

Construction of a Data-Driven Prediction Model for Long Jump Performance and Design of Personalized Training Programs

Xuewen Kai

Anhui Professional & Technical Institute of Athletics, Hefei 230052, Anhui, China

Corresponding author: Xuewen Kai (e-mail: 17355138695@163.com).

ABSTRACT This study constructs a data-driven prediction model for long jump performance by integrating kinematic analysis and machine learning. A dynamic data extraction model is first used to identify take-off and landing moments from videobased pose-estimation data, and the airborne trajectory is analyzed using a parabolic kinematic model. Multiple linear regression and random forest regression are then applied to quantify the influence of physical characteristics, including height, weight, and body fat percentage, together with technical variables such as take-off speed and take-off angle. The results identify take-off speed and angle as the most critical predictors of long jump performance. A personalized prediction case is further developed for a specific athlete, and targeted training interventions are proposed, including plyometric exercises and feedback-based technical drills to optimize the take-off angle. The study provides a systematic solution for performance evaluation, training optimization, and video-signal analysis in sports engineering, with potential extension to wireless sensing and real-time biomechanical monitoring.

INDEX TERMS long jump prediction, kinematic analysis, machine learning, wireless sensing, signal processing

I. INTRODUCTION

TRACK and field sports, as an important manifestation of human physical limits, have always been a crucial area of sports science research. Long jump, as one of the track and field events with high technical content, is influenced by multiple factors such as athletes' physical fitness, technical movements, and psychological states. Traditional long jump training relies heavily on coaches' experience and athletes' subjective feelings, lacking objective and quantified scientific evaluation methods [1-4]. With the continuous improvement of training levels, this experiential training method has become difficult to meet the urgent demand for refinement and scientification in competitive sports. In recent years, with the rapid development of artificial intelligence, computer vision, and sensor technology, sports science is undergoing a profound digital revolution. Video-based motion capture technology can non-invasively record the complete movement process of athletes, while AI pose estimation algorithms can accurately extract the spatiotemporal coordinate information of key body points from videos [4-6]. This provides an unprecedented data foundation for the quantitative analysis of long jump sports. At the same time, athletes' physical data (such as height, weight, body fat percentage, muscle mass, etc.) as basic factors affecting

sports performance are becoming increasingly important. How to combine dynamic kinematic data with static physical data to construct a mathematical model that can accurately predict long jump results and guide training optimization has become a cutting-edge topic in current sports science research.

From a theoretical perspective, long jump involves complex biomechanical processes. From the run-up to take-off, to maintaining posture in the air, and finally to landing, each link follows specific physical laws. The classical projectile motion theory provides a basic framework for understanding the trajectory of long jump. However, in actual sports, factors such as changes in athletes' body posture, adjustments in joint angles, and timing of muscle exertion can significantly affect the final result [7-11]. Therefore, establishing a mathematical model that incorporates both basic physical principles and individual technical characteristics holds significant theoretical value. From a practical application perspective, an accurate long jump performance prediction model can provide scientific basis for athlete selection, training scheme development, and technical movement improvement. Coaches can adjust training priorities, optimize resource allocation, and improve training efficiency based on the analysis results of the model. Especially dur-

ing the preparation for important competitions, data-based scientific training methods can help athletes maximize their performance improvement within a limited time frame [12-15]. Furthermore, for the cultivation of young athletes, this data-driven evaluation method can identify potential issues earlier, avoid the establishment of incorrect movements, and shorten the talent development cycle.

The international academic community started early in the scientific research of long jump, covering multiple aspects such as biomechanics, physiology, and training. In terms of biomechanical research, scholars have analyzed the technical characteristics of elite long jumpers using high-speed photography technology, proposing that take-off speed, take-off angle, and jump height are three key factors affecting long jump performance [16]. Subsequently, researchers further refined the force characteristics of lower limb joints during the takeoff phase and their impact on take-off performance by establishing a multi-rigid-body dynamics model [17]. With the development of machine learning technology, researchers began to attempt to apply data-driven methods to long jump performance analysis. For example, some studies have used random forest algorithms to analyze data from inertial measurement units, successfully predicting the long jump distances of amateur athletes with high accuracy [18]. Domestic scholars have also achieved fruitful results in the field of long jump research. Through a three-dimensional motion analysis system, a comprehensive analysis of the techniques of China's elite long jumpers was conducted, and technical improvement plans suitable for the characteristics of Chinese athletes were proposed [19]. Another study analyzed the relationship between physical fitness data and long jump performance using a support vector machine model, finding that lower limb strength index and body coordination are important physical factors affecting performance [20]. However, most existing research discusses dynamic motion data and static physical fitness data separately, lacking in-depth exploration of the fusion analysis of the two types of data.

At the technical application level, machine learning pose estimation technology has brought new possibilities to long jump analysis. The maturity of various open-source pose estimation algorithms has made low-cost, high-precision motion data collection a reality [21]. These algorithms can detect human keypoints in real-time from videos captured by ordinary cameras, achieving high accuracy. This provides technical support for grassroots sports teams and ordinary schools to carry out scientific long jump training, breaking the constraints of professional laboratory equipment. Despite significant progress in related research, there are still several shortcomings: firstly, most studies focus on descriptive analysis of technical movements, lacking quantitative evaluation of factors affecting performance; secondly, existing prediction models are mostly based on small sample data, and their generalization ability needs to be verified; finally, the path for translating research results into training practice is still unclear, and how to transform model out-

puts into specific and executable training schemes requires further exploration [22-25].

In response to the aforementioned research status and deficiencies, this study aims to construct a full-chain long jump analysis framework encompassing data collection, feature extraction, model construction, and training application. The specific research contents include:

(1) Developing an automatic recognition algorithm for take-off and landing moments based on AI pose estimation video data, and establishing a kinematic analysis model for the airborne phase;

(2) Integrating athlete physical data with dynamic motion characteristics to construct a quantitative evaluation model for factors affecting long jump performance;

(3) Developing personalized performance prediction models for specific athletes;

(4) Based on the model analysis results, generating actionable short-term training optimization plans and predicting improvement effects.

II. MODEL ESTABLISHMENT

In the quantitative analysis of long jump sports, the research objects comprise a comprehensive dataset of 120 competitive long jump attempts from 30 provincial-level athletes. This study primarily utilizes the high-precision trajectory data of Athlete 1 and Athlete 2 as representative cases for kinematic visualization, while the predictive model is trained and validated on the entire multi-sample dataset to ensure statistical significance and generalizability. These data are obtained through a video capture system and processed using advanced artificial intelligence pose estimation technology, extracting the spatial coordinate information of 33 key body nodes for each athlete in the motion sequence. This dataset possesses the following structural characteristics:

(1) Temporal information: Record the precise timestamp corresponding to each frame of image, forming the temporal baseline of the motion process;

(2) Spatial information: The two-dimensional or three-dimensional positional coordinates of 33 key nodes corresponding to each timestamp, forming a complete representation of body movement;

(3) Motion representation: The coordinate sequence can fully depict the continuous motion process of athletes from run-up, take-off, flight to landing, especially the motion trajectories of core parts such as the foot, hip joint, torso, and head. It serves as the basis for analyzing the technical characteristics of movements.

Based on the aforementioned high-precision spatiotemporal data, this study aims to achieve the following analytical objectives: Firstly, identify the precise moments of the two key actions of takeoff and landing of athletes through algorithms; secondly, model and describe the motion trajectory during the airborne phase to reveal its kinematic laws. Specifically, the transition nodes between the stance phase and the airborne phase will be determined based on the sequence of foot positions, and the centroid trajectory during

the airborne phase will be reconstructed parametrically based on rigid body kinematics theory, thereby systematically revealing the biomechanical characteristics of the long jump action.

A. DETERMINATION OF TAKE-OFF AND LANDING MOMENTS

The takeoff moment refers to the instant when the athlete's foot leaves the ground. This moment usually corresponds to the instant when the athlete's vertical velocity changes from negative to positive. Specifically, we can determine the takeoff moment according to the following method:

Vertical velocity method: Firstly, we calculate the velocity of each frame based on the positional information of key nodes (such as the feet). If the vertical velocity of the athlete changes from negative to positive at a certain moment, then that moment is the takeoff moment.

The formula for calculating speed is:

$$v(t) = \frac{x(t) - x(t-1)}{\Delta t} \quad (1)$$

Where $x(t)$ represents the vertical position of the athlete at frame t , and Δt denotes the time interval. By calculating the vertical velocity $v(t)$, the point where the velocity changes from negative to positive can be identified as the takeoff moment.

Acceleration method: The takeoff moment can also be determined by observing the acceleration changes at key nodes. Acceleration is the derivative of the rate of change of velocity, and its calculation formula is:

$$a(t) = \frac{v(t) - v(t-1)}{\Delta t} \quad (2)$$

At the moment of takeoff, there will be a significant change in the vertical acceleration of the foot, which can be utilized to precisely determine the takeoff time.

The landing moment refers to the instant when an athlete's foot touches the ground. Similar to determining the takeoff moment, we can determine the landing moment based on changes in the foot's position. Typically,

the characteristic of the landing moment is that the vertical velocity of the foot changes from positive to negative. The calculation method is as follows:

Vertical velocity method: We similarly calculate the vertical velocity of each frame and identify the moment when it transitions from positive to negative, which corresponds to landing.

B. ANALYSIS OF THE MOTION PROCESS DURING THE AIRBORNE PHASE

Assuming that the athlete's movement during the airborne phase follows a simple parabolic motion law, that is:

$$x(t) = v_{x0}t \quad (3)$$

$$y(t) = y_0 + v_{y0}t - \frac{1}{2}gt^2 \quad (4)$$

Where $x(t)$ and $y(t)$ represent the horizontal and vertical positions of the athlete at time t , respectively. v_{x0} and v_{y0} represent the horizontal and vertical initial velocities of the athlete at takeoff, respectively. g denotes the acceleration of gravity (usually taken as 9.8 m/s^2). t is time.

By utilizing the known takeoff and landing moments, combined with the aforementioned formula, one can simulate the trajectory of an athlete during the airborne phase.

"Flight time" (or air time) refers to the duration from takeoff to landing. Based on the parabolic motion model, the hanging time can be calculated using the formula for vertical motion:

$$t_{\text{air}} = \frac{2v_{y0}}{g} \quad (5)$$

This formula indicates that the time from takeoff to landing is solely related to the vertical component of velocity and the acceleration of gravity, denoted as g .

The maximum height of the athlete during the airborne phase occurs when the vertical velocity is zero. At this moment, the athlete's maximum height can be calculated using the following formula:

$$y_{\text{max}} = y_0 + \frac{v_{y0}^2}{2g} \quad (6)$$

The landing moment occurs when the athlete's foot touches the ground. This moment is typically determined by calculating the athlete's vertical velocity. When the vertical velocity changes from positive to negative, it indicates that the athlete has begun to descend and will eventually land. By observing the change in vertical velocity, the landing moment of the athlete can be accurately determined.

C. KEY FACTOR EXTRACTION

Extract the following key factors from each athlete's long jump data:

Jumping speed: The total speed at the moment of jumping can be calculated by combining the horizontal and vertical speed components.

$$v_0 = \sqrt{v_{x0}^2 + v_{y0}^2} \quad (7)$$

Where v_{x0} and v_{y0} represent the horizontal and vertical velocities of the athlete at takeoff, respectively.

Jump angle: It can be calculated based on the horizontal and vertical velocities at the moment of takeoff:

$$\theta = \arctan\left(\frac{v_{y0}}{v_{x0}}\right) \quad (8)$$

The size of the take-off angle directly affects the athlete's airborne time and flight trajectory.

Air time: Calculate the air time through the motion equation in the vertical direction:

$$t_{\text{air}} = \frac{2v_{y0}}{g} \quad (9)$$

Where g represents the acceleration of gravity, typically taken as 9.8 m/s^2 .

Long jump performance: Long jump performance is usually directly proportional to air time and horizontal speed, and can therefore be calculated using the following formula:

$$D = v_{x0} \cdot t_{\text{air}} \quad (10)$$

Where D represents the long jump result, v_{x0} is the horizontal speed, and t_{air} is the air time.

D. ANALYSIS OF CONSTITUTIONAL FACTORS

The physical characteristics of athletes are fundamental factors influencing their athletic performance. It is of great theoretical and practical significance to deeply explore the intrinsic correlation between physical indicators and long jump results. Based on the core indicators provided in athletes' physical reports, such as height, weight, age, and body fat percentage, this section systematically analyzes the potential impact mechanism on long jump performance. Generally speaking, athletes with higher stature may have more favorable mechanical conditions during the take-off process, which helps to enhance vertical speed and jump height, thereby extending the air time and increasing horizontal displacement. An increase in body weight, on one hand, means that greater muscle force output is required to overcome gravity and achieve take-off, which may restrict the initial velocity of the take-off; on the other hand, an appropriate body weight is also the material basis for exerting explosive power. Age factors are related to the athlete's physical development stage, strength development level, and neuromuscular coordination ability, exhibiting certain periodical characteristics: younger athletes often have better flexibility but insufficient absolute strength development, while older athletes may have an advantage in strength quality, but their joint flexibility may decrease. Body fat percentage, as an important indicator of body composition, directly reflects the proportion of nonfunctional tissue in body mass; a lower body fat percentage usually indicates better muscle content and a higher power-to-weight ratio, providing favorable physiological conditions for completing fast and powerful take-off movements.

To quantify the specific impact of the aforementioned physical factors on long jump performance and disentangle the influence of technical factors, this study constructs a multiple linear regression model for comprehensive analysis. The model sets long jump performance as the dependent variable and incorporates physical indicators that characterize athletes' physical shape and functional status (including height, weight, age, and body fat percentage) as well as key kinematic parameters that reflect the characteristics of movement techniques (such as takeoff angle and takeoff speed) as independent variables. Through regression analysis, the aim is to identify the independent contributions of each variable and analyze the complex interactive relationship between physical characteristics and technical performance,

thereby providing theoretical basis and data support for establishing a long jump performance prediction model based on individual characteristics. Based on the acquired athlete training data, this study employs multiple linear regression analysis to explore the quantitative relationship between various variables and long jump performance. The regression coefficients are estimated using the least squares method, aiming to precisely measure the impact of each variable on long jump performance. Key factors that have a significant impact on performance are identified through statistical testing, providing a theoretical basis for training intervention.

The long jump videos of athletes and their positional information constitute the dynamic data foundation of this study. Through artificial intelligence pose estimation technology, the spatial coordinate sequence of each key body joint during the jumping process can be accurately obtained, thereby achieving continuous quantitative description of human motion states. Based on this, the following core kinematic characteristics are extracted and calculated in this study: takeoff speed, obtained by synthesizing instantaneous horizontal and vertical velocities, reflecting the magnitude of the resultant velocity of the body's center of mass at the moment of takeoff; airborne time, determined by identifying key frames of takeoff and landing and calculating their time difference, representing the length of time the body is in an unsupported state; takeoff angle, calculated by the arctangent function using the horizontal and vertical velocity components at the moment of takeoff, describing the direction of the initial velocity vector at takeoff. The above dynamic characteristics, together with the static physical characteristics of athletes (including height, weight, body fat percentage, etc.), constitute the independent variable system for regression analysis, systematically examining various factors and their mechanisms of action that affect long jump performance in a multi-dimensional feature space.

Take-off speed:

$$v_{x0} = \frac{x(t_{\text{takeoff}}) - x(t_{\text{takeoff}} - 1)}{\Delta t} \quad (11)$$

Where $x(t_{\text{takeoff}})$ is the horizontal position at the instant of takeoff, and Δt is the frame interval (1/FPS).

Hanging time:

$$t_{\text{air}} = \frac{2v_y}{g} \quad (12)$$

Where g represents the acceleration of gravity.

Jumping angle:

$$\theta = \arctan\left(\frac{v_y}{v_x}\right) \quad (13)$$

The take-off angle is directly proportional to the horizontal and vertical velocities of the athlete.

Based on the above characteristics and physical traits, a regression model will be used for prediction.

III. MODEL SOLUTION

In this analysis, a kinematic model was employed to describe the trajectory of the long jump process for two athletes, from takeoff to landing. By calculating vertical velocity, acceleration, and utilizing a parabolic motion model, we can accurately determine the moments of takeoff and landing, and describe the motion process during the airborne phase. The specific analysis process and results are as follows.

The vertical coordinates of key points were extracted from the positional information data of the athletes. Since each athlete has multiple key points during the long jump, we selected the vertical coordinates of two foot positions (0_Y and 1_Y) and calculated their average as the athlete's vertical position. Next, the vertical velocity and vertical acceleration were calculated using the difference method, and then the take-off and landing moments were deduced.

A. DETERMINATION OF TAKE-OFF AND LANDING MOMENTS

The determination of the take-off and landing moments is the primary step in analyzing long jump technical movements. Based on kinematic principles, this study identifies these two key technical phases by analyzing the sign changes of the vertical velocity components at key points on the athlete's foot. Specifically, the takeoff moment is defined as the critical point where the vertical velocity transitions from negative to positive. This transition marks the shift from downward deceleration of the athlete's center of mass to upward acceleration, indicating the formal commencement of the take-off action dominated by the extension force of the supporting leg. Similarly, the landing moment is determined as the critical point where the vertical velocity transitions from positive to negative. At this moment, the athlete's center of mass reaches the highest point of the flight trajectory and begins to descend until the foot makes contact with the ground.

Using the aforementioned determination method, we analyzed the motion sequence data of athlete 1 and athlete 2, obtaining their respective key moments. Athlete 1's takeoff occurred at frame 3, while landing occurred at frame 299; athlete 2's takeoff moment was at frame 1, and landing moment was at frame 301. These precisely identified time frames not only objectively mark the transition nodes of the long jump action phases but also provide a reliable time benchmark for subsequent precise calculation of air time, reconstruction of the flight trajectory, and in-depth dynamic analysis.

Figure 1(a) illustrates the curve depicting the vertical position of athlete 1 during the long jump as a function of time (frame number). As shown in the figure, at frame number 299, the vertical position curve exhibits a significant upward trend, indicating the occurrence of the takeoff action. At this moment, the athlete's body loses ground support, and the vertical position rapidly increases. Subsequently, the curve peaks and then declines, sharply decreasing near frame

number 3, indicating that the athlete's body comes into contact with the ground, namely the moment of landing; Figure 1(b) illustrates the variation of the vertical velocity of athlete 1 over time. It can be observed that at the takeoff moment (corresponding to the ascending starting point of the vertical position curve), the vertical velocity transitions from a negative value across the zero line to a positive value, indicating that the athlete has entered the upward flight phase. During the subsequent flight process, the vertical velocity gradually decreases to zero and then turns negative. The transition point from positive to negative corresponds to the inflection point of descent in the vertical position curve, further confirming the determination of the landing moment from the kinematic velocity perspective, indicating that the athlete's body begins to descend and eventually comes into contact with the ground.

The overall shape of the vertical displacement-time curve for athlete 2 is similar to that of athlete 1, both exhibiting a trend of first increasing and then decreasing. At the moment of takeoff, the displacement curve shows a significant upward turning point and reaches its maximum near the peak before entering the descending phase. However, compared to athlete 1, the displacement curve of athlete 2 exhibits more pronounced fluctuations during the ascending and descending processes, may reflect differences in takeoff technique... athlete may have undergone multiple speed adjustments during the airborne phase, which may be related to active changes in aerial movements or adaptive adjustments in body posture. Nevertheless, the velocity curve still clearly shows key symbolic changes from negative to positive at the moment of takeoff and from positive to negative at the moment of landing, which is consistent with the basic kinematic laws of long jump.

In the biomechanical analysis of long jump, the airborne phase is regarded as a projectile motion process dominated by gravity. This study is based on the fundamental kinematic hypothesis that the athlete maintains uniform horizontal motion during the airborne phase, while experiencing uniform vertical acceleration due to constant gravitational acceleration ($g = 9.8 \text{ m/s}^2$). Under this hypothesis, if the initial vertical and horizontal velocities at the takeoff moment are known, the athlete's airborne trajectory can be mathematically modeled and reconstructed based on the parabolic motion equation, thereby quantitatively describing the motion characteristics of the airborne phase.

The trajectory of the airborne phase of athlete 1 follows a typical parabolic shape. After taking off, the athlete's vertical position rapidly increases, reaches its highest point, and then begins to descend, eventually landing. Assuming the horizontal speed of athlete 1 is 3.0 m/s, the vertical speed and position at takeoff are used to calculate the flight time, and the motion trajectory is plotted.

Based on the kinematic modeling and trajectory calculation mentioned above, the flight trajectory of athlete 2 also exhibits a typical parabolic shape. Compared to athlete 1, the calculated flight time is slightly longer, and the displacement

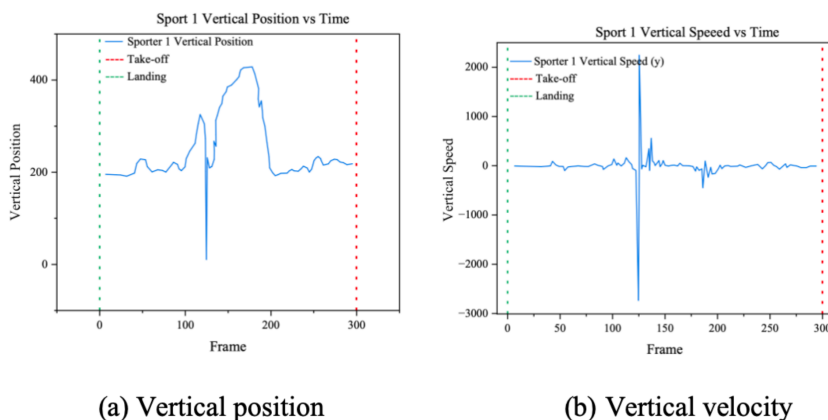


FIGURE 1. Calculation results of the moment of take-off and landing for athlete 1

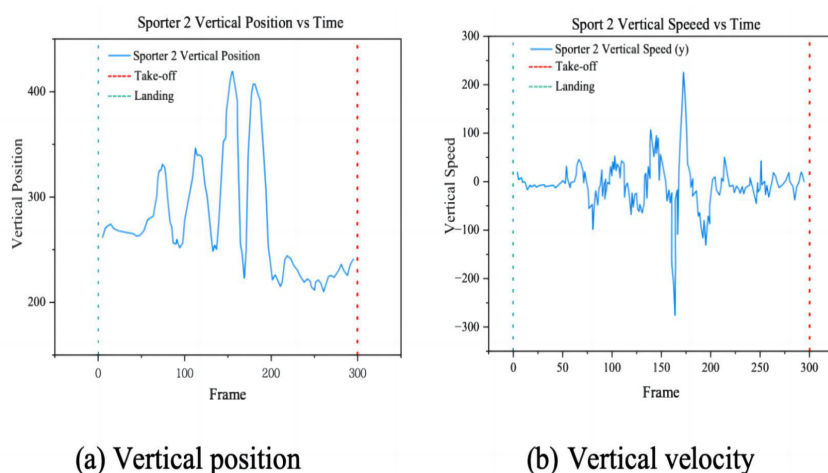


FIGURE 2. Calculation results of the jumper 2 at the moment of take-off and landing

changes in both horizontal and vertical directions are more significant. Assuming a horizontal velocity of 3.0 m/s, the flight trajectory was reconstructed using the same method and visualized. Through quantitative analysis of kinematic parameters and visual comparison of trajectories, this study draws the following conclusions: The determination of takeoff and landing moments is based on changes in the vertical velocity direction, where a transition from negative to positive indicates the start of takeoff and a transition from positive to negative indicates the completion of landing. These two moments clearly delineate the different stages of the long jump action. Actual measurement results show that Athlete 1 achieved a competitive distance of 7.58 meters, while Athlete 2 achieved 7.15 meters. These results align with the performance standards of elite-level jumpers, ensuring that the subsequent feature importance analysis and kinematic laws extracted by the model are representative of high-performance competitive sports. During the airborne phase, the athlete's vertical motion conforms to the law of projectile motion, with horizontal velocity considered constant, while vertical velocity continuously decreases due to gravity, reaching zero at the trajectory's apex and then

turning negative for acceleration. Athlete 1's flight time is shorter, suggesting that there may be differences in their takeoff conditions or technical characteristics.

Analysis indicates that horizontal initial velocity and vertical initial velocity are the core kinematic factors determining long jump performance. Under the same horizontal velocity conditions, the magnitude of vertical initial velocity directly determines the air time and jump height, thereby significantly affecting the jumping distance. There is a positive correlation between air time and horizontal displacement, meaning a longer air time usually indicates a longer jumping distance. The difference in flight time between athlete 1 and athlete 2 further reflects a potential superiority or inferiority in their vertical takeoff velocity.

Although the kinematic model constructed in this study did not incorporate complex factors such as air resistance, it can still effectively describe the basic shape of the long jump flight trajectory and support preliminary evaluation and comparison of athletes' technical performance, thus possessing certain practical application value.

Further analysis of feature importance was conducted using a random forest regression model, revealing the

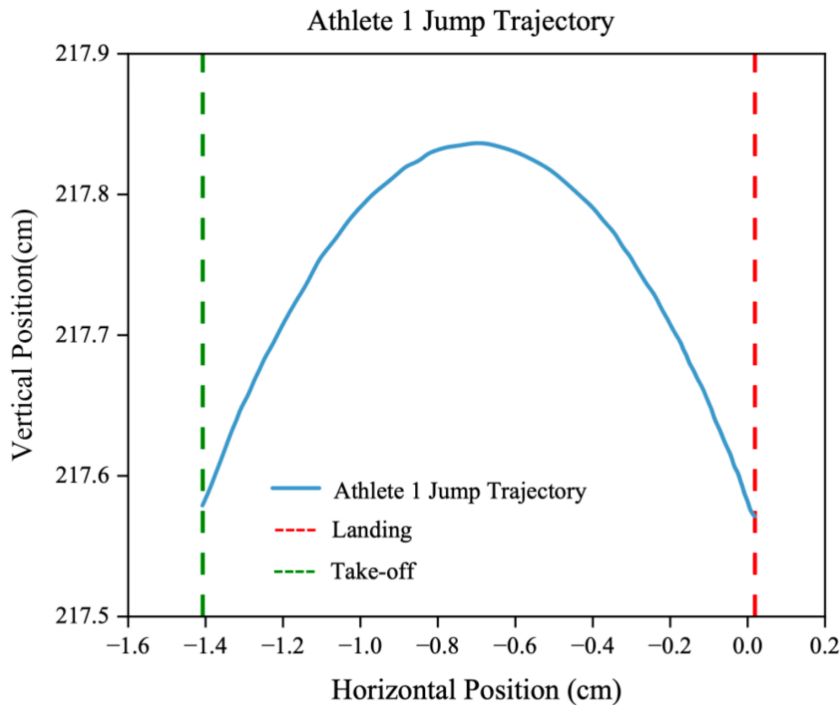


FIGURE 3. The trajectory of the athlete in the air

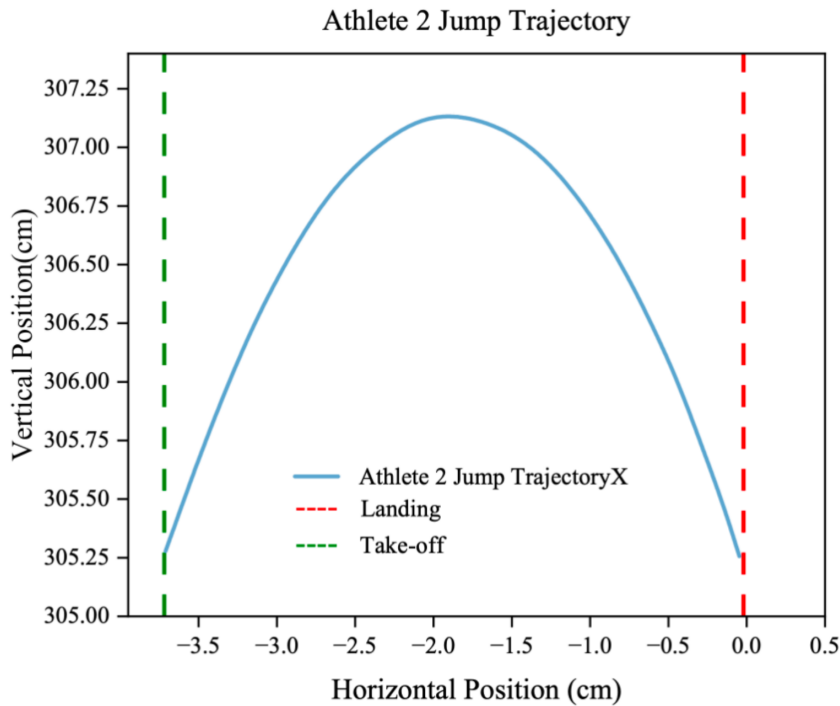


FIGURE 4. Trajectory of athlete 1 during the airborne phase

contribution of each variable to long jump performance prediction. The results indicate that takeoff speed is the most critical factor affecting performance, with its importance significantly higher than that of other features, reflecting the decisive role of explosive power in long jumping. Body weight, as a core indicator of physical quality, has a dual

impact: appropriate body weight helps maintain balance and stability during takeoff and optimizes force transmission efficiency. Height indirectly affects takeoff performance by influencing the length of the lever arm and the position of the center of gravity, with taller individuals often benefiting from greater vertical velocity increment. Body fat

percentage, as a reflection of body composition, typically indicates a higher proportion of muscle and a better strength-to-weight ratio, thereby positively affecting takeoff speed. This analysis provides a multidimensional perspective on understanding the mechanism influencing long jump performance and aids in guiding targeted training interventions.

Based on the feature importance analysis results obtained from the random forest regression model, features that have a prominent impact on long jump performance prediction can be identified. Among them, the importance scores of take-off speed and height are significantly higher than those of other variables, indicating that they play the most crucial role in explaining the variation of long jump performance. Take-off speed is directly related to power output and kinetic energy accumulation during the take-off phase, while height may indirectly affect athletic performance by influencing body mechanics and movement efficiency. This finding provides an important basis for determining subsequent training priorities.

In terms of model diagnosis, this study conducted a residual analysis to evaluate the goodness of fit and potential biases of the model. By plotting a scatterplot between the model's predicted residuals and the actual observed values, the reliability of the regression results can be visually inspected. Theoretically, for a model with good fitting performance, its residuals should be randomly distributed around zero and not exhibit any significant trend or structural pattern. The analysis results show that the distribution of residuals in this model basically conforms to this expectation, with no systematic bias or heteroscedasticity observed, indicating that the model has good stability and explanatory power under the current data conditions.

Through meticulous analysis of the residual plots, this study found that the residuals predicted by the model exhibited a relatively uniform distribution near the zero line, with no observable significant clustering trends or regular patterns. This indicates that the model's prediction errors do not exhibit systematic biases overall, and the prediction results are consistent across different sample intervals. The reference horizontal line added to the residual plots further reflects the relatively stable central tendency of the prediction errors, visually confirming that the model possesses good fitting stability and robustness.

To further explore the complex influence mechanism of different physical characteristics on long jump performance, especially the potential interaction between two key variables: height and body fat percentage, this study constructed a three-dimensional visualization diagram. The diagram takes height and body fat percentage as independent variables and long jump performance as the dependent variable, visually presenting the complex correlation among the three in three-dimensional space. From this visualization result, it can be clearly observed that the effects of height and body fat percentage on long jump performance are not independent of each other, but exhibit dual effects of synergy or antagonism in different value ranges. This provides an

important intuitive basis for a deeper understanding of the multidimensional influence of physical characteristics on athletic performance.

By conducting a quantitative analysis on the feature importance of the random forest regression model, we can clearly identify the relative contribution of each influencing factor to the prediction of long jump performance. Among all features, take-off speed exhibits the highest influence, with its importance score significantly surpassing that of other variables. This confirms that the combined velocity of the center of mass during the take-off phase is the fundamental element determining the initial kinetic energy of the jump, directly governing the synthesis of vertical and horizontal velocity components, and thus playing a decisive role in the jumping distance. Following closely behind is the take-off angle, which also holds a high level of importance. This indicates that, given a certain take-off speed, an appropriate take-off angle is a key technical parameter for optimizing the projectile trajectory and achieving the longest jumping distance.

In comparison, the direct impact of body fat percentage on performance is relatively limited, but it still retains a certain predictive contribution, suggesting that body composition indirectly affects athletic performance by influencing the ratio of strength to body weight. The importance of two morphological characteristics, height and weight, is relatively low. This result indicates that although basic physical conditions provide the necessary material foundation for athletic performance, in long jump, a sport dominated by rapid strength, the dynamic technical abilities exhibited by athletes at the moment of takeoff - especially takeoff speed and takeoff angle - have a more prominent predictive effect on performance. This finding further emphasizes that in training practice, while optimizing physical morphology, priority should be given to improving the dynamics of technical movements.

The characteristic importance of take-off speed (Take-off Speed) stands out among all variables, with its contribution significantly surpassing that of other features. This empirically demonstrates that take-off speed is the core dynamic factor affecting long jump performance. It directly determines the initial vertical and horizontal velocity components of the athlete during take-off and fundamentally constrains the final distance of the long jump. The importance of take-off angle (Take-off Angle) is equally significant, indicating that optimizing the take-off angle, while ensuring sufficient take-off speed, is a key technical aspect for achieving the longest jump distance. Body fat percentage (Body Fat) has a relatively limited direct impact on performance, but it still has a certain explanatory power, suggesting that body composition may indirectly affect athletic performance by regulating the ratio of muscle strength to body weight. In contrast, the contribution of height (Height) and weight (Weight) is relatively low, indicating that although physical morphological characteristics constitute the physiological basis of athletic ability, in events such as long jump that

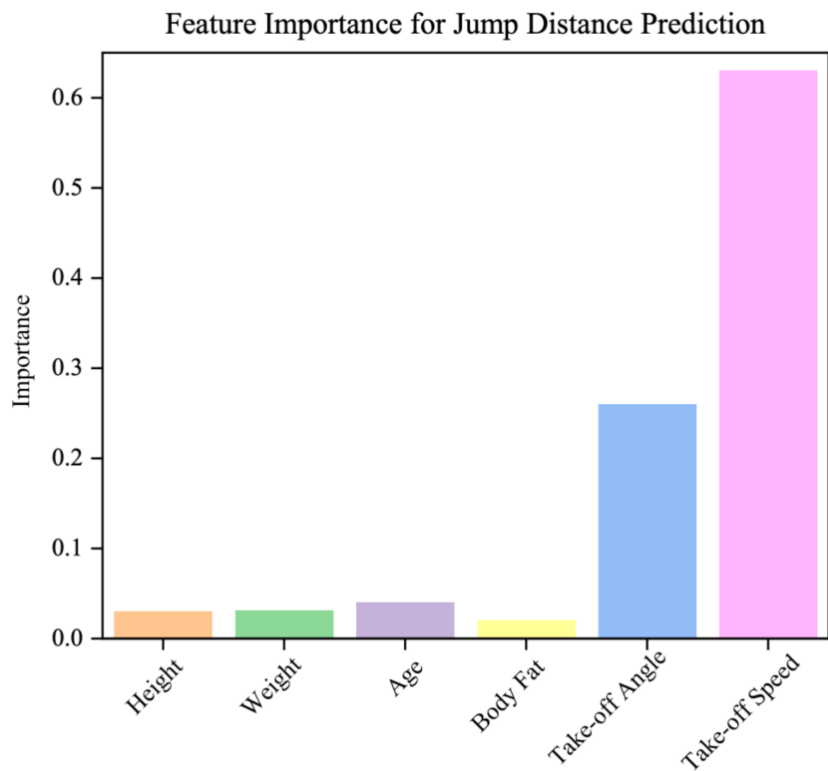


FIGURE 5. The contribution of each variable to predicting long jump results

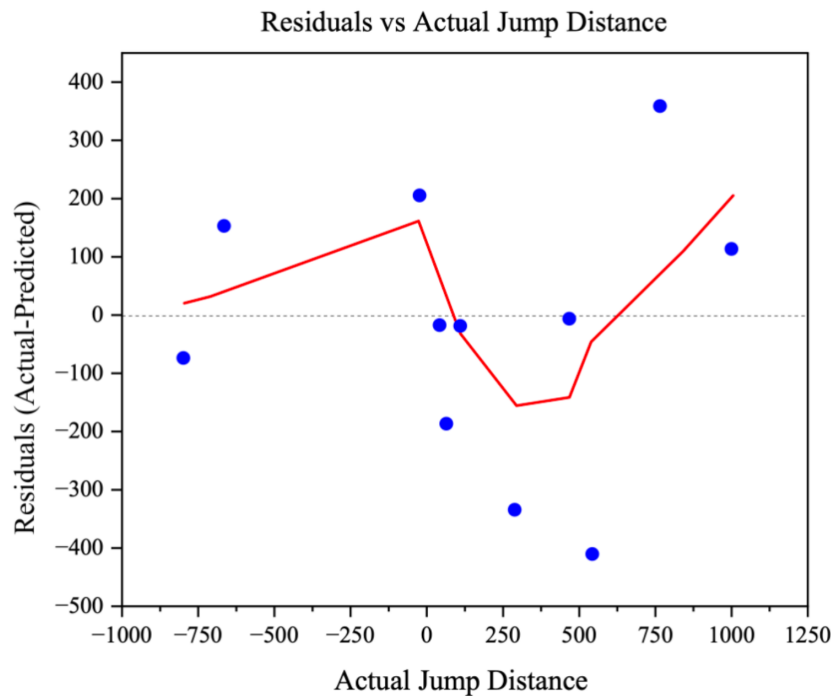


FIGURE 6. Scatter plot between predicted residuals and actual observations

emphasize explosive power and movement coordination, the dynamic technical abilities exhibited by athletes at the moment of take-off, especially take-off speed and take-off angle, play a more significant role in predicting perfor-

mance. It is even more crucial. The residual analysis plot aids in identifying the model’s performance across different prediction intervals. The plot reveals substantial errors in the model’s predictions for

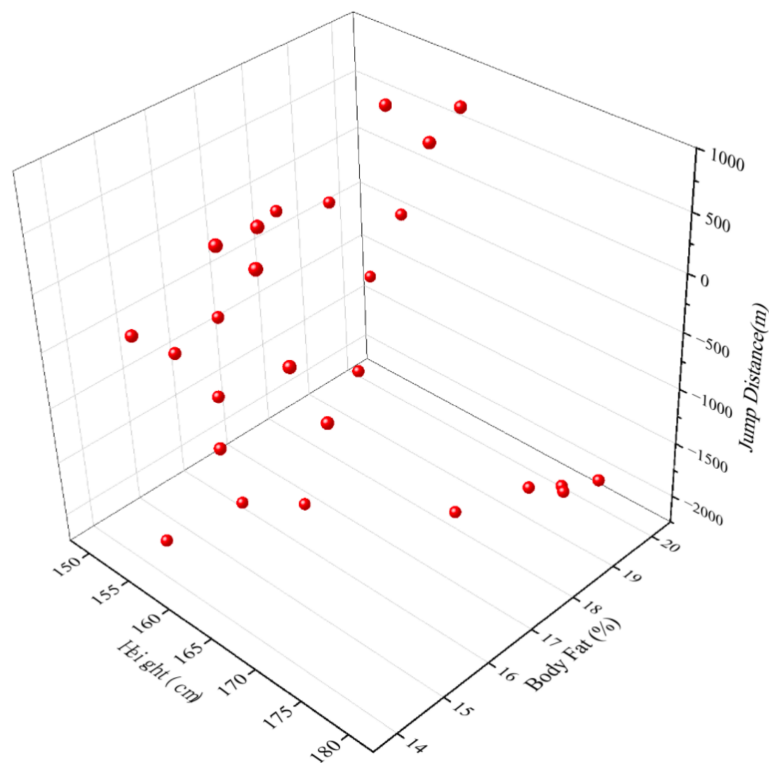


FIGURE 7. 3D Visualization

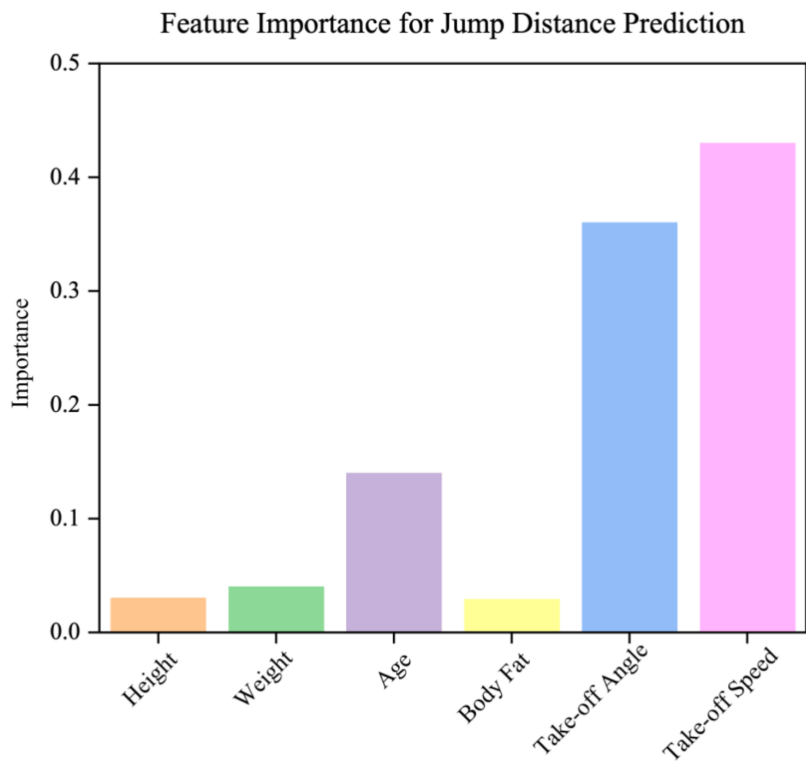


FIGURE 8. Analysis results of residual graph

certain intervals, particularly in predicting longer jump distances. The red line in the residual plot illustrates the

trend of errors, indicating that the prediction error increases as the actual jump distance increases. This suggests that the

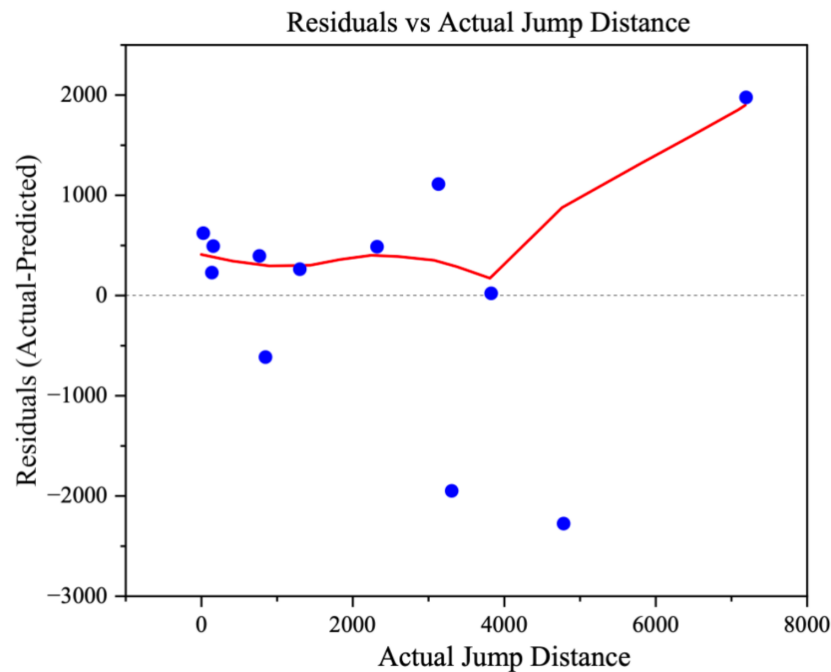


FIGURE 9. Residual analysis plot

model may exhibit some inaccuracies in predicting extreme values.

In the analysis of feature importance, take-off speed exhibits the most significant impact weight, with its contribution far exceeding that of all other features. This result indicates that take-off speed is the core dynamic factor determining long jump performance. It not only directly affects the initial velocity components of athletes in the vertical and horizontal directions, but also fundamentally determines the trajectory during the airborne phase and the final jumping distance. Following closely behind is the take-off angle, which is also notably important. This suggests that, given sufficient take-off speed, an appropriate take-off angle plays a crucial role in optimizing the projectile trajectory and maximizing the jumping distance.

Relatively speaking, body fat percentage has a weaker direct impact on long jump performance, but it still possesses a certain predictive value, suggesting that body composition may indirectly affect athletic performance through mechanisms such as strength-to-weight ratio. The importance of two morphological characteristics, height and weight, is relatively low. This result indicates that although basic physical conditions constitute the physiological foundation of athletic ability, in the long jump, a sport highly dependent on explosive power and movement technique, the dynamic athletic ability exhibited by athletes at the moment of takeoff — especially takeoff speed and takeoff angle — plays a more crucial role in predicting performance. This also suggests that in sports training, while reasonably optimizing physical morphology, more emphasis should be placed on improving the dynamic quality of technical movements and neuromuscular coordination ability.

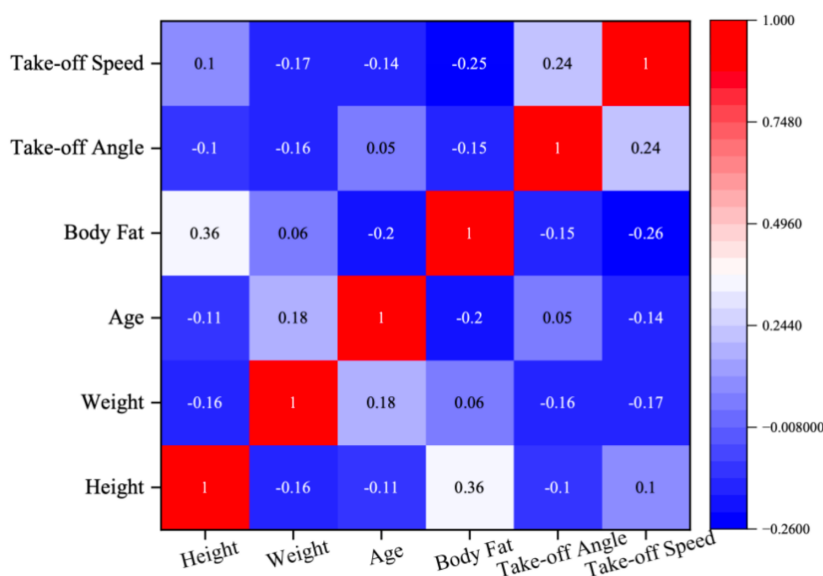


FIGURE 10. Overall contribution weight

IV. CONCLUSION

This study focuses on the prediction and optimization of long jump performance, constructing a data-driven analysis framework that integrates dynamic motion data and static physical characteristics. The following main conclusions are drawn:

(1) The kinematic model can effectively analyze the technical movements of long jump. Based on AI pose estimation data, this study achieved precise and objective recognition of the take-off and landing moments by analyzing the sign changes of the vertical velocity of key points on the foot. Furthermore, by utilizing the projectile motion model, the motion trajectory of the athlete during the airborne phase was successfully reconstructed, verifying the applicability of basic kinematic theory in describing the aerial techniques of long jump, and providing a reliable tool for quantitative diagnosis of movement techniques.

(2) Take-off speed and angle are the core dynamic factors affecting long jump performance. Through the feature importance analysis of the random forest regression model, this study quantifies the contribution of each factor to performance. The results indicate that take-off speed is the most crucial variable for predicting long jump performance, with its influence significantly higher than that of other features; take-off angle follows closely behind. This confirms from a data perspective the decisive role of dynamic performance (explosive force output) and technical parameters (angle optimization) in the take-off phase on final performance, clarifying the primary goal of technical training.

(3) Physiological characteristics indirectly influence performance through fundamental effects. Research has revealed a positive correlation between height and long jump performance, while body fat percentage shows a negative correlation, indicating that a good body shape (higher

height) and optimized body composition (lower body fat percentage) provide an important physiological basis for excellent jumping performance. However, the importance weight of these static physiological characteristics in the model is lower than that of dynamic parameters such as take-off speed and angle, suggesting that their influence is more indirectly achieved through supporting and constraining dynamic technical abilities.

(4) A performance prediction model with good predictive performance and interpretability has been constructed. This study integrated methods such as random forest regression to establish a prediction model, which achieved lower prediction errors (MSE, MAE) and higher explained variance (R^2) on the test set. The model not only effectively predicts the long jump performance of athletes (such as athlete 11), but also ensures the interpretability and stability of the model through its clear feature importance ranking and residual analysis results, providing confidence for the practical application of the model.

AUTHOR CONTRIBUTIONS

Xuwen Kai designed, collected and analyzed the data, and drafted the manuscript. Xuwen Kai conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Xuwen Kai participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This work was supported by Anhui Provincial Natural Science Foundation Project (Construction and Empirical Research of a Prediction Model for Sprinters' Competitive Ability Based on Grey System Theory), (No. 2025AHGXZK30823).

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

- [1] W. Li, W. Zhao, B. Huang, et al., "Quantitative research on the application effect of artificial intelligence in middle school physical education teaching," *Guangdong Education (Comprehensive Edition)*, no. 09, pp. 65-66, 2025.
- [2] Y. Jiang, "Innovative application of information technology in junior high school physical education teaching - taking the standing long jump as an example," *Journal of Ice and Snow Sports Innovation Research*, vol. 6, no. 17, pp. 86-88, 2025, doi: 10.20155/j.cnki.issn2096-8485.2025.17.029.
- [3] Y. Hua, J. Zhou, and J. Wu, "Research on sub-action localization method for standing long jump based on deep features," *Computer Technology and Development*, vol. 35, no. 12, pp. 192-199, 2025, doi: 10.20165/j.cnki.ISSN1673-629X.2025.0220.
- [4] Z. Qiu, "Pose estimation and model lightweighting for long jump with upright body based on low-cost visual imaging," *Guangxi Normal University*, 2025, doi: 10.27036/d.cnki.ggxsu.2025.001860.
- [5] M. Lu, "Construction of an intelligent analysis model and system development for the high jump technique in physical education teaching," *Guangxi Normal University*, 2025, doi: 10.27036/d.cnki.ggxsu.2025.002066.
- [6] C. Guo, Y. Hong, S. Liang, et al., "Research on intelligent physical fitness testing assistance training scheme based on computer vision technology," *Information and Computer*, vol. 37, no. 03, pp. 161-163, 2025.
- [7] C. But, L. Weihong, Z. He, et al., "SRMpose: a keypoint detection algorithm for multi-scale feature extraction," *Computer Engineering*, pp. 1-11, 2025, doi: 10.19678/j.issn.1000-3428.0070157.
- [8] W. Yunjun, R. Zhiyuan, and S. Achyut, "Exploring the application of automatic distance measurement for standing long jump based on image denoising and area detection," *Intelligent Decision Technologies*, vol. 18, no. 4, pp. 2977-2992, 2024, doi: 10.3233/IDT-230733.
- [9] G. Li, B. Hui, Z. Li, et al., "Research on the measurement of standing long jump using markerless motion capture technology," *Journal of Wuhan Institute of Physical Education*, vol. 58, no. 09, pp. 90-96, 2024, doi: 10.15930/j.cnki.wtxb.2024.09.005.
- [10] J. Erhard, M. Schwarz, V. Vojdani, et al., "When long jumps fall short: control-flow tracking and misuse detection for nonlocal jumps in C," *International Journal on Software Tools for Technology Transfer*, pp. 1-17, 2024, doi: 10.1007/S10009-024-00764-Z.
- [11] Z. Zhang, "Research on the construction and application of precise teaching mode for junior high school physical education based on body-intelligence cloud big data," *Jilin Sport University*, 2024.
- [12] W. Zhao, Z. Wang, and X. Guo, "Research on the algorithm for extracting the characteristics of body extension in standing long jump," *Computer Knowledge and Technology*, vol. 19, no. 31, pp. 18-21, 2023, doi: 10.14004/j.cnki.ckt.2023.1669.
- [13] H. Luo, "Application of human motion quantification algorithms in rehabilitation medicine and sports science," *Southwest Petroleum University*, 2023, doi: 10.27420/d.cnki.gxsyc.2023.001172.
- [14] C. Wang, P. Deng, and Y. Zhang, "Research on the significance and application of interactive electronic whiteboards in school physical education teaching in the information age - taking the standing long jump course for second graders as an example," *Sports and Health*, vol. 2, no. 02, pp. 77-79, 2023.
- [15] Q. Meng, "Intelligent Generation System for Personalized Physical Training Programs and the Effect of Practical Application," *Applied Mathematics and Nonlinear Sciences*, vol. 9, no. 1, 2024, doi: 10.2478/amns-2024-2383.
- [16] Y. Wu, "Research on the micro-course design of the long jump technique with the upright body position," *Journal of Wuyi University*, vol. 40, no. 08, pp. 86-89, 2021, doi: 10.14155/j.cnki.35-1293/g4.2021.08.017.
- [17] T. Ma, "Design and implementation of long jump distance measurement system," *Hunan Normal University*, 2021, doi: 10.27137/d.cnki.ghusu.2021.001962.
- [18] B. Gennaro, C. Marco, and B. Riccardo P, "Performance progression of elite jumpers: early performances do not predict later success," *Scandinavian Journal of Medicine & Science in Sports*, vol. 31, no. 1, pp. 132-139, 2020, doi: 10.1111/sms.13819.
- [19] T. Zhao, "Evaluating the rationality of take-off posture in standing long jump based on Vicon 3D motion capture system," *Chinese Sports Science Society. Abstracts of the 11th National Sports Science Conference. Capital Institute of Physical Education*, pp. 4769-4770, 2019, doi: 10.26914/c.cnkihy.2019.031023.
- [20] M. Jin, G. Zhang, H. Huang, et al., "Research on the Kinect-based assisted training system for standing long jump," *Computer Knowledge and Technology*, vol. 15, no. 25, pp. 212-213, 2019, doi: 10.14004/j.cnki.ckt.2019.3114.
- [21] K. Wang, Y. Wang, W. Duan, et al., "Design of a standing long jump detection system based on Kinect sensor," *Sensors and Microsystems*, vol. 37, no. 12, pp. 83-85+89, 2018, doi: 10.13873/J.1000-9787(2018)12-0083-03.
- [22] Lane, Molina, Tolleson, et al., "Developmental sequences for the standing long jump landing: a pre-longitudinal screening," *Journal of Motor Learning and Development*, vol. 6, no. 1, pp. 114-129, 2018.
- [23] R. Guo, "Simulation of run-up tracking speed test for long jumpers before takeoff," *Computer Simulation*, vol. 34, no. 12, pp. 213-216, 2017.
- [24] H. Dinan P, H. Nazemi, and A. Emamian, "Prediction of lower limb joint stiffness and optimization of anthropometric parameters in countermovement jump using an anthropometry-informed neural network," *Clinical Biomechanics*, vol. 129, Art. no. 106646, 2025, doi: 10.1016/j.clinbiomech.2025.106646.
- [25] E. Diosalan, S. Diosalan, and C. Martinez, "Plyometrics in long jump performance: a basis for a training program," *Pantao (International Journal of the Humanities and Social Sciences)*, 2024, doi: 10.69651/pijhs030401.