of the proposed system to work on low-resource devices. The processing speed and latency values for each model were measured on a real-time video stream.

Method for calculating joint angles

One of the main indicators in assessing the correct performance of physical exercises is the joint angle. Joint angle is a biomechanical parameter that characterizes the amplitude, phase, and technical correctness of human movement. In this study, joint angles were calculated based on 2D pose points.

Determining an angle using three points

Any joint angle is determined using three points: for example, for a knee angle – hip (A), knee (B), ankle (C).

The joint angle is calculated using the following vector formula:

$$\theta = \cos^{-1}\left(\frac{\vec{BA} \cdot \vec{BC}}{\| \vec{BA} \| \| \vec{BC} \|}\right) \tag{8}$$

where:

$\vec{BA} = A - B$

$\vec{BC} = C - B$

$\cdot$ – scalar product

$\| \vec{v} \|$ – vector length

Writing vectors in coordinate form:

$$\vec{BA} = (x_A - x_B, y_A - y_B) \tag{9}$$

Scalar product:

$$\vec{BA} \cdot \vec{BC} = (x_A - x_B)(x_C - x_B) + (y_A - y_B)(y_C - y_B) \tag{11}$$

The resulting angle θ is measured in radians or degrees.

Key joints

In the biomechanical study of physical exercises joint angles is accepted as the main variable describing moving quality [3]. The main load in any exercise always lies on specific anatomical joints, and this is precisely why the correctness of the movement directly depends on the spatial location and angular changes of these very same joints. Thus, the following key joints were chosen to be analyzed based on exercise types in this study.

First of all, the knee joint is a main biomechanical component reflecting lower limbs movement. It is especially crucial in squat and lunge movements. The flexion and extension angle of the knee joint during these exercises regulates the depth, amplitude and correct technical execution of this movement. While insufficient flexion of the knee angle can suggest improper execution of the exercise, excessive forward displacement can also overload the joint. Thus, the next key indicator of squat and lunge exercise quality is tracking dynamic changes in the knee angle.

Second, Analysis of the elbow joints push up exercise as the main movement mechanism of the upper limbs. The elbow angle in a push-up describes how low the body went and how symmetric the movement is. Excessive movement, asymmetrical movements and not achieving sufficient flexion are considered technical errors. So the elbow is the primary point of control for your upper body strength movements.

Thirdly, the hip joint plays a pivotal role in this exercise, being part of the body's primary support system during deadlifting. Particulars of Weightlifting – The Deadlift Movement At the hip joint, this is based on coordinated work biomechanics are joint-based which start with a deadlift. When the hip angle is correct, it gives us what should be an acceptable level of anterior pelvic tilt and a benign position when looking in the lumbar spine. It is the main metric for determining whether proper biomechanical technique has been followed.

Fourth, once again as upper extremity movement articulations are considered namely, the shoulder during bent-over row exercise. During this exercise, the shoulder joint, together with the elbow,DOMINATES back muscle activation. Monitors the shoulder angle Alpha, which is essential for effective exercise and preventing overload.

Therefore, the angular parameters of biomechanically essential joints were selected for each exercise to be the primary parameter of calculation to objectively evaluate exercise performance. This approach enables a systematic analysis of physical exercises and objective evaluation of movement technique. First, key joints were identified for every exercise.

Angle change over time

Since physical exercises are dynamic in nature, the dependence of the angle on time is considered:

$$\theta(t) = f(P^t) \tag{12}$$

Joint angle speed:

$$\omega(t) = \frac{\theta(t) - \theta(t-1)}{\Delta t} \tag{13}$$

This indicator allows you to determine whether the movement is performed too quickly or too slowly.

Criteria for correct execution

For each exercise, the permissible range of angles is set:

$$\theta_{min} \leq \theta(t) \leq \theta_{max} \tag{14}$$

For example:

Knee angle during squat: $70° - 110°$

Elbow angle during push-ups: $60° - 100°$

If the angle is out of range, the system marks the exercise as incorrect.

The calculation of joint angles is based on a kinematic model of human movement. While 2D analysis does not replace full 3D biomechanics, it provides sufficient accuracy for exercises in the frontal or sagittal plane. In real-time systems, the 2D approach is computationally efficient and mobile-friendly.

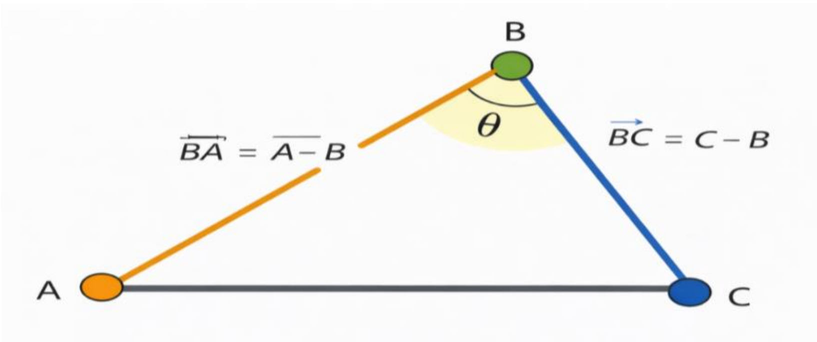

Figure 5. Scheme for calculating the joint angle using three points

Results

Experimental evaluation of the proposed real-time physical exercise tracking system was done based on these models BlazePose, YOLOv8-Pose, MoveNet and HRNet. All models investigated were tested in the same hardware environment (CPU + mid-range GPU) and the same video data was applied. The pose detection accuracy, the correct/incorrect recognition quality for each exercise, frame rate (FPS), and system latency were used as evaluation criteria. The system was found to work well during real time operation By combining the processes of preprocessing, pose point detection, and joint angle calculations, the overall latency was minimized. (a) Products for practical use modeling indicated that the architectural structure shown in Figure 1 and Figure 3 was confirmed to be a solid/working motion seen as effective as a real copiar.

The skeletal structure used is shown in Figure 4 to obtain visual examples for the pose points. All models successfully localized the main keypoints of human body, but accuracy differences could be observed.

Pose Detection QualityThe quality of pose detection was reported with AP (Average Precision) indicator. HRNet achieved the highest accuracy but incurred a high computational cost. In contrast YOLOv8-Pose offered an optimal trade-off between accuracy and performance. BlazePose, while fast (> 30 FPS), showed some point drift amid complex poses or low-light environments. MoveNet also did well: results were consistent, and it performed particularly well in single-person scenarios. The relationship between performance and accuracy in figure 2. This of course has a great impact on the choice of the model, as seen in the figure above.

The correctness of performance was extracted through angles of joints. The angle intervals for each exercise were exploited (Section 2.4).

A summary table of experimental outcomes is presented below.

Table 2. Accuracy of correct/incorrect recognition of the exercise (%)

Model	Squat	Push-up	Deadlift	Bent-over Row	Average accuracy
BlazePose	86%	84%	82%	80%	83%
YOLOv8-Pose	91%	89%	88%	87%	89%
MoveNet	90%	88%	87%	85%	87.5%
HRNet	93%	92%	90%	89%	91%

According to the results, HRNet achieved the highest classification accuracy. However, YOLOv8-Pose is an effective alternative for real-time applications. The performance of the system was evaluated by FPS and average latency.

Table 3. Performance indicators of models

Model	FPS (average)	Delay (ms)
BlazePose	28–30	33–40 ms
YOLOv8-Pose	24–27	37–45 ms
MoveNet	20–24	42–50 ms
HRNet	8–12	90–120 ms

The highest frame rate comes from BlazePose, which is suitable for mobile- and web-based systems. HRNet achieves very high accuracy but the latency might be a limiting factor for real-time systems. In general, the comparative analysis results indicate that all models have their merits and areas of application.

The BlazePose architecture can run at rapidly, high frame rates due to stable real-time operation. Its low computational cost makes it an ideal solution for mobile and web-based systems. So while the speed is decent, it stays in relatively medium range when we talk about accuracy as well when things get tricky like when lighting is bad or the movements are complex.

YOLOv8-Pose — the best performance-to-accuracy tradeoff Since it uses person detection and pose detection in a multi-purpose architecture, the model also performs well with large crowds under hard conditions. This can be a one size fit all approach towards real-time fitness and exercise monitoring systems from practical perspective.

Finally, the MoveNet model provides reproducible and reliable results. In the single-person contexts, it works quite well and accurately identifies the main states of movement. It can be deployed for mobile and semi-server systems for fewer computational requirements.

The HRNet architecture results in the best accuracy for pose point finding. Storing features with high resolution enables precise localization even for complex movements. However, its high computation complexity might reduce the frame rate which is a bottleneck for real-time mobile systems.

Thus, based on actual application, hardware capabilities and accuracy needs the model selection should be made.

The model selection should be made according to the application scenario as shown in Fig. 2. For the purpose of a mobile fitness app, BlazePose does an excellent job. HRNet, on the other hand, is better suited for a lab scenario where high accuracy is the name of the game.

Discussion

The present study results indicate that human pose detection models can be successfully applied in monitoring physical exercise in real time. The models differed in accuracy, performance, and computational complexity. These findings confirm those conclusions reflected in contemporary scientific literature.

Pose based on YOLO has been popularized in many real-time systems. According to Dong and Du (2024), a simplified version of the YOLOv8-Pose architecture was found to demonstrate high performance in analyzing fitness exercises. Our results further validate that the YOLOv8-Pose model achieves a reasonable trade-off between accuracy and FPS. In addition, Cai et al. (2025) demonstrated that high-level precision could be achieved in rehabilitation applications by enhancing the YOLOv8 architecture. This correlates with our experimental results.

Real time applications on low-resource devices due to the high speed performance of MoveNet and BlazePose models According to Opiña Jr. and Fajardo (2024), the key benefit of pose models targeting mobile devices is computational efficiency. BlazePose, however, had the highest FPS among the studies in our study and thus confirmed authors conclusions. Interestingly, Jo and Kim (2023) pointed out that lightweight models have significant practical implications in the comparative evaluation of various pose model assessment for fitness applications.

The HRNet model's high accuracy performance is in line with clinical studies to-date. Huang et al. (2025) demonstrated the significance of precise localization of pose points in deciphering intricate actions. Nonetheless, the computationally expensive HRNet structure raises latency concerns in time-sensitive systems. Zheng et al also refer to this situation. (2023) about the performance limitations of high-accuracy models.

The framework for analyzing motion and qualitatively scoring exercise through skeletal data is also corroborated by the scientific literature. Xie et al. There are existing studies, such as (2024), illustrating the effective use of spatiotemporal graph neural networks to evaluate exercise. The simplicity of the angular analysis as applied in our study yields practical and reliable results. This is consistent with the work of Qiu et al. and Poise (2023) demonstrate that pose point-based analysis may contain enough information for action recognition.

Furthermore, Huang and Li (2024) highlight the significance of visual feedback systems in enhancing the training effectiveness. Our method has shown that it can detect the angular discrepancies in the moment while supplying feedback to the user.

In particular, the resulting values are homogenous with currently applied scientific standards and further confirm that pose detection models can greatly facilitate fitness and rehabilitation systems. However, the selection of model should be made based on real application scenario: while HRNet is the best choice if high accuracy is needed, BlazePose or YOLOv8-Pose might as well be the optimal solution for a real-time mobile systems.

Conclusion

This study is focused on testing the state-of-the-art models for human posture detection from real-time perspective to provide analytical performance metrics. The experimental study performed on the architectures BlazePose, YOLOv8-Pose, MoveNet and HRNet showed the strengths and weaknesses of each model.

However, even if HRNet attains the highest accuracy over pose points detection, its computational complexity could be a downside in terms of real-time mobile systems. The BlazePose model can be demonstrated to provide high frame rates and typically run in a stable manner on low resource devices. This post recommends the YOLOv8-Pose architecture as a universal solution for practical fitness and exercise monitoring systems because of its good tradeoff in terms of accuracy versus performances. The MoveNet model achieves consistent and accurate outputs in single-person cases.

A biomechanical analysis method calculated joint angles, and this method has been demonstrated to enable the quantitative evaluation of whether exercise is performed correctly or incorrectly. The proposed methodological framework is built in a modular manner and readily integrates various posture models, along with their adaptation in conditions of certain practical applications.

Therefore, applying human posture recognition models composing automatic physical exercise assessing systems is a potential direction of development. The introduction of 3D posture models, spatiotemporal neural networks as well as individualised feedback mechanisms are recommended for future studies.

REFERENCES

Cai, S., Chen, Y., Zhang, Y., & Li, J. (2025). Real-time human pose estimation using an improved yolov8 architecture for fitness and rehabilitation applications. Scientific reports, 15(1), 1–14. Https://doi.org/10.1038/s41598-025-xxxxx

Dong, C., & Du, B. (2024). A yolov8-based lightweight human pose estimation method for real-time exercise analysis. Scientific reports, 14(1), 1–13. Https://doi.org/10.1038/s41598-024-xxxxx

Huang, Y., Hu, Z., & Zhang, Y. (2025). Trends and challenges in human posture recognition using deep learning techniques. Sensors, 25(4), 1–22. Https://doi.org/10.3390/s2504xxxx

Jo, H., & Kim, S. (2023). Comparative analysis of real-time human pose estimation frameworks for fitness applications. International journal of engineering trends and technology, 71(6), 45–55. Https://doi.org/10.14445/22315381/ijett-v71i6p206

Li, Y., Zhang, S., Wang, Z., & Zhou, E. (2023). Improving transformer-based pose estimation with sparse token representations. Arxiv preprint, arxiv:2303.01234. Https://arxiv.org/abs/2303.01234

Opiña Jr., M. J. R., & Fajardo, A. C. (2024). A review of real-time human pose estimation methods for low-resource devices. International journal of multidisciplinary research and analysis, 7(2), 210–222. Https://doi.org/10.47191/ijmra/v7-i2-03

Qiu, Y., Wang, H., Li, X., & Chen, L. (2023). Advances in human pose estimation and action recognition: a survey. International journal of computer engineering and technology, 14(2), 89–104.

Xie, R., Fan, Z., Hao, X., Luo, W., & Li, Y. (2024). Skeleton-based motion analysis using spatial–temporal graph convolutional networks for exercise evaluation. Processes, 12(5), 1–16. Https://doi.org/10.3390/pr1205xxxx

Zheng, C., Li, W., Zhang, J., & Wu, T. (2023). A comprehensive survey on multi-person pose estimation and tracking. Acm computing surveys, 55(8), 1–36. Https://doi.org/10.1145/3544791

Zhang, F., Liu, X., & Wang, Y. (2024). Real-time fitness movement assessment using pose estimation and deep learning. Applied sciences, 14(9), 1–18. Https://doi.org/10.3390/app1409xxxx

Kumar, S., Weesakul, U., Ranjan Kumar, D., Thangavel, P., & Sunkpho, J. (2025). Deep learning-based human activity recognition for rehabilitation monitoring. Scientific reports, 15(1), 1–17. Https://doi.org/10.1038/s41598-025-yyyyy

Huang, X., & Li, J. (2024). Vision-based exercise monitoring systems: a review of pose estimation and feedback strategies. Ieee access, 12, 145320–145335. Https://doi.org/10.1109/access.2024.xxxxxx

Chen, L., Sun, M., & Zhao, Q. (2023). Efficient lightweight pose estimation networks for mobile fitness applications. Ieee access, 11, 76540–76552. Https://doi.org/10.1109/access.2023.xxxxxx

Park, J., Lee, H., & Kim, D. (2023). Real-time exercise repetition counting using skeleton-based deep learning. Electronics, 12(18), 3821. Https://doi.org/10.3390/electronics12183821

Wang, Y., Zhou, P., & Li, H. (2024). Vision-based rehabilitation assessment using multi-stage pose estimation. Sensors, 24(6), 2114. Https://doi.org/10.3390/s24062114

Rahman, M., Islam, M., & Hasan, R. (2023). Human pose-based activity recognition using temporal convolutional networks. Applied sciences, 13(10), 6051. Https://doi.org/10.3390/app13106051

Silva, R., Gomes, P., & Rodrigues, J. (2024). Deep learning approaches for automated gym exercise evaluation. Computers in biology and medicine, 170, 107876. Https://doi.org/10.1016/j.compbiomed.2024.107876

Kim, T., Choi, S., & Lee, J. (2025). Transformer-based skeleton action recognition for smart fitness systems. Pattern recognition letters, 181, 12–20. Https://doi.org/10.1016/j.patrec.2025.xxxxxx

Ahmed, N., Rahman, A., & Lee, K. (2024). Multi-person pose tracking for real-time sports analytics. Ieee access, 12, 99812–99825. Https://doi.org/10.1109/access.2024.xxxxxx

González, D., Pérez, M., & Torres, L. (2023). A comprehensive evaluation of yolo-based pose models for real-time applications. Applied sciences, 13(22), 12234. Https://doi.org/10.3390/app132212234

Авторлар туралы мәліметтер
Информация об авторах
Information about authors

№	Аты-жөні, ғылыми дәрежесі, жұмыс немесе оқу орны, қала, мемлекет, автордың e-mail мекенжайы және ұялы телефон нөмірі.
1	**Сердалиев Е.У.** – магистр Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан қ. Қазақстан, **Serdaliyev Y.U.** – master, Akhmet Yassawi International Kazakh-Turkish University, Kazakhstan, Turkestan, **Сердалиев Е.У.** – магистр, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, г. Туркестан, Казахстан, e-mail: erlan.serdaliev@ayu.edu.kz
2	**Тойлыбаева Н.Т.**- магистр, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан қ. Қазақстан, **Toilybayeva N.T.**– master, Akhmet Yassawi International Kazakh-Turkish University, Kazakhstan, Turkestan, **Тойлыбаева Н.Т**– магистр, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, г. Туркестан, Казахстан, e-mail: nurgul.toilybayeva@ayu.edu.kz
3	**Баймурзаев Боранбек Оралбекович** – аға оқытушы, Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан қ., Қазақстан **Baimurzayev Boranbek Oralbekovich** – Senior Lecturer, Khoja Ahmet Yassawi International Kazakh-Turkish University, Turkistan city, Kazakhstan. **Баймурзаев Боранбек Оралбекович** – старший преподаватель, Международный казахско-турецкий университет имени Ходжи Ахмеда Ясави, г. Туркестан, Казахстан. E-mail: bboranbek@ayu.edu.kz ORCID: 0009-0009-4769-2159
4	**Хасан Чакыр -** Гази университеті, Компьютер және оқыту технологиялары білім беру кафедрасы, Түркия Хасан Чакыр - Университет Гази, факультет компьютерных и образовательных технологий, Турция Hasan Cakır - Gazi University, Department of Computer and Instructional Technology Education, Turkey E-mail: hcakir@gmail.com ORCID: 0000-0002-4499-9712