

# Action Recognition and Sports Evaluation of Running Pose Based on Pose Estimation

Yongchang Yang<sup>1,\*</sup>, Yang Zeng<sup>2</sup>, Li Yang<sup>3</sup>, Yifan Lu<sup>4</sup>, Xinwei Lee<sup>5</sup>, Yasushi Enomoto<sup>6</sup>

<sup>1</sup>Doctoral Program in Physical Education, Health and Sport Sciences, University of Tsukuba, Japan

<sup>2</sup>School of Intelligent Technology, Geely University of China, China

<sup>3</sup>School of Electronic Information Engineering, Geely University of China, China

<sup>4</sup>Viterbi School of Engineering, University of Southern California, the United States

<sup>5</sup>Graduate School of Systems and Information Engineering, University of Tsukuba, Japan

<sup>6</sup>Faculty of Health and Sport Sciences, University of Tsukuba, Japan

Received September 26, 2023; Revised December 13, 2023; Accepted January 8, 2024

## Cite This Paper in the Following Citation Styles

(a): [1] Yongchang Yang, Yang Zeng, Li Yang, Yifan Lu, Xinwei Lee, Yasushi Enomoto, "Action Recognition and Sports Evaluation of Running Pose Based on Pose Estimation," *International Journal of Human Movement and Sports Sciences*, Vol. 12, No. 1, pp. 148-163, 2024. DOI: 10.13189/saj.2024.120118.

(b): Yongchang Yang, Yang Zeng, Li Yang, Yifan Lu, Xinwei Lee, Yasushi Enomoto (2024). *Action Recognition and Sports Evaluation of Running Pose Based on Pose Estimation*. *International Journal of Human Movement and Sports Sciences*, 12(1), 148-163. DOI: 10.13189/saj.2024.120118.

Copyright©2024 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

**Abstract** At present, deep convolutional neural networks (CNNs) have made impressive progress on human pose estimation. Sports pose is one of the important bases for evaluating the athletes' skill level. Until now, the determination and analysis of athletes' pose is still very time-consuming manual work. Therefore, how to adapt and retrain the state-of-the-art human pose detection system and apply it to specific sports competitions is of great significance. Videos recorded on a regular basis can quickly provide a large amount of pose data, and the video recording work is a preliminary preparation for the development of mining algorithms. In this study, we use running as an example of a sport. Using the High-Resolution Net (HRNet) neural network algorithm, we obtain 2D pose data from the video recordings of individual movements to realize the recognition and evaluation of athletes' movement pose during running. This breaks the research boundaries of traditional laboratories and realizes the analysis of movements during actual running competitions. Our experimental results showed the effectiveness of this method, which has the potential to be applied to other cycle-based or phase-based types of sport. The exploration of this study has achieved the identification of running motions and the analysis of relevant motion variables, such as joint angle, stride frequency, stride length, etc. This provides a more scientific technical basis for us to adopt this method in

practical sports video analysis in the future.

**Keywords** Running Pose, Pose Estimation, HRNet, Action Recognition, Object Tracking

## 1. Introduction

Research on cross-discipline integration is becoming more extensive with various developments in science and technology. The commercialization of computers had a significant impact on sports modernization. Nowadays, computers are widely used to improve the movement of athletes and also to assess technical aspects that are difficult to judge with the naked eye [1]. In the traditional human posture recognition system, the tester needs to wear some customized sensors to measure body position, speed, inertia, etc. However, wearable devices have certain security risks due to their intrusiveness. With the advance of computer vision and image processing, image-based non-contact human motion analysis has emerged [2]. For example, recognizing sports activities requires a highly accurate estimation of skeletal joints. Human pose estimation can precisely predict the pose of human body parts in images or videos and make judgments and analysis on their movements [3].

Human pose estimation (HPE) has been an active research topic for decades. In recent years, convolutional neural networks (CNNs) have become increasingly popular in tracking various tasks, such as object recognition, image classification, etc. CNN can also be applied in human pose estimation [4]. Many datasets have been published in recent years to evaluate human pose estimation algorithms. Many of these studies extensively used 2D human pose benchmark datasets [5,6]. Meanwhile, 3D human pose datasets are relatively less diverse due to the constraints of 3D pose capturing sensors [3].

In many sports disciplines, HPE is an important method for performance analysis and hence the improvement of athletic performances [7]. In recent years, with the deepening of cross-integration of different subjects, in-body research on human pose recognition and evaluation has gradually become popular in the field of sports. Huang & Liu [8] used OpenPose as a human key points detection model for 2D pose recognition and converted the detected data into clinical test metrics. Through metric scores, they corrected the sports training methods and reduced athletes' sports injuries. De Souza Vicente et al. [9] used latent-dynamic conditional random fields on RGB-d skeleton estimates of Taekwondo fighters to identify specific kicks and punches in a fight sequence [10]. Lei et al. [4] proposed a network structure for the task of single person pose estimation in a complex environment. They improved the stacked hourglass model and applied it to the pose estimation of a single diver, increasing the detection accuracy of human key points. In addition, Wu & Koike [11] used a residual convolutional neural network to estimate a table tennis player's continuous upper body pose. Bao et al. [12] proposed a video-based method for estimating the pose of ski jumpers. They used High-Resolution Net (HRNet) to transfer the learning of feature knowledge from the public dataset COCO 2017 to the task of ski jumper's pose estimation, and it has improved higher than other mainstream human pose estimate methods.

At present, computer vision has been used to achieve certain research findings in sports, but there are not many applications to running events. Since video data acquisition is relatively simple, this study proposed a video-based method for estimating the pose of runners. This method can greatly reduce the data processing time for traditional technicians to manually label athletes and then perform motion analysis. Analyzing these movements allows us to quantify parts of the race, which can generate sports performance indicators, evaluate the driving factors for success, help make more accurate judgments, and provide more reasonable suggestions.

The main contributions of this study may include the following:

- 1) Faster R-CNN is a classic algorithm with high accuracy and a wide range of use in the field of object detection, and Cascade R-CNN is an improved version of Faster R-CNN. In this study, the method of Cascade

R-CNN is applied to the detection of runners, and the deep convolutional neural network (D-CNN) is used to extract image features and also to realize the object detection and tracking of runners.

- 2) Processing human body detection frames uses the HRNet neural network model, realizing the detection of the tracked object human joint points and outputting the motion coordinate trajectory of each joint point.
- 3) By calculating the coordinate points of the detected human skeleton's joints, we were able to calculate the angle changes of the human body pose in each frame of the image, as well as relevant mechanical variables such as stride frequency, stride length, and body center of mass during running. This helped us to realize the biomechanical evaluation of athletes' sports skills and performances.

## 2. Related Works

The related works will be introduced from three different perspectives: Datasets for human pose estimation, body pose estimation and tracking, and the application of pose estimation in sports.

### 2.1. Datasets for Human Pose Estimation

In recent years, many datasets have been published to evaluate human pose estimation algorithms. The datasets are becoming increasingly complex and more challenging. Usually, 2D human pose datasets are widely used. Due to the limitations of 3D pose-capturing sensors, the 3D human pose datasets are relatively less diverse [3]. The databases for 2D pose estimation mainly include MSCOCO, MPII, FLIC, etc. The current mainstream research databases for pose estimation are the two databases COCO and MPII. The COCO dataset was developed by Microsoft in 2014 for image segmentation detection [13]. The COCO dataset contains 200,000 images and 250,000 examples with 17 key points of human bones. The MPII dataset is an evaluation benchmark for human pose estimation released in 2014, and it is the first large-scale benchmark dataset for human pose estimation [14]. The image frames are selected from YouTube videos, which include about 25,000 images, of which more than 40,000 body key points are annotated, and each human body is annotated with 16 feature key points. MPII can be used for single and multi-person pose estimation evaluation. The FLIC images are collected from popular Hollywood movies [15]. On the other hand, HumanEva-I&II, Human3.6 M and MPI-INF-3DHP are three commonly used datasets for 3D pose evaluation. Compared to the above 2D human pose datasets, 3D pose datasets have less diversity in terms of the background or environment. All samples from Human3.6 M are captured in an indoor environment. Thus models trained with the RGB images from it are usually not well generalizable to real-life scenarios [3].

## 2.2. Body Pose Estimation and Tracking

Before the emergence of CNN, traditional methods relied on various features designed by researchers. Subsequently, with the progress of scientific research, significant improvements have been made in human pose estimation. For example, Newell et al. [16] proposed a stacked hourglass network, which is the first multi-resolution network for human pose estimation. The hourglass based network effectively aggregates features with the same spatial orientation to obtain rich low-level information for accurate key point localization [17]. Chen et al. [18] proposed a Cascade Pyramid Network (CPN) to integrate feature representations at all levels to improve the human pose estimation process, which won the COCO 2017 key point challenge [19]. Subsequently, Sun et al. [20] proposed a novel HRNet that maintains a high-resolution table throughout the whole architecture. HRNet continuously learns high-resolution features by connecting convolutional streams from high-to-low and low-to-high resolution, and achieves state-of-the-art estimation performance [21].

In terms of human body pose estimation algorithms, 2D human body pose estimation has become a basic but challenging problem in computer vision, with the goal of locating key points of the human body (such as elbows and wrists). Its application is also wide, such as Human Action Recognition and Human-Computer Interaction. There are two main types of 2D human pose estimation algorithms: Top-Down and Bottom-Up. Top-Down first detects the human body in the image and then predicts the key points for each human body individually. The computational cost of this method increases with the increase of the number of people, and the accuracy is higher. Bottom-Up first detects all the key points, then clusters the key points and combines them into a human body. The calculation amount of this method does not increase with the increase of the number of people. The 3D human pose estimation algorithm mainly predicts the position of a person in 3D space. According to different inputs, it can be divided into the following three methods: 1) predict 3D Pose from 2D Pose based on monocular images [22]; 2) based on multi-view information [23]; 3) based on video (multi-frame) information [24].

## 2.3. The Application of Pose Estimation in Sports

HPE plays a crucial role in analyzing and enhancing the performance of athletes across various sports disciplines. Ludwig et al. [7] proposed two methods in the research on self-supervised learning for human pose estimation in sports (taking the triple and long jump as an example). The two methods fine-tune a 2D HPE system trained on general poses to a sports discipline-specific HPE model using only a few labeled images. The evaluation results show the sufficiency of a training dataset containing 50 labeled images and some video sequences to train a deep neural

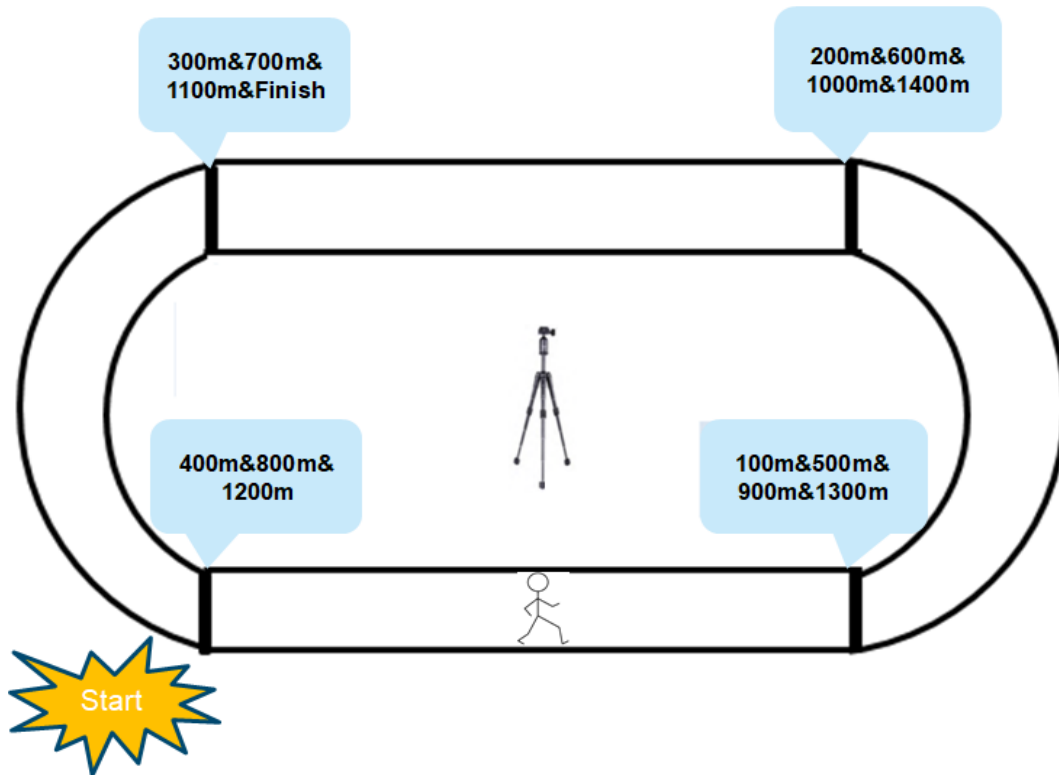
network for a new sports domain such that it generates acceptable results. These methods could open the usage of human pose estimation performance measurements to a wide range of sports disciplines in the future. Also, the expense to collect video material is very low as it only requires a smartphone or a small camera, so it is very cost-friendly. In addition, the tracking and evaluation of soccer players' poses during the competition [25], the detection of long jumpers' landing, jumping and stride events [26], and the flight parameters of ski jumpers such as ski angle, lower body angle, and upper body angle are used to improve their body posture during the flight phase, to achieve long-distance flight [7,12], and to predict the landing point of the serve for table tennis players [11]. Although pose estimation has been applied in many sports events, there is almost no application in track and field competitions (such as sprinting, long-distance running, etc.).

Typically, sports technique analysis is primarily confined to local sports biomechanics laboratories. The use of laboratory measurements is constrained due to the inability of sports equipment to perfectly simulate natural activities and the challenge of replicating outdoor environmental conditions [27]. Moreover, athletes are equipped with various sensors capable of quickly and accurately detecting and tracking movements, providing valuable information. However, this traditional method imposes an additional burden on athletes, significantly affecting their competitiveness [28,29]. With the advancement of artificial intelligence, numerous studies that were previously only achievable through experiments have been changed. In a previous study, we employed OpenCV to track 1500 m runners in mid to long-distance running [30], outputting their actual coordinates, calculating their real running distance and speed. Building upon this foundation, the current study utilizes middle distance running as an exemplar to achieve 2D posture tracking and evaluation during running movements. If the recognition and evaluation of runners' posture in videos can be achieved, it will further transcend the confines of traditional laboratories, fostering the advancement of sports technology action analysis, and enabling the analysis of middle distance runners' movements in competitive settings.

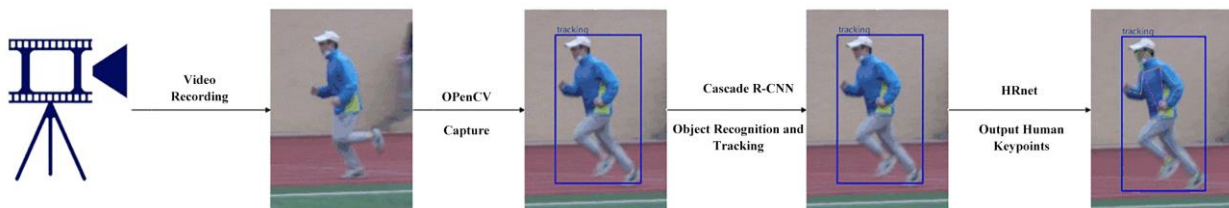
## 3. Method

### 3.1. Video Shooting

We used a high-speed camera to track and film runners. The camera is securely fixed onto a tripod and placed at the center of the university-standard, 400-meter track and field playground. It has a shooting speed of 25 frames per second and an image resolution of 1920 \* 1088 pixels. The detailed setup of the shooting venue is shown in Figure 1.



**Figure 1.** Drawing of video shooting location. The tripod in the center shows the location of the camera. The blue boxes show the positions of 100m intervals in a 1500m race



**Figure 2.** Runner Pose Estimation Method Diagram. First, the recorded video is passed into OpenCV to capture the runner. Then, we use Cascade R-CNN to recognize and track the runner. Finally, we use HRNet to obtain and output the human key points

### 3.2. The Process of the Runner Pose Estimation Method Based on Deep Learning

The runner pose estimation method includes several steps: data acquisition, data annotation, data enhancement, HRNet training, and testing of continuous frame images. The overall process is illustrated in Figure 2.

Figure 3 presents the human pose estimation algorithm (2D), which outlines the steps involved in detecting and

estimating the pose of human subjects in two-dimensional space. The algorithm depicted in this figure includes key components such as feature extraction, candidate generation, pose estimation, and post-processing. The human pose estimation algorithm (2D) can be primarily divided into two methods: Top-Down and Bottom-Up. In the Top-Down method, the algorithm first detects human bodies within an image and then predicts key points for each individual body.

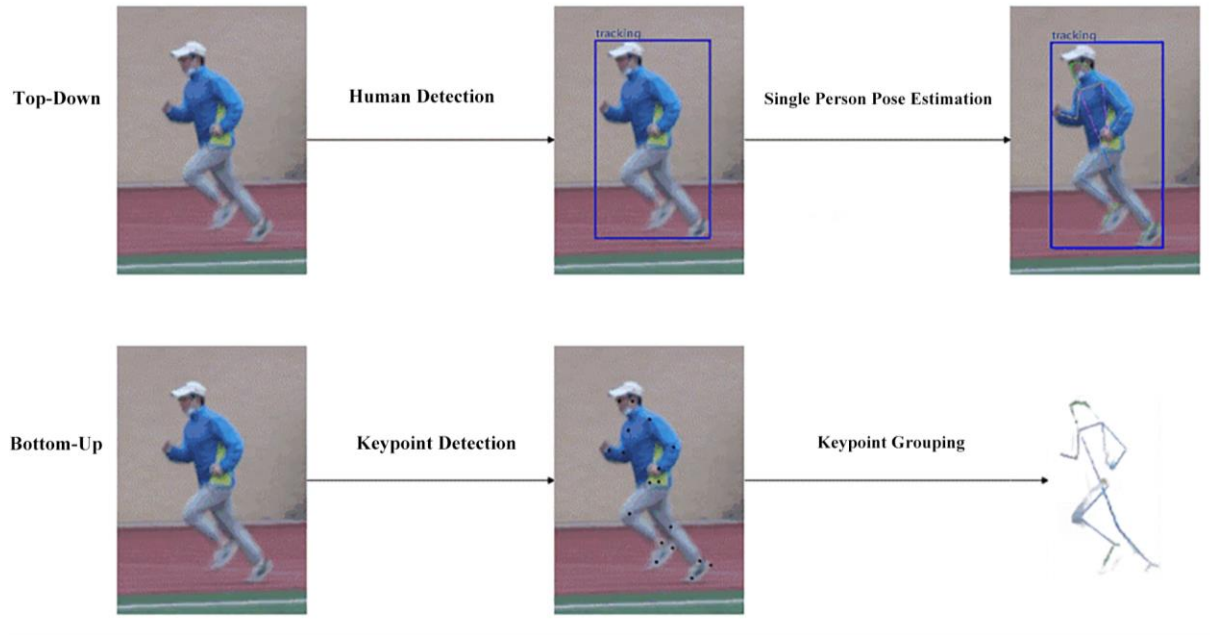


Figure 3. Human Pose Estimation Algorithm (2D)

This method is more robust to different body sizes and provides higher accuracy, but the computation cost increases with the number of bodies present in the image. On the other hand, the Bottom-Up method detects all keypoints and then clusters them to form human bodies, with a constant computation cost regardless of the number of bodies. This method is faster but provides lower accuracy [31]. Considering that the Top-Down algorithm is more accurate, this study adopted the Top-Down algorithm to evaluate the posture of runners, including three main steps: (a) input image, (b) detected by human detector, and (c) single-person pose detection results (Figure 3).

Before we dive into the calculation of mechanical variables, it is important to describe how data is fed into the system. Consider a sample  $x$ , which is a part of the dataset  $X$ . Each sample  $x$  is a pair of a video sequence and the corresponding labelled mechanical variables. The model's training process begins with the extraction of frames from the video sequence, which is used as input for our dual-stage model composed of Cascade R-CNN and HRNet.

### 3.2.1. Preprocessing

Firstly, the video sequence is preprocessed, where each frame in the video is analyzed and reduced to its essential components. The frames are then resized to a predefined size and normalized to adjust for variations in brightness and contrast.

### 3.2.2. Model Training: Cascade R-CNN and HRNet

The Cascade R-CNN is a multi-stage object detection framework that increases detection performance by adding a sequence of detectors trained with increasing IoU thresholds to form a cascaded detector [32].

The formal representation of the model is as follows: given the training set of tuples  $\{(x_1, y_1), \dots, (x_n, y_n)\}$ , where  $x_i$  represents the  $i$ -th image, and  $y_i$  represents the corresponding bounding box and class label, the model is trained by minimizing the following objective function:

$$L = \frac{1}{n} (\sum_i L_{i,cls} + \sum_i L_{i,box}) \quad (1)$$

Where  $L_{i,cls}$  represents the classification loss term for the  $i$ -th sample,  $L_{i,box}$  represents the bounding box loss term for the  $i$ -th sample.  $\sum_i L_{i,cls}$  represents the sum of classification loss terms across all samples,  $\sum_i L_{i,box}$  represents the sum of bounding box loss terms across all samples, and  $n$  represents the number of samples. This equation calculates the average loss value by summing up the loss terms for each sample and dividing by the number of samples.

Once the initial bounding boxes are detected by the Cascade R-CNN, these detected regions are then passed to the HRNet for pose estimation.

The HRNet maintains a high-resolution representation through the entire process and starts from a high-resolution subnetwork, which is parallel to a low-resolution subnetwork. These two subnetworks are connected by multi-scale fusion cells, allowing for multi-scale integration of high-resolution and low-resolution representations.

The formal representation of the HRNet can be as follows: given the detected region of interest (RoI) from the Cascade R-CNN, the objective function of the HRNet is to minimize the mean squared error (MSE) between the predicted and true joint locations:

$$L = \frac{1}{N} \sum_i \|J_{i,pred} - J_{i,true}\|^2 \quad (2)$$

Where  $N$  is the number of joints,  $J_{i,pred}$  is the predicted joint location, and  $J_{i,true}$  is the true joint location.

After training these models, we have a system that can accurately detect a person in a video frame using Cascade R-CNN and subsequently predict their pose using HRNet. These predicted poses can then be utilized to calculate the mechanical variables as described in the following sections.

### 3.3. Calculation of Mechanical Variables

This study aimed to obtain the mechanical variables during the running process, mainly by calculating the joint angles, body center of mass, stride frequency, and stride length of runners. The following formulas were used for the calculations:

- 1) By  $P_i (i = 1,2,3)$  three points, calculate the distances between every two points as follows: distance  $a$  (Equation(3)) between point 1 and point 2, distance  $b$  (Equation(4)) between point 2 and point 3, and distance  $c$  (Equation(5)) between point 1 and point 3.

$$a = \sqrt{(P_{x2} - P_{x1})^2 + (P_{y2} - P_{y1})^2} \quad (3)$$

$$b = \sqrt{(P_{x3} - P_{x2})^2 + (P_{y3} - P_{y2})^2} \quad (4)$$

$$c = \sqrt{(P_{x3} - P_{x1})^2 + (P_{y3} - P_{y1})^2} \quad (5)$$

According to the cosine rule of trigonometry (Equation (6)), this formula is used to calculate the joint angles of the runner (Equation(7)).

$$\cos x = \frac{a^2 + b^2 - c^2}{2ab} \quad (6)$$

$$x = \frac{180}{\pi} \cos^{-1} \left( \frac{a^2 + b^2 - c^2}{2ab} \right) \quad (7)$$

- 2) We calculate stride frequency as the number of times of both legs cross in one minute. In terms of stride length calculation, based on the 'Inference\_top\_down\_pose\_model' in the OpenPose library, detection is performed step by step from the top [33]. First, all body parts in the image are detected and filtered and then combined into key points. To obtain real-time stride length, the attributes  $x_l$  (left ankle point) and  $x_r$  (right ankle point) corresponding to the left and right ankle key points in the 'Inference\_top\_down\_pose\_model' are taken separately, and the real-time stride is set to  $X$  (Equation (8)).

$$X = |x_l - x_r| \quad (8)$$

- 3) Regarding the calculation of body center of mass, we first calculate the real-time center point. 18 key points  $P_i (i = 1,2,3 \dots \dots, i \leq 18)$  are extracted using OpenPose. Fit 18 key points and take the average point  $P_{(x,y)}$  (Equation (9)(10)).

$$P_x = \frac{1}{18} \sum_{i=1}^{18} P_{x_i} \quad (9)$$

$$P_y = \frac{1}{18} \sum_{i=1}^{18} P_{y_i} \quad (10)$$

The real-time center point can be represented as (Equation (11)):

$$P_{(x,y)} = \left( \frac{1}{18} \sum_{i=1}^{18} P_{x_i}, \frac{1}{18} \sum_{i=1}^{18} P_{y_i} \right) \quad (11)$$

Next, using real-time center points to fit center of mass points. Fitting the center point of each frame in a video with the center point of the previous frame can continuously optimize the center point position, thereby achieving the goal of approaching the center of gravity position. Get the center of mass point  $P_{(x_{i/2}, y_{i/2})}$  from two center points  $P_i (i = 1,2)$ .

The center of mass point is recalculated in real-time with each frame of video read to optimize and reduce the error. The error decreases as the video continues to be read (Equation (12)(13)).

$$x_{i/2} = \frac{P_{x1} + P_{x2}}{2} \quad (12)$$

$$y_{i/2} = \frac{P_{y1} + P_{y2}}{2} \quad (13)$$

### 3.4. Calibration Method

In this study, we used pixels as the unit of length. To obtain actual length, we converted pixels to centimeters using the following method. First, we need to obtain the actual height of runner  $L$  and the number of pixels that represent the runner's height in the video  $P$ . Then, we can calculate the ratio of actual height to the height in pixels  $L/P$ . Finally, we can multiply our experimental data by the ratio  $L/P$  to convert the unit from pixels to centimeters.

Example: Experimental data stride length  $X = 79px$ , Actual height of the runner  $L = 175cm$ , Height in pixels  $P = 163px$ , Actual stride length  $S = X * L/P = 79px * 175cm/163px = 84.82cm$ . The vertical displacement of the body center of mass also needs to be calibrated using a similar method. In this study, we attempted to output the actual height and the center of gravity of the person in the image for more accurate data calibration and analysis. As we discussed earlier, we first capture the runner's image and then identify their poses. The runner's bounding box changes as they move, so the number of pixels that represent their height is constantly changing. However, the runner's center of mass remains relatively constant during running. To reduce measurement errors, we subtracted 5 cm from the actual height of the subjects to determine the actual length of the runners more accurately in the OpenPose dynamic capture frame. Due to the possibility of slight squatting or leaning forward during the start and movement process, the overall height may be slightly reduced, so we reduced the actual height of the human body by 5 cm same during the calculation process to ensure that the data obtained from our research is closer to the true value. This method can ensure that the data we obtain from the study matches the actual values.

Due to the lack of actual height information for the person in the image, the current algorithm could not

perform actual calibration of stride and center of mass displacement. To solve this problem, we will continue to improve the algorithm to output the actual height information of the person in the image and include it in the analysis of stride and center of gravity displacement. This way, we can assess the athletic ability and skill of athletes more accurately and provide a more scientific and reliable calculation method for future data calibration.

### 3.5. Data Processing

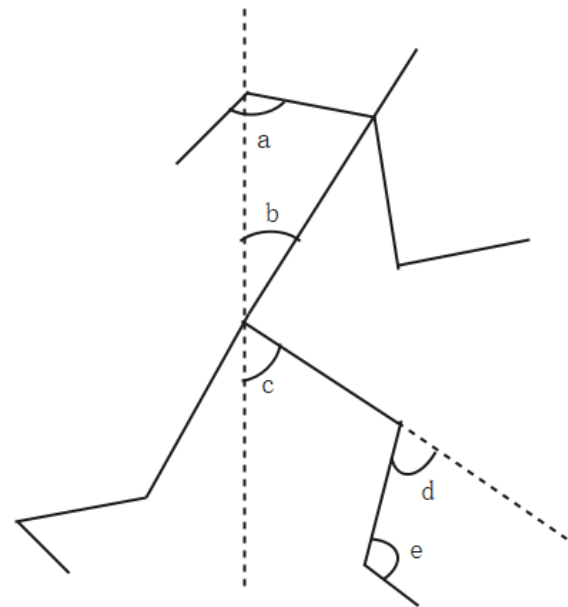
In this study, we extracted a 1-minute video clip from a 1500m running video for motion recognition and variable output. In this study, Python was used to apply Gaussian smoothing to the obtained data to remove noise. Gaussian smoothing is a smoothing filter method based on Gaussian distribution, mainly used to eliminate noise in images or data. In practical usage, we set the Sigma value to 5. After Gaussian filtering, the data noise was effectively removed, and the extraction of important feature points resulted in more accurate results.

## 4. Results and Discussions

### 4.1. Body Joint Angle Features of Runners

This study has primarily calculated and analyzed the elbow joint angle, thigh joint angle, knee joint angle, ankle joint angle, and trunk angle of runners (shown in Figure 4). In a complete exercise cycle, the most critical part of completing technical movements is the arm swing technique. It is the motivation for athletes to continue to run forward and fast, which plays a vital role in improving the performance of competitors. In addition, the thrust obtained by kicking leg backwards can be enhanced by adjusting the posture of the swing arm. If the athlete's upper limbs do not do any exercise, even if the athlete speeds up the running speed or increases the frequency and length of running, it is difficult for them to obtain effective driving force for pushing backwards. Therefore, when an athlete is running, the arm swing of the upper limbs is extremely important, for more propulsion when your swing leg strides [34]. When running, the angle formed by the connection line between the hip joint and the knee joint and the axis perpendicular to the hip joint is called the thigh angle (Figure 4(c)). The movement of the thigh is mostly related to the contraction and extension of the frontal axis of the hip joint [35]. Thigh angle and knee joint angle have a significant impact on the running economy [36]. The knee joint angle is defined as the line extended from the thigh and the calf (Figure 4(d)). The change of the knee joint angle also determines the athlete's technical movements to a certain extent and has a significant impact on the running speed [37]. The angle of the ankle joint is the angle between the line connecting the knee joint and ankle joint, and the line connecting the ipsilateral heel

tubercle and the toe joint (90 degrees). Research has shown that as running speed increases, the plantarflexion angle shows a decreasing trend [38]. The angle of the ankle joint at the moment of lift-off and the range of plantar flexion of the ankle joint during the support stage show an increasing trend with the increase of speed. The maximum plantarflexion angle in the swing stage shows a decreasing trend with the increase of speed. There is a correlation between changes in ankle joint angle and running speed. With the improvement of the running technique and the improvement of the running speed, the forward angle of the trunk should become larger and larger. But increased torso lean angles shift the center of mass forward, which may increase the metabolic cost of running. In addition, forward trunk-leaning yielded significantly lower peak knee extension moments and higher hip extension moments, whereas the opposite effects occurred with backward trunk-leaning [39]. The joint angle division of this study is shown in Figure 4.



**Figure 4.** Elbow joint angle(a), trunk angle(b), thigh joint angle(c), knee joint angle(d), and ankle joint angle(e) conventions used during data analysis [40]

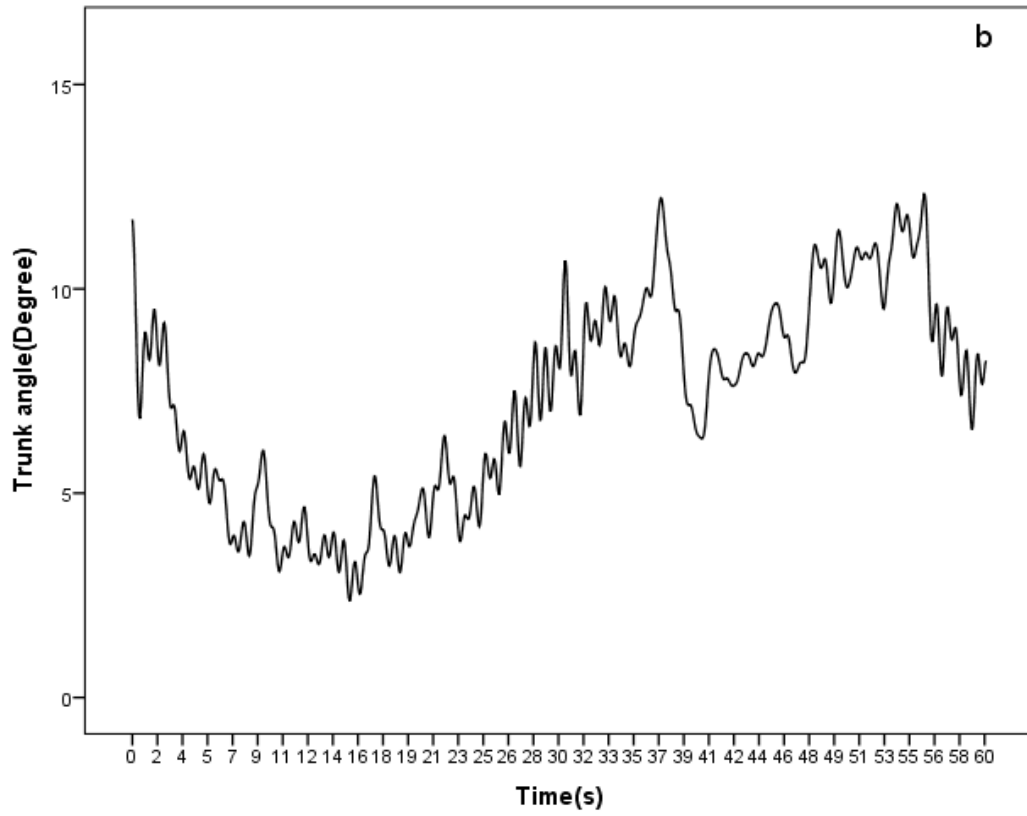
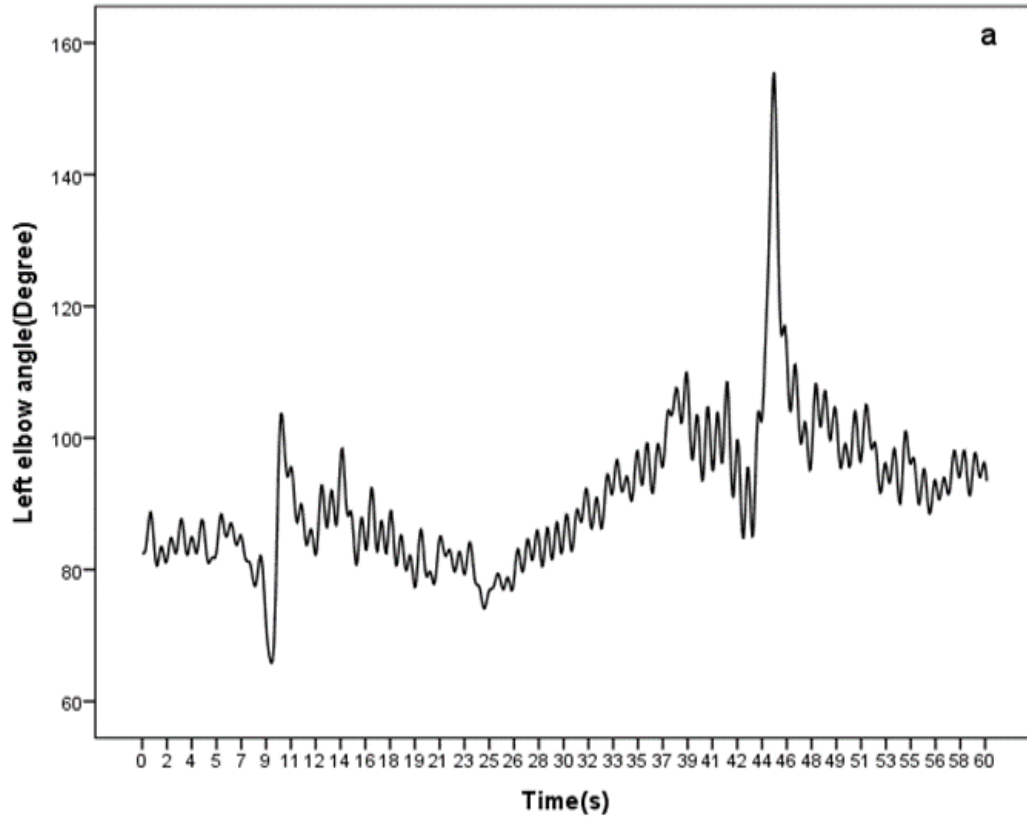
Figure 5 is a joint angle-time relationship graph. We analyzed a one-minute video of a runner, and Figure 5(a) shows the schematic of the changing left elbow joint angle. The average angle of the left elbow joint of this runner is 91 degrees. At the 9th second, it shows a minimum value of 66 degrees. At around 45 seconds, it shows a maximum value of 155 degrees. From the graph, it can be seen that the runner maintains a basic swinging posture of the left elbow joint at 90 degrees during running. However, due to fatigue and other phenomena, the elbow joint angle may change to alleviate its tense state. Figure 5(b) presents the changes in the forward inclination angle of the runner's torso. Analysis

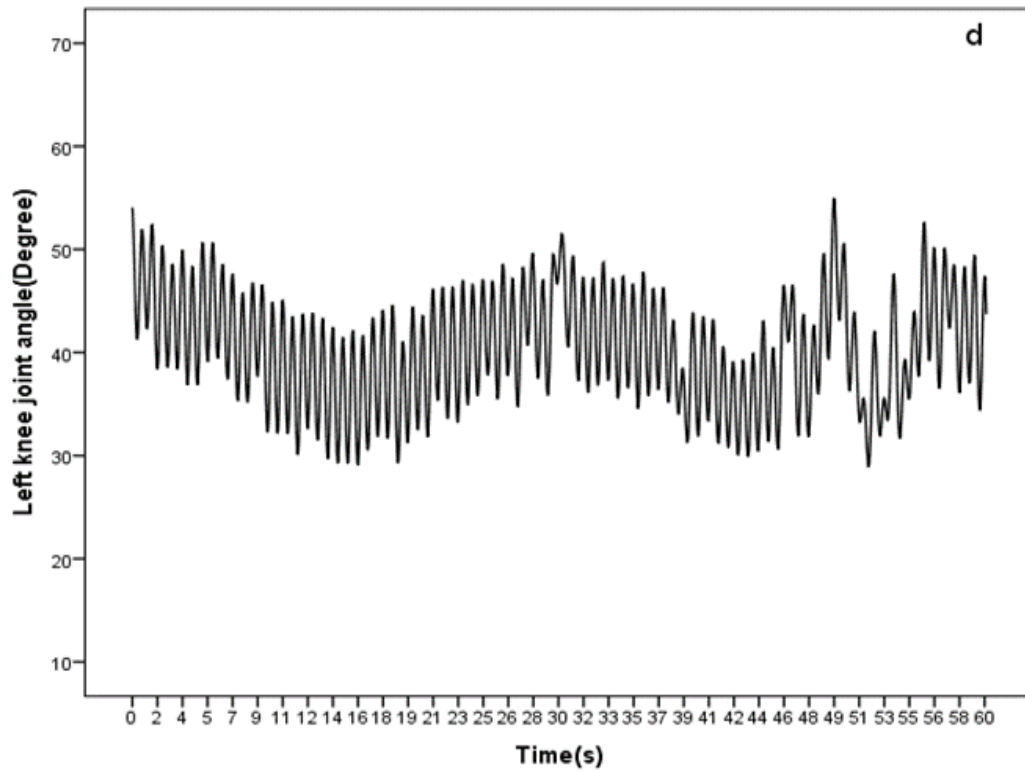
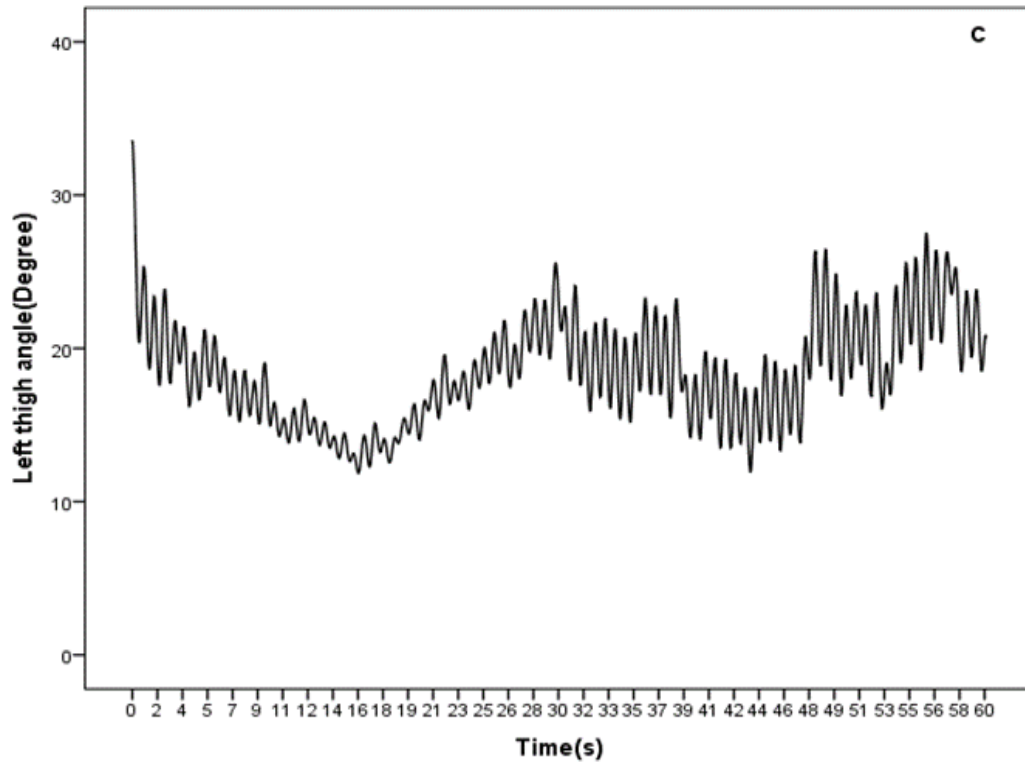
of the posture of this runner in the one-minute video shows that the average forward inclination angle of the torso is 7.3 degrees, with a minimum of 2.4 degrees and a maximum of 12.3 degrees. In Yamaguchi's analysis of the movements of women's mid- distance running events, he found that athletes of different levels and sports have slightly different degrees of forward inclination angle of the runner's torso during mid to long- distance running, with an average upper body forward inclination angle of 10.7 degrees [41]. As the running speed increases, the forward inclination angle of the torso should correspondingly increase. However, the excessive forward inclination of the torso can shift the body's center of gravity forward, thus affecting the extension of the runner's knees. Therefore, during running, an appropriate forward inclination of the torso should be maintained to prevent the negative effects of excessive or backward inclination of the torso. Figure 5(c) shows the changes in the forward swing angle of the left thigh. The forward swing angle of the thigh refers to the angle of the thigh swinging forward during running. The average forward swing angle of the thigh is 18.5 degrees, with a maximum swing of 33.6 degrees and a minimum swing of 11.9 degrees. It can be seen that the size of the forward swing amplitude of the thigh varies due to factors such as individual physical flexibility, running posture habits, and stride length. If the forward swing amplitude of the thigh is too large or insufficient, it may have an adverse effect on stride length, stride frequency, and running efficiency. Therefore, runners should control the forward swing amplitude of their thighs reasonably to achieve the optimal running state.

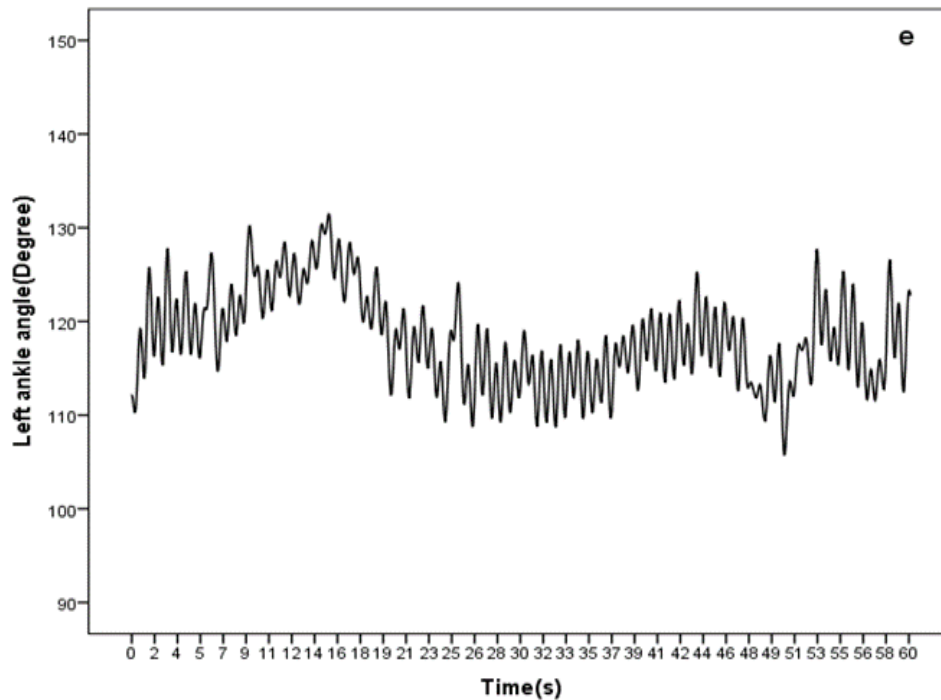
Figure 5(d) shows a schematic of the bending changes of the runner's left knee joint. When running, the knee joint will undergo a certain degree of bending and extension, and the angle change range is approximately between 30 and 80 degrees, and the activity range of the knee joint is around 60 to 115 degrees, which is suitable for running [42]. However, it should be noted that the degree of knee joint bending during running will vary depending on different body structures and running postures, so not everyone will fall within this range. Generally speaking, the bending angle of the knee joint is considered normal at approximately between 30 and 55 degrees. However, some people may be below 30 degrees or above 55 degrees, which does not necessarily mean that there is a problem. According to the report by Hamner &

Delp [43], the knee joint biomechanics are mainly affected by running speed. In addition, this study only analyzed one side photographed by the camera, so the conclusion is relatively limited. According to the research by Ma & Ding [44], by observing and analyzing the data, it was found that the knee joint angle characteristics of the same person are relatively stable when walking with the same stride length. In addition, the left and right knee joint angles of some individuals are symmetrical, but this is not universal, and there are significant differences between different people. Therefore, this study assumed that there was symmetry between the left and right knee joint angles of the runner. It should be noted that the body posture and stride size during running will also affect the bending angle of the knee joint, and there may be differences between different people. Therefore, it is recommended to maintain, adjust, and improve an appropriate degree of knee joint bending during running, depending on personal body conditions, to achieve the best results.

Figure 5(e) shows a schematic of the changes in the ankle angle of the runner. The range of ankle joint angle changes usually refers to dorsiflexion and plantarflexion of the feet, which is the degree to which the feet move up or down. According to the analysis of ankle joint angle changes in this study, the maximum value observed was 131 degrees (plantarflexion), the minimum value recorded was 106 degrees (dorsiflexion), with an average angle of 118 degrees for dorsiflexion and plantarflexion. During running, the ankle begins from dorsiflexion and then undergoes plantarflexion as the foot moves forward. Afterwards, during the landing, stability, and swing phases of running, the foot is involved in varying degrees of dorsiflexion and plantarflexion movements, so the range of ankle joint angle changes has a strong correlation with the motion trajectory and action changes of the foot. In addition, research by Cui & Ren [38] shows that with an increase in running speed, the plantarflexion angle tends to decrease. The angle of the ankle joint at the moment of lift-off and the range of plantar flexion of the ankle joint during the support stage show an increasing trend with the increase of speed. The maximum plantarflexion angle in the swing stage tends to decrease with increasing speed. This indicates that there is a correlation between ankle joint angle changes and running speed. Therefore, when running, attention should be paid to controlling the range of ankle joint angle changes to improve running efficiency.







**Figure 5.** Diagram of the relationship between joint angles and time (a, Left elbow angle; b, Trunk angle; c, Left thigh angle; d, Left knee joint angle; e, Left ankle angle)

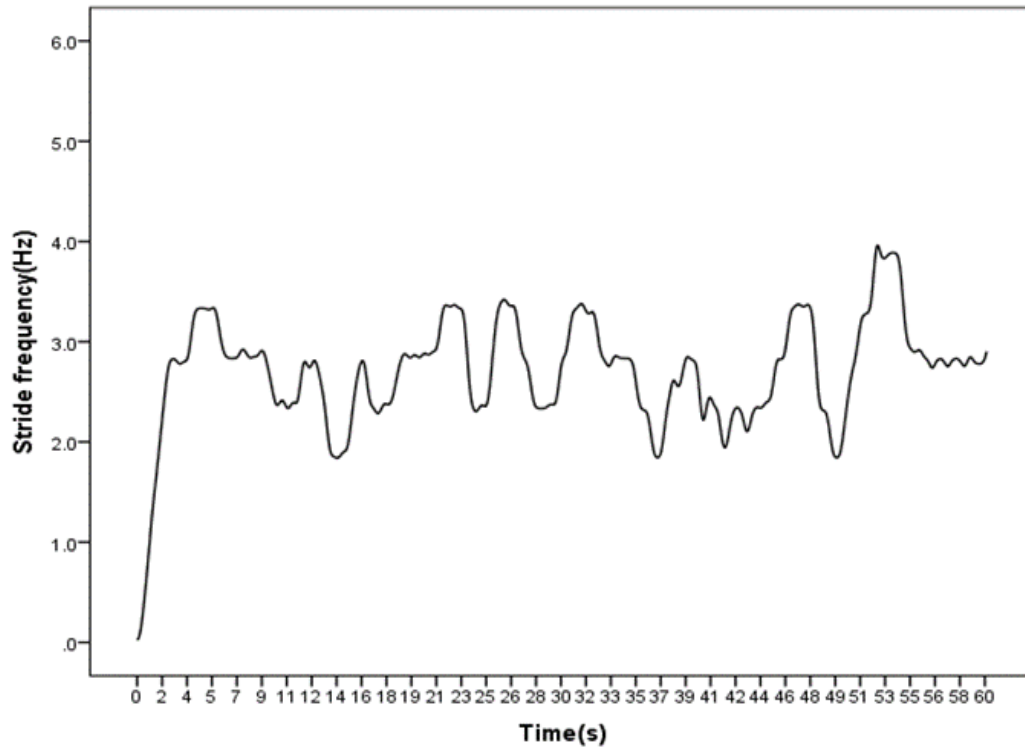
#### 4.2. Characteristics of Runners' Stride Frequency and Stride Length

Figure 6 presents the relationship between a runner's one-minute stride frequency changes and time, displaying the number of steps taken per minute over a certain period of time. Stride frequency refers to the number of steps taken within a certain period of time, usually measured in steps per minute. The stride frequency directly affects the effectiveness and speed of running. This runner's average stride frequency is 2.70 Hz. For the general population, a stride frequency between 2.0 Hz and 3.5 Hz is common during running [45], although the actual stride frequency may vary depending on individual factors and training levels. When running, maintaining a reasonable and stable stride frequency is important for staying at an appropriate level of exercise intensity and reducing exercise-related injuries. Therefore, understanding one's own stride frequency and adjusting accordingly can help runners achieve the best running results.

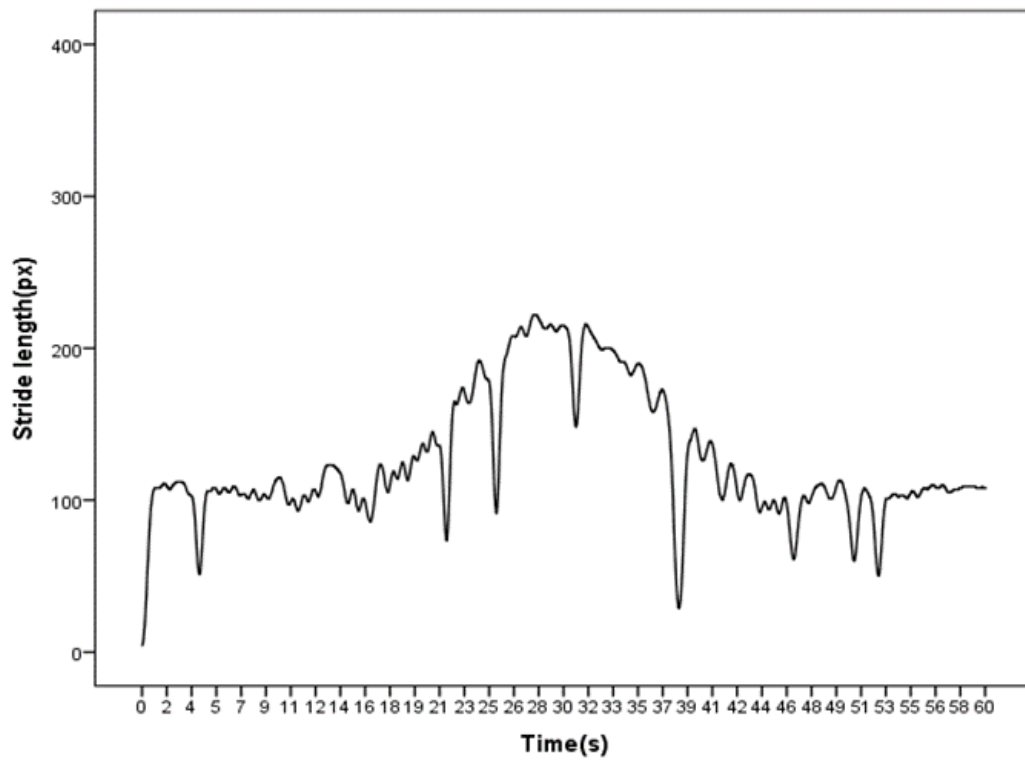
The size of a person's stride when running can vary due to individual differences, height, step rate, running speed, and other factors. Generally speaking, both too large or too small stride length can affect running efficiency and speed, and need to be scientifically and reasonably adjusted according to individual conditions. Within the normal range, an appropriate stride length can help runners better utilize the body's natural motor mechanism and principles of dynamics, making running more economical and stable. The average stride length for people when running is approximately between 0.95-1.5 meters [46]. The results of this study are the stride length of the person in the image,

not the actual stride length. The average stride length of the runner is 1.26 pixels (see Figure 7), but this data has not been calibrated yet. Based on our research method, we also need to output the actual height of the person in the image to calibrate and calculate the actual stride length of the runner. In addition, when identifying the stride length of the runner, there may be situations where the distance between the legs is almost overlapping, which may result in abnormal cliff-like drops in data. To improve the accuracy and stability of data processing, we will continue to improve this method in future research.

Stride frequency and stride length are two important factors that affect running speed and efficiency, with a certain correlation and interaction between them. As we know, running speed equals the product of stride frequency and stride length. Generally speaking, the faster the stride frequency and the longer the stride length, the faster one runs. According to research by Hogberg [47], Luhtanen & Komi [48], among others, the increase in stride length has a greater impact than the increase in stride frequency before running speed reaches 70-80% of maximum speed. If running speed continues to increase, stride frequency increases rapidly while stride length decreases. Therefore, for beginners in running, gradually increasing stride frequency and appropriately adjusting stride length can help develop good running habits and body awareness. This also reduces unnecessary exercise burdens and training risks. For high-level runners, a reasonable ratio of stride frequency and stride length can help improve running efficiency and stability, and to some extent, avoid running injuries and fatigue.



**Figure 6.** Schematic diagram of changes in stride frequency and time of runner



**Figure 7.** Schematic diagram of changes in stride length and time of runner

### 4.3. Characteristics of Center of Mass Displacement of Runners

The distance of the vertical displacement of the body center of mass is related to a person's height, running speed, stride length, stride frequency, and other factors. Typically, the vertical displacement of the body center of mass during running is approximately 6-10 cm [49], and this displacement will increase with an increase in speed and stride length. High-level athletes or runners, due to their well-trained and good running skills, usually have a smaller vertical displacement of the body center of mass, typically between 3-5 cm. According to the calculation of Yang & Enomoto [50] on the vertical displacement of the body center of mass for 1500-meter runners of different levels, the vertical displacement of the center of gravity for competitive runners is approximately 4 centimeters. Strohrmann et al. found longer ground-contact times and greater vertical displacement in novice runners [51], it's important to note that maintaining a stable center of gravity and reducing excessive vertical displacement during running can reduce joint stress and the risk of injury, as

well as make running more efficient and comfortable [52].

The results of this study are the actual vertical displacement of the body center of mass of the person in the image. The average vertical displacement of the body center of mass of this runner is 10 pixels (Figure 8). Like stride length, this data has not been calibrated. Therefore, all distance-related data needs to be calibrated based on the actual height of the person in the image before calculating the actual vertical displacement of the runner's body center of mass. Thus, our next step is to output and calibrate the height of the person in the image.

In addition, we also faced some key problems in the research process, such as low camera performance, inexperienced video shooting skills, and output and calibration of the actual height of the person in the video. We will continuously improve and optimize these technical issues in future research to improve the accuracy and practical value of the data. Next, we will apply this method to analyze the athletic ability of athletes in competition and further explore the practical application value of this method.

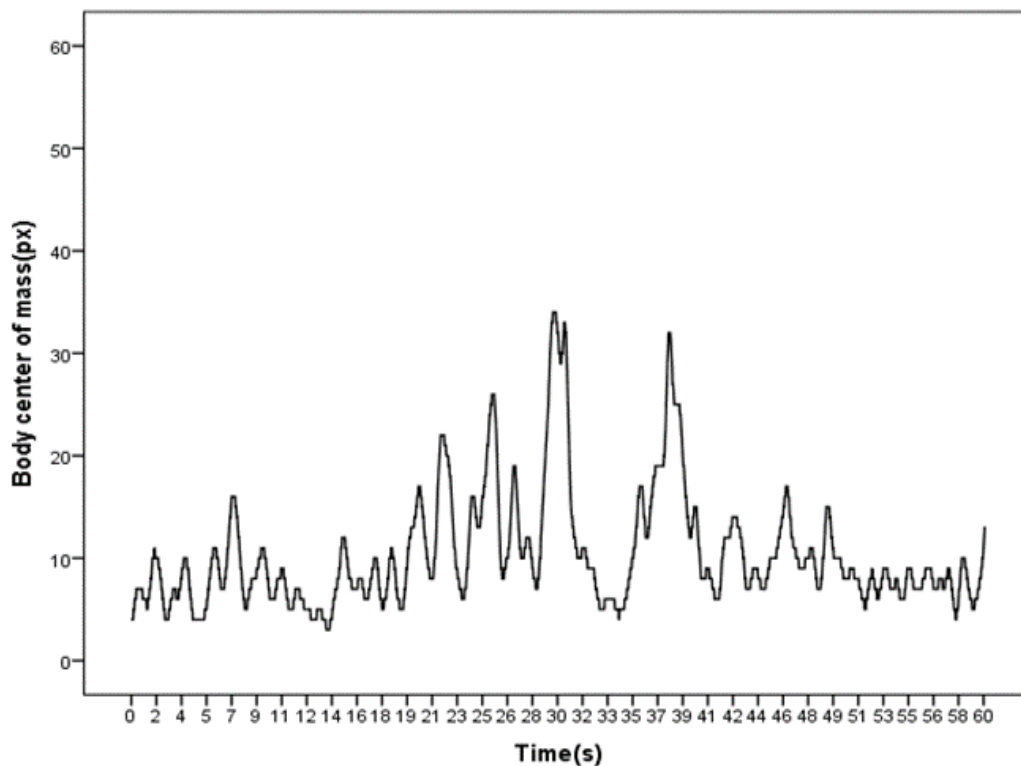


Figure 8. Schematic diagram of changes in body center of mass and time of runner

## 5. Conclusions

We applied posture estimation technology to evaluate running motions. By identifying running motions and analyzing relevant motion variables, effective evaluation of running motions was achieved. Traditional posture estimation usually only identifies motions and rarely conducts detailed analysis and evaluation of motions in practical sports. The exploration of this study has achieved the identification of running motions and the analysis of relevant motion variables, which provides a more scientific technical basis for us to adopt this method in practical sports video analysis in the future.

## REFERENCES

- [1] Kuruppu, G., Manoj, C., Kodituwakku, S.R., Pinidiyaarachchi, U.A.J. Comparison of different template matching algorithms in high speed sports motion tracking. In 2013 IEEE 8th International Conference on Industrial and Information Systems, pp. 445-448, 2013. DOI: 10.1109/ICIInfS.2013.6732025
- [2] Su, P., Chai, Z., Wang, L., Cui, S. Eight-Section Brocade Sequence Action Recognition and Evaluation Based on Pose Estimation. *Electronic Science and Technology*, 35(12), pp.84-90, 2022. DOI: 10.16180/j.cnki.issn1007-7820.2022.12.012 (in Chinese)
- [3] Song, L., Yu, G., Yuan, J., Liu, Z. Human pose estimation and its application to action recognition: A survey. *Journal of Visual Communication and Image Representation*, 76, pp.103055, 2021. DOI: 10.1016/j.jvcir.2021.103055
- [4] Lei, F., Yan, J., Wang, X. Human Pose Estimation of Diver Based on Improved Stacked Hourglass Model. In *Proceedings of the 3rd International Conference on Video and Image Processing*, pp.178-182, 2019. DOI:10.1145/3376067.3376111
- [5] Johnson, S., Everingham, M. Clustered pose and nonlinear appearance models for human pose estimation. In *bmvc*, Vol. 2, No. 4, pp. 5, 2010. DOI: 10.5244/C.24.12.
- [6] Iqbal, U., Milan, A., Gall, J. Posetrack: Joint multi-person pose estimation and tracking. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2011-2020, 2017. DOI: 10.1109/CVPR.2017.495
- [7] Ludwig, K., Scherer, S., Einfalt, M., Lienhart, R. Self-supervised learning for human pose estimation in sports. In 2021 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), pp.1-6, 2021. DOI: 10.1109/ICMEW53276.2021.9456000
- [8] Huang, L., Liu, G. Functional motion detection based on artificial intelligence. *The Journal of Supercomputing*, 78(3), pp.4290-4329, 2022. DOI: 10.1007/s11227-021-04037-3
- [9] De Souza Vicente, C. M., Nascimento, E. R., Emery, L. E. C., Flor, C. A. G., Vieira, T., Oliveira, L. B. High performance moves recognition and sequence segmentation based on key poses filtering. In *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp.1-8, 2016. DOI: 10.1109/WACV.2016.7477711
- [10] Lienhart, R., Einfalt, M., Zecha, D. Mining automatically estimated poses from video recordings of top athletes. *International Journal of Computer Science in Sport*, 17(2), 94-112, 2018. DOI: 10.2478/ijcss-2018-0005
- [11] Wu, E., Koike, H. Futurepong: Real-time table tennis trajectory forecasting using pose prediction network. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1-8, 2020. DOI: 10.1145/3334480.3382853
- [12] Bao, W., Niu, T., Wang, N., Yang, X. Human pose estimation of ski jumpers based on video. In 7th International Symposium on Advances in Electrical, Electronics, and Computer Engineering, Vol. 12294, pp. 732-737, 2022. SPIE. DOI: 10.1117/12.2639724
- [13] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L. Microsoft coco: Common objects in context. 686 In *Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Springer* 687 International Publishing, Part V 13, pp.740-755,2014. DOI:10.48550/arXiv.1405.0312
- [14] Andriluka, M., Pishchulin, L., Gehler, P., Schiele, B. 2d human pose estimation: New benchmark and state of the art analysis. In *Proceedings of the IEEE Conference on computer Vision and Pattern Recognition*, pp. 3686-3693, 2014. DOI: 10.1109/CVPR.2014.471
- [15] Sapp, B., Taskar, B. Modec: Multimodal decomposable models for human pose estimation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp.3674-3681, 2013. DOI: 10.1109/CVPR.2013.471
- [16] Newell, A., Yang, K., Deng, J. Stacked hourglass networks for human pose estimation. In *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, Proceedings, Part VIII*14, pp. 483-499, 2016. DOI: 10.1007/978-3-319-46484-8\_29
- [17] Cai, Y., Wang, Z., Luo, Z., Yin, B., Du, A., Wang, H., Sun, J. Learning delicate local representations for multi-person pose estimation. In *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, Proceedings, Part III*16, pp.455-472, 2020. DOI: 10.1007/978-3-030-58580-8\_27
- [18] Chen, Y., Wang, Z., Peng, Y., Zhang, Z., Yu, G., Sun, J. Cascaded pyramid network for multi-person pose estimation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp.7103-7112, 2018. DOI: 10.1109/CVPR.2018.00742
- [19] Li, Y., Wang, C., Cao, Y., Liu, B., Luo, Y., Zhang, H. A-hrnet: Attention based high resolution network for human pose estimation. In 2020 Second International Conference on Transdisciplinary AI (TransAI), pp. 75-79, 2020. IEEE. DOI: 10.1109/TransAI49837.2020.00016
- [20] Sun, K., Xiao, B., Liu, D., Wang, J. Deep high-resolution representation learning for human pose estimation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5693-5703, 2019. DOI: 10.1109/CVPR.2019.00584

- [21] Wang, J., Sun, K., Cheng, T., Jiang, B., Deng, C., Zhao, Y., ... Xiao, B. Deep high-resolution representation learning for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, 43(10), pp.3349-3364, 2020. DOI: 10.1109/TPAMI.2020.2983686
- [22] Martinez, J., Hossain, R., Romero, J., Little, J. J. A simple yet effective baseline for 3d human pose estimation. In *Proceedings of the IEEE international conference on computer vision*, pp.2640-2649, 2017. DOI: 10.1109/iccv.2017.288
- [23] Tu, H., Wang, C., Zeng, W. Voxelpose: Towards multi-camera 3d human pose estimation in wild environment. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, Proceedings, Part I 16*, pp. 197-212, 2020. DOI: 10.1007/978-3-030-58452-8\_12
- [24] Pavllo, D., Feichtenhofer, C., Grangier, D., Auli, M. 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*, pp.7753-7762, 2019. DOI:10.48550/arXiv.1811.11742
- [25] Bridgeman, L., Volino, M., Guillemaut, J. Y., Hilton, A. Multi-person 3d pose estimation and tracking in sports. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 2019. DOI: 10.1109/CVPRW.2019.00304
- [26] Einfalt, M., Dampeyrou, C., Zecha, D., Lienhart, R. Frame-level event detection in athletics videos with pose-based convolutional sequence networks. In *Proceedings of the 2nd International Workshop on Multimedia Content Analysis in Sports*, pp. 42-50, 2019. DOI: 10.1145/3347318.3355525
- [27] Wixted, A. J., Billing, D. C., & James, D. A. Validation of trunk mounted inertial sensors for analysing running biomechanics under field conditions, using synchronously collected foot contact data. *Sports Engineering*, 12, pp.207-212, 2010. DOI: 10.1007/s12283-010-0043-2
- [28] Ghasemzadeh, H., & Jafari, R. Coordination analysis of human movements with body sensor networks: A signal processing model to evaluate baseball swings. *IEEE Sensors Journal*, 11(3), pp.603-610, 2010. DOI: 10.1109/JSEN.2010.2048205
- [29] Sheng, M., Wang, W., Qin, H., Wan, L., Li, J., & Wan, W. A Novel Changing Athlete Body Real-Time Visual Tracking Algorithm Based on Distractor-Aware SiamRPN and HOG-SVM. *Electronics*, 9(2), 378, 2020. DOI: 10.3390/electronics9020378
- [30] Yang, Y., Lee, X., Enomoto, Y. Implementing the tracking of 1500 m runners using Open CV. *Journal of Physical Education and Sport*, 23(7), 1698-1705, 2023. DOI: 10.7752/jpes.2023.07208
- [31] Dang, Q., Yin, J., Wang, B., Zheng, W. Deep learning based 2d human pose estimation: A survey. *Tsinghua Science and Technology*, 24(6), pp.663-676, 2019. DOI: 10.26599/TST.2018.9010100
- [32] Cai, Z., Vasconcelos, N. Cascade r-cnn: Delving into high quality object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 6154-6162, 2018. DOI: 10.48550/arXiv.1712.00726
- [33] Cao, Z., Simon, T., Wei, S. E., Sheikh, Y. Realtime multi-person 2d pose estimation using part affinity fields. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7291-7299, 2017. DOI: 10.48550/arXiv.1812.08008
- [34] Macadam, P., Cronin, J. B., Uthoff, A. M., Johnston, M., & Knicker, A. J. Role of arm mechanics during sprint running: A review of the literature and practical applications. *Strength & Conditioning Journal*, 40(5), pp.14-23, 2008. DOI: 10.1519/SSC.0000000000000391
- [35] Novacheck, T. F. The biomechanics of running. *Gait & posture*, 7(1), pp.77-95, 1998. DOI:10.1016/S0966-6362(97)00038-6
- [36] Zhao, L., Ma, X., Zhang, B., Wang, K. Biomechanical research of knee joint during the process of running. *Journal of Southeast University*, 33(01), pp.27-32, 2017. DOI: 10.3969/j.issn.1003-7985.2017.01.005 (in Chinese)
- [37] Dos' Santos, T., Thomas, C., Comfort, P., & Jones, P. A. The effect of angle and velocity on change of direction biomechanics: An angle-velocity trade-off. *Sports medicine*, 48, pp.2235-2253, 2018. DOI: 10.1007/s40279-018-0968-3.
- [38] Cui, X., Ren, Z. Influence of running speed on gait of the middle and long distance runners. *Journal of Shenyang Sport University*, 33(03), pp.122-126, 2014. (in Chinese)
- [39] Amini Aghdam, S., Epro, G., James, D., Karamanidis, K. Leaning the Trunk Forward Decreases Patellofemoral Joint Loading During Uneven Running. *The Journal of Strength & Conditioning Research*. 36(12), pp.3345-3351, 2022. DOI: 10.1519/JSC.00000000000004128
- [40] Cronin, J., Hansen, K., Kawamori, N., McNair, P. Effects of weighted vests and sled towing on sprint kinematics. *Sports Biomechanics*, 7(2), pp.160-172, 2008. DOI: 10.1080/14763140701841381
- [41] Yamaguchi, T. A Cinematographical Analysis of Middle Distance Running Form. *Bulletin of Tokyo Women's College of Physical Education* (10), pp.43-53, 1975. (in Japanese)
- [42] Zhang, L., Liu, G., Han, B., Wang, Z., Yan, Y., Ma, J., Wei, P. Knee joint biomechanics in physiological conditions and how pathologies can affect it: a systematic review. *Applied bionics and biomechanics*, 2020. DOI: 10.1155/2020/7451683
- [43] Hamner, S. R., Delp, S. L. Muscle contributions to fore-aft and vertical body mass center accelerations over a range of running speeds. *Journal of biomechanics*, 46(4), pp.780-787, 2013. DOI: 10.1016/j.jbiomech.2012.11.024
- [44] Ma, Y., Ding H. Characteristic Angle of Human Knee Joint in Walking. *Forensic Science and Technology*, 44(4), pp.322-326, 2019. DOI: 10.16467/j.1008-3650.2019.04.009 (in Chinese)
- [45] Máca, J., & Štěpánek, J. Pedestrian load models of footbridges. In *MATEC Web of Conferences* (Vol. 107, p. 00009). EDP Sciences, 2017. DOI: 10.1051/mateconf/201710700009
- [46] Danion, F., Varraine, E., Bonnard, M., & Pailhous, J. Stride variability in human gait: the effect of stride frequency and stride length. *Gait & posture*, 18(1), pp.69-77, 2003. DOI:

10.1016/S0966-6362(03)00030-4

- [47] Högberg, P. Length of stride, stride frequency, “flight” period and maximum distance between the feet during running with different speeds. *Arbeitsphysiologie*, 14(6), pp.431-436, 1952. DOI: 10.1007/BF00934422
- [48] Luhtanen, P., Komi, P. V. Mechanical factors influencing running speed. *Biomechanics VI-B*, 2, pp.23-28, 1978.
- [49] Van Oeveren, B. T., de Ruiter, C. J., Beek, P. J., & van Dieën, J. H. The biomechanics of running and running styles: a synthesis. *Sports biomechanics*, 1-39, 2021. DOI: 10.1080/14763141.2021.1873411
- [50] Yang, Y., Enomoto, Y. The pace pattern of the men’s 1500 m race with different levels. *Journal of Physical Education & Sport*, 22(10). pp.2549-2556, 2022. DOI: 10.7752/jpes.2022.10323
- [51] Strohrmann C., Harms H., Kappeler-Setz C., Tröster G. Monitoring kinematic changes with fatigue in running using body-worn sensors. *IEEE Trans. Inf. Technol. Biomed*, 16: pp.983–990, 2012. DOI: 10.1109/TITB.2012.2201950
- [52] Möhler, F., Fadillioglu, C., Scheffler, L., Müller, H., & Stein, T. Running-Induced Fatigue Changes the Structure of Motor Variability in Novice Runners. *Biology*, 11(6), pp. 942, 2022. DOI: 10.3390/biology11060942