IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 8 October 2023, accepted 27 October 2023, date of publication 7 November 2023, date of current version 13 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3330757

RESEARCH ARTICLE

Normative Evaluation Method of Long Jump Action Based on Human Pose Estimation

XIUGANG GONG[1], XINYUAN GENG[1], GUANGJUN NIE[2], TAO WANG[1], JIAJUN ZHANG[1], AND JIEBING YOU[2]

[1]School of Computer Science and Technology, Shandong University of Technology, Zibo 255100, China
[2]Department of Neurology, Zibo Central Hospital, Zibo 255036, China

Corresponding author: Jiebing You (youjiebing2009@163.com)

This work was supported by the Shandong Provincial Natural Science Foundation under Grant ZR2020QF069.

ABSTRACT Addressing the issue of the lack of objective quantitative evaluation in training long jump events, this study presents a normative analysis method based on human pose estimation and similarity measures. By training a lightweight human pose estimation model, this method can run on low-delay embedded devices. In line with key movements, the proposed method designs a normative analysis for long jump actions, which yields a measurement of the movements' adherence to the standard and provides corrective suggestions. Experimental results indicate that the accuracy of this approach in analyzing the standardization of long jump action reaches 91.3%. As a result, it holds significant application value in various scenarios, including students' long jump training and correction of long jump techniques. Furthermore, it can be extended to other practical applications beyond sports.

INDEX TERMS Human pose estimation, similarity measures, action recognition and correction.

I. INTRODUCTION

In the realm of sports, the standardization of movements plays a pivotal role in determining the effectiveness of our training efforts. Failure to adhere to standardized actions can result in several grave consequences, such as the development of incorrect conditioned reflexes, heightened susceptibility to sports-related injuries, compromised muscle coordination transfer, diminished proprioceptive abilities, and a decline in overall training efficiency. Unfortunately, in conventional long jump training, teachers commonly resort to subjective assessments of students' actions, leaving students without a tangible means to gauge the adherence of their actions to the prescribed standards. Hence, there arises a pressing need for an objective method to comprehensively evaluate the standardization of long jump actions, bridging the gap between subjective judgment and concrete, measurable criteria in training assessment.

To analyze the standardization of long jump action, the initial step involves capturing the action posture. Early human action recognition necessitated the use of external equipment to perceive changes in human pose for action identification. For instance, Dowling et al. employed inertial sensor equipment to detect knee bending angles, trunk tilt, and thigh coronal velocity, applying this approach to the detection and recognition of ACL injuries [1]. Similarly, Pansiot et al. proposed a micro-sensor based on an accelerometer for swimming motion analysis [2]. This method involved recording pitch angle and side angle features extracted from acceleration to identify the posture and basic motion index of the human body, culminating in the development of a system to detect swimming performance, thus facilitating training guidance. Although sensor-based approaches offer precise inference of human actions, the inconvenience of having to wear sensors for each analysis remains a notable limitation.

The normativity of an action is typically determined by assessing its similarity to a standard action, and various methods have been proposed for calculating this similarity [3], [4], [5], [6], [7], [8]. In a previous study [9], the Euclidean distance was employed to measure similarity, where differences between nodes were compared based on the obtained coordinates. However, this approach necessitates that the motion

The associate editor coordinating the review of this manuscript and approving it for publication was Liang-Bi Chen.

height of the two videos align at the same time point, and variations in student height, weight, or build can significantly distort the calculation results due to coordinate position transformations. As a result, the Euclidean distance method is gradually being supplanted by alternative approaches. In another investigation [10], interpolating wavelet was utilized to extract key frames from reference actions, and the Dynamic Time Warping (DTW) algorithm was employed to match reference and contrast actions. The DTW method has been widely applied in various fields, including gesture recognition, information retrieval, and language recognition [11], [12], [13], [14], [15], [16]. Subsequently, the average distance between matched keyframes was normalized to derive the similarity of the two action sequences. Although DTW-based similarity calculation effectively assesses the similarity between time series, it offers only a general indication of action similarity and lacks the capacity to provide detailed action-specific recommendations.

This study addresses the aforementioned challenges by introducing a normative analysis method for standing long jump. A deep learning-based pose estimation algorithm is employed to identify the key points of the long jump motion. By comparing the similarity between each frame's action and the desired target action, four critical actions are extracted. Subsequently, a comprehensive normative analysis method is devised to assess the long jump action and offer suggestions for improvement. This approach holds promise in enhancing the training and performance evaluation of standing long jump, contributing to the advancement of athletic techniques and sports sciences.

II. HUMAN POSE ESTIMATION

Owing to the limitations of conventional motion recognition techniques, this study adopts a deep learning-based human body pose estimation algorithm to effectively capture motion within video streams. In recent years, intelligent human posture estimation founded on vision has emerged as a highly challenging domain within the realm of computer vision. This approach enables the recognition of human behavior within video sequences by detecting human actions and subsequently extracting and learning relevant action features. Notably, numerous human pose estimation methods have been proposed in recent times [17], [18], [19], [20], [21], which are mainly classified into two categories: top-down methods, such as CPM [22] and AlphaPose [23], and bottom-up methods, such as OpenPose [24] and PersonLab [25].

The detection of human key points plays a critical role in describing human pose and predicting human behavior. However, traditional human detection methods encounter two major issues: imprecise localization of key points and sluggish model performance. To address these challenges effectively, this paper employs the following strategies.

1. This study utilizes a proprietary human pose estimation dataset, which includes 1500 standing long jump samples. Among these, 1362 samples were judiciously curated from the COCO (2017) dataset [26], and the remainder were collected and labeled by our research team. Furthermore, to enhance the dataset's diversity, some original images were mirrored before labeling. By improving the quantity and quality of annotated images, the problem of misplaced keypoints has been effectively minimized.

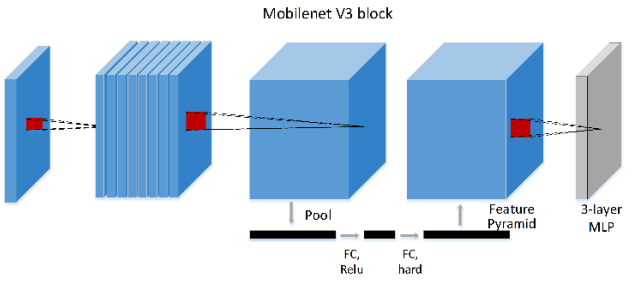


FIGURE 1. Model structure.

2. In model training, downsampling of the input images was implemented to reduce training complexity and resource consumption [26]. To facilitate thermal map-based training, the coordinates of bone key points in the original images were transformed to align with the resolution after downsampling. This involved converting the coordinates to the thermal map format, employing Gaussian fuzziness [27]. Once the thermal map prediction was completed, the resolution of the downscaled image was restored to its original pixel dimensions, and the predicted key point coordinates were mapped back to the original coordinate space. This approach ensures accurate localization and estimation of key points despite the initial downsampling process.

3. The lightweight human pose estimation model was trained using the innovative MoveNet algorithm, which employs heat maps to accurately identify key points on the human body. The proposed model structure of this study is shown in Figure 1.

Where the MobileNet-v3 is utilized as the backbone network [28], and features are extracted through a combination of the feature pyramid (FPN) technique [29] and a 3-layer Multi-Layer Perception (MLP), resulting in the generation of four essential heat maps: The **Center** of the human body, the set of all key points (**Keypoints**), the Offset of the key points (**Reg**), and the quantization error of the key points (**Offset**).

Following the generation of the heat map, a series of post-processing steps are executed to refine the key point positions. First, $2K$ values corresponding to the coordinate positions of the header_Reg channels are extracted, after which the center point coordinates are added to yield a rough key point position. Next, the header_Keypoints is divided by a weight matrix, and the coordinates of the maximum values are then computed through K channels, resulting in the final refined coordinates of 17 key points. The 17 key points of the human body are presented in Figure 2.

The key points encompass a comprehensive set of human body landmarks, arranged from top to bottom, namely: the left and right eyes, left and right ears, nose, left and right shoulder joints, left and right elbow joints, left and right

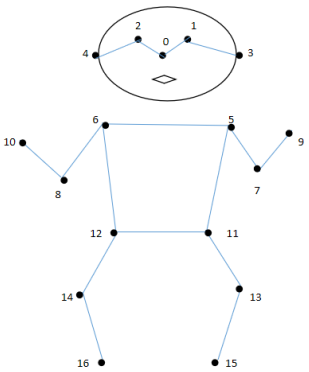


FIGURE 2. 17 key points of the human body.

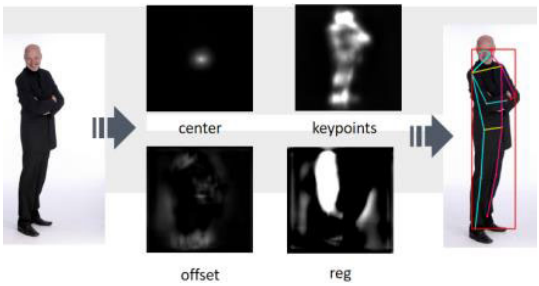


FIGURE 3. The flow chart of the human pose estimation.

hands, left and right hip joints, left and right knee joints, and left and right feet.

4. Different loss functions are used for each heatmap. The loss function incorporates weighted *MSE* and *L1* Loss, as shown in equations (1) and (2). **Keypoints** and **Center** use weighted *MSE* to balance positive and negative samples, while **Reg** and **Offset** use *L1* Loss. The weights assigned to each loss function are equally distributed.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \tag{1}$$

$$L1 = \sum_{i=1}^{n}|Y_i - f\ (x_i)| \tag{2}$$

The flow chart of the whole human pose estimation process is shown in Figure 3.

III. EXTRACT KEY ACTION

Sample video of the long jump practice consists of hundreds of frames of action. Analyzing every action not only increases the complexity of the system, but also reduces its robustness. Therefore, it is necessary to extract the key actions of the long jump before conducting normative analysis. Four key actions are identified in the long jump: take-off action, air action, landing action, and buffer action. A key point sequence is constructed for each of these four actions, and then compared frame by frame with the test video to select the key action.

After reviewing the previously published literature concerning normative recommendations for the long jump technique [30], [31], [32] and consulting with national-level sports educators, we have devised a set of standardized movement sequences specific to the standing long jump, as shown in Figure 4.

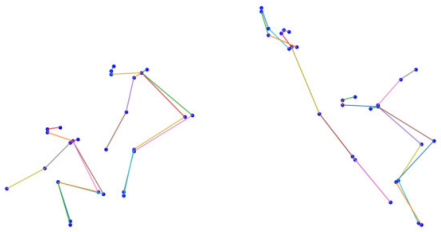


FIGURE 4. Standardized movement sequences of standing long jump.

The following sections describes the movement requirements for the four key elements of the standardized movement sequences:

Preparatory Action: The preparatory stance involves positioning the feet at shoulder-width apart, arms oscillating anteriorly and posteriorly in synchrony with the leg drive. During the arm swing, a simultaneous flexion of the knees occurs, lowering the center of gravity, and a slight forward inclination of the upper body is assumed. The hands are extended as far back as feasible while the legs apply rapid force to the ground. Typically, the knees of the takeoff leg undergo a flexion of approximately 90 degrees or more, serving the dual purpose of energy storage and elasticity provision. The alignment of the knees with the toes is meticulously maintained to preserve equilibrium.

Aerial Action: Rapidly propelling the body upward and forward is achieved through the forceful push-off of both feet's forefeet from the ground. Simultaneously, the arms undergo a substantial anterior and superior swing, in conjunction with the leg thrust. This phase entails a synchronized combination of thrust and swing to achieve an elevated trajectory while maximizing body extension. A linear body posture is maintained, with an approximate 45-degree angle between the legs and the ground.

Landing Action: Upon landing, a slight flexion at the hips assists in attenuating the transmission of impact forces to the upper body. Simultaneously, the arms are projected forward, contributing to balance, and a slight anterior inclination of the waist is observed.

Buffering Action: The knee joints undergo flexion as the feet make deliberate and steady contact with the ground. The landing is buffered through the coordinated use of both feet and knees, reducing the impact. The thigh-to-leg angle is approximately 30 degrees to mitigate impact. The upper body maintains a slight flexion of approximately 5 degrees at the shoulders and elbows, contributing to equilibrium preservation.

The key action matching is guided by the application of the following rules:

1. Sequence Alignment of Standard Actions: The matching process commences by aligning the sequence of standard actions. Specifically, the takeoff action is initially matched,

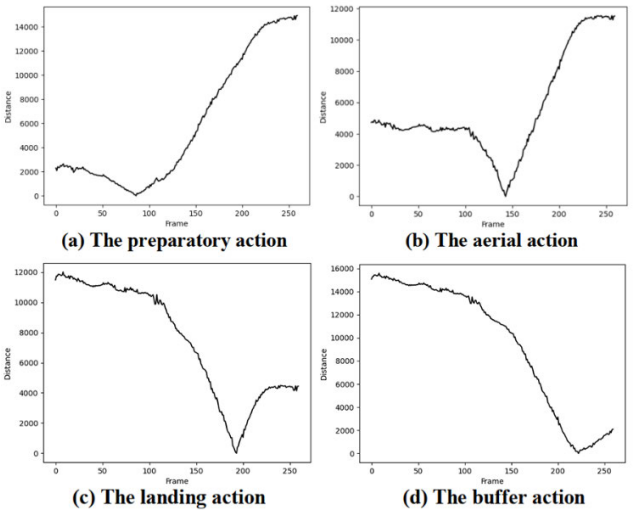


FIGURE 5. Key actions extraction.

and once successfully aligned during the air phase, any further comparison for the takeoff action and any preceding actions is dispensed with.

2. Prioritization of High Overall Similarity: The matching priority is given to achieving a high overall similarity between the test action and the standard action. To quantitatively assess this similarity, this study employs the Dynamic Time Warping (DTW) method [33], which is a nonlinear warping technique that combines distance measurement and time warping. It can effectively calculate the global similarity between action sequences.

3. Leg keypoint feature similarity prioritized matching: When the overall similarity between several test actions and standard actions is the same, the angle features of key leg points between actions are compared, and the actions with high angle feature similarity are matched as key actions. The calculation equation is as follows:

$$\Delta = \omega_1 \cdot x_1 + \omega_2 \cdot x_2 + \omega_3 \cdot x_3 \tag{3}$$

On the basis of the above rules, the DTW method is utilized to measure the similarity between the unequal length series of test video and key actions.

The equation for calculating DTW distance of the target point (x, y) is defined as:

$$DTW(x, y) = min\left(\frac{1}{K}\left(\sum_{i=1}^{K} W_i\right)^{\frac{1}{2}}\right) \tag{4}$$

where $K \in [max(x, y), x+y-1]$, W_i is the sequence distance value of each path, and the optimal path W must satisfy the following conditions:

Boundary: The starting point and ending point of W must be the starting point and ending point of the plane diagonal, namely $W_1 = (1, 1)$, $W_K = (x, y)$.

Continuity: For two adjacent points W_i (x_i, y_i) and W_{i-1} (x_{i-1}, y_{i-1}), where $x_i - x_{i-1} \leq 1$ and $y_i - y_{i-1} \leq 1$, the adjacent points are continuous.

Monotonicity: For two adjacent points W_i (x_i, y_i) and W_{i-1} (x_{i-1}, y_{i-1}), where $x_i - x_{i-1} \geq 0$, $y_i - y_{i-1} \geq 0$, and the elements on W cannot be backtracked.

The DTW distance between the action sequence and the standard action sequence are calculated for all frames. The actions with minimum distance from the following four curves are chosen as the corresponding key actions. Figure 5 presents the process of key actions extraction.

IV. ACTION NORMATIVE ANALYSIS

Upon the extraction of the key actions, a normative analysis employing the standard actions becomes imperative. Conventional action norms have been predominantly reliant on subjective judgments, leading to inefficiencies in the assessment process. Hence, there arises a pressing need to devise a quantitative and objective calculation methodology to evaluate action norms more effectively. In this study, we propose a novel approach that combines joint angle difference and feature index to ascertain the normalization of actions.

The devised method involves the segmentation of the test time series based on joint angle differences, followed by the utilization of DTW method to compute the distance disparity between the test sequence and the standard sequence. This computation allows us to determine the similarity between the tester's actions and the standard actions. To further heighten the precision of similarity calculation, a feature index is introduced in conjunction with joint angle calculation [34], [35].

To account for variations in human body types, a dynamic programming method is employed to calculate the motion similarity. The cosine similarity calculation method is used to determine the angles of each joint. Cosine similarity is a measure of the difference in direction between two vectors [36]. For two n-dimensional vectors $A = (a_1, a_2, \ldots, a_n)$ and $B = (b_1, b_2, \ldots, b_n)$, the cosine similarity ranges from -1 to 1. The cosine value is inversely proportional to the angle between the vectors. A value of -1 indicates that the vectors are in the opposite direction, while a value of 1 indicates that the vectors are in the same direction.

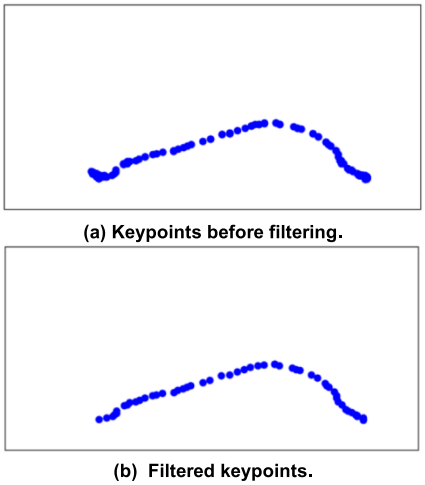
The equation for calculating cosine similarity $\cos(\theta)$ is as follows:

$$\cos(\theta) = \frac{\sum_{i=1}^{n} A_i * B_i}{\sqrt{\sum_{i=1}^{n} (A_i)^2} * \sqrt{\sum_{i=1}^{n} (B_i)^2}} \tag{5}$$

Based on the characteristics of the standing long jump event and in accordance with the priority matching rules, features with higher priority are assigned larger feature indices. Subsequently, the feature similarity is determined based on the feature indices, serving as the evaluation result. This approach allows for the calculation of the overall similarity between the key point vectors of the measured posture and the standard posture, thereby obtaining the final assessment [37].

$$C_i = (1 - \lambda)^w d + \lambda \left(|\cos(\alpha_i)| + 1\right)^q \tag{6}$$

In the given context, C_i represents the characteristic similarity of each joint, where a smaller value indicates a more



(a) Keypoints before filtering.

(b) Filtered keypoints.

FIGURE 6. Filtering keypoints.

TABLE 1. Comparison of different lightweight models.

Lightweight models	Accuracy (mAP)	Frame rate
Proposed model	68.3	10FPS
PoseNet [38]	52.2	3FPS
LightWeight OpenPose [39]	71.2	1FPS
CenterNet [40]	56.5	7FPS
CPM [41]	53.6	9FPS
MobileNet-v2 [42]	50.4	10FPS
AlphaPose[43]	72.7	1FPS
ShuffleNet-v2[44]	60.5	9FPS

standardized action. $\cos(\alpha_i)$ represents the cosine similarity of each joint. The value of d corresponds to the DTW value of two actions. Parameter λ lies between 0 and 1, denotes the weight ratio of the overall similarity to the cosine similarity of each joint. Experimental results have shown that a value of 0.73 yields the best effect. The variables w and q are characteristic indices. Due to variations in the activity angles of each joint, there can be significant differences in the obtained cosine similarity. Therefore, different values are assigned to each joint of w and q to ensure equal weighting of the scores.

To achieve the standardization of the overall long jump process, the trajectory of the key points of the foot is subjected to a filtering process. Subsequently, the integral value between the key points of the foot and the lowest horizontal line is calculated, followed by the assessment of the absolute difference from the standard action. Prior to conducting these calculations, it is imperative to filter the key points of the foot and exclude any non-long jump related key points from the process. In summary, this method allows for the establishment of a standardized long jump process by carefully analyzing the foot key point trajectories and accurately measuring their deviations from the standard action. The elimination of irrelevant key points ensures the accuracy and reliability of the assessment procedure. Figure 6 shows the filtering effect of keypoints.

Since each node is discrete, the parabola of the whole is not a continuous smooth curve, and the standard answer cannot be calculated directly through the existing equation. The approximate value can be obtained through the definition of Riemann integral, and the equation is as follows:

$$\int_b^a f(x)\,dx \approx \sum_{i=0}^{n-1} f(t_i)(x_{i+1} - x_i) \tag{7}$$

The left side is the integral value of the continuous function, and the right side is the approximate integral value of the discrete function. Where $x_{i+1} - x_i$ is the difference between the horizontal coordinate values of adjacent points, and $f(t_i)$ is a function value within the interval between x_{i+1} and x_i. In order to unify standards, the maximum value within the interval is taken by default.

The final normative calculation of motion was performed using a step scale, and the characteristic cosine similarity of each joint was weighted with the integral difference of the foot trajectory. The equation is as follows:

$$S_i = \begin{cases} 5, C_i > k^3 \\ 10, k^2 < C_i < k^3 \\ 15, k < C_i < k^2 \\ 20, C_i < k \end{cases} \quad \left(k = \sqrt{p+q}\right) \tag{8}$$

$$\text{Score} = Sum(S_i) - \int_b^a |\Delta f(x)|\,dx \tag{9}$$

In order to unify the score, S_i is a cumulative score of all joints, and the highest score of each joint is 20. The score of each joint is accumulated and then the integral difference of the whole action is subtracted to obtain the score of the key action.

V. EXPERIMENTAL RESULTS AND ANALYSIS

A. EXPERIMENTAL RESULTS OF HUMAN POSE ESTIMATION

To substantiate the feasibility of deploying this system on intelligent edge devices, the proposed method is compared with other lightweight models on a Raspberry 4B (1.8GHz). The comparison result of different lightweight models is presented in Table 1.

Figure 7 presents the pose estimation results of different models using the same long jump sample.

The human body detection frame is determined by employing the maximum external matrix encompassing the key points. For this specific long jump sample, the CPM, CenterNet and MobileNet-v2 showed a large deviation in the prediction of key points. Although the effect of LightWeight OpenPose is still good, the device lags seriously during the operation; PoseNet showed a situation in which the key points

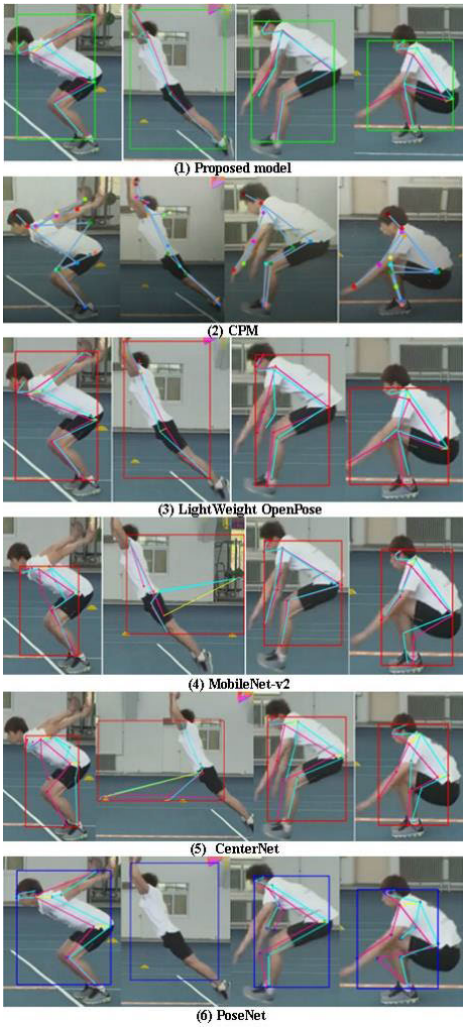


FIGURE 7. Visualization of pose estimation.

TABLE 2. Results of normative analysis and subjective judgment of three physical education teachers.

The name of the key actions	Normative analysis method		Subjective judgment of three physical education teachers	
	Correct number	Accuracy rate	Average correct number	Accuracy rate
Preparatory action	42	91.3%	36.7	79.8%
Aerial action	44	95.6%	39.3	85.4%
Landing action	38	82.6%	34.7	75.4%
Buffering action	44	95.6%	38.3	83.3%

TABLE 3. Results of normative analysis of long jump action.

The name of the key actions	Standard action	Test action	Suggestion	Score
Preparatory action			The forearm is a little lower, a little higher is recommended	106.52
Aerial action			The upper body is too high. Lower is recommended	105.43
Landing action			The upper forearms are too low, and elevation is recommended	97.88
Buffering action			Forearms are too high. Lower is recommended	94.84

of the vacating action were not recognized. By comparing the recognition effect of each model on the key points of the long jump action, it can be seen that this experimental model has the best effect for this specific sample.

B. EXPERIMENTAL RESULTS OF NORMATIVE ANALYSIS OF ACTIONS

Based on the action normative analysis method proposed in Section IV, 46 groups of standing long jump actions were tested and analyzed, as shown in Table 2.

As a comparison, the subjective judgment of three physical education teachers are also presented. The criterion for correctness is based on the standardized movement sequences mentioned in Section III.

The average accuracy rate of the four key actions is 91.3% with the proposed method, which indicates that it is a reliable and unbiased method of assessing the standardization of the long jump movement.

With the proposed method, the normalization suggestions of a test long jump action are depicted in Table 3. Notably, the disparities between the standard action and test action are observed in the pronounced deviation between the actions of the tester's upper and lower extremities when compared to the established standard practices.

Encouragingly, the recommendations generated by the system exhibit a remarkable alignment with the subjective assessments rendered by human evaluators. This congruence underscores the system's potential utility as a valuable tool in assisting athletes and coaches in the refinement of long jump techniques, thereby fostering improvements in long jump performance.

VI. CONCLUSION

In this study, we present a normative evaluation method of long jump action based on human pose estimation and similarity measures. Leveraging joint feature points extracted from the image, the learner's posture is deduced, and a normative calculation approach is devised to assess the variance between the learner's posture and the standard posture, consequently determining the adherence to standard actions. The

method facilitates the restoration of standard actions through ladder similarity calculation, offering valuable feedback and correction suggestions to learners, facilitating adjustments in their actions and poses.

Experimental findings showcase the effectiveness of the proposed method in analyzing the standard of long jump actions, thereby exhibiting its practical significance in contexts such as physical testing training in primary and middle schools and the correction of long jump techniques for athletes. By introducing an objective evaluation index to the long jump domain, our approach contributes to a more systematic and reliable means of assessing and improving athletic performance, thus holding promise for broader applications in various training and coaching settings.

REFERENCES

[1] A. V. Dowling, J. Favre, and T. P. Andriacchi, "Inertial sensor-based feedback can reduce key risk metrics for anterior cruciate ligament injury during jump landings," *Amer. J. Sports Med.*, vol. 40, no. 5, pp. 1075–1083, May 2012.

[2] J. Pansiot, B. Lo, and G.-Z. Yang, "Swimming stroke kinematic analysis with BSN," in *Proc. Int. Conf. Body Sensor Netw.*, Jun. 2010, pp. 153–158.

[3] P. Wang, C. F. Yuan, W. M. Hu, B. Li, and Y. N. Zhang, "Graph based skeleton motion representation and similarity measurement for action recognition," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 370–385.

[4] G. Wei and Y. Wei, "Similarity measures of Pythagorean fuzzy sets based on the cosine function and their applications," *Int. J. Intell. Syst.*, vol. 33, no. 3, pp. 634–652, Mar. 2018.

[5] G. Wei, J. Wang, M. Lu, J. Wu, and C. Wei, "Similarity measures of spherical fuzzy sets based on cosine function and their applications," *IEEE Access*, vol. 7, pp. 159069–159080, 2019.

[6] B. Fernando and S. Herath, "Anticipating human actions by correlating past with the future with Jaccard similarity measures," in *Proc. Comput. Vis. Pattern Recognit.*, Jun. 2021, pp. 13224–13233.

[7] R. Li, X. Wang, Y. Liu, and S. Zhang, "Improved technology similarity measurement in the medical field based on subject-action-object semantic structure: A case study of Alzheimer's disease," *IEEE Trans. Eng. Manag.*, vol. 70, no. 1, pp. 280–293, Jan. 2023.

[8] X. Jiang, F. Zhong, Q. Peng, and X. Qin, "Action recognition based on global optimal similarity measuring," *Multimedia Tools Appl.*, vol. 75, no. 18, pp. 11019–11036, Sep. 2016.

[9] J. Ying, "Research on sports assisted training based on Kinect," *Autom. Technol. Appl.*, vol. 38, no. 9, pp. 151–153, 2019.

[10] Y. J. Hua, W. Qing, and C. Hong, "Somatosensory dance interaction system based on motion evaluation algorithm," in *Proc. 3rd Int. Conf. Mechatronics Inf. Technol.*, 2018, pp. 16–28.

[11] S. Cai, Z. Lu, B. Chen, L. Guo, Z. Qing, and L. Yao, "Dynamic gesture recognition of A-mode ultrasonic based on the DTW algorithm," *IEEE Sensors J.*, vol. 22, no. 18, pp. 17924–17931, Sep. 2022.

[12] X. Li, "Human–robot interaction based on gesture and movement recognition," *Signal Process., Image Commun.*, vol. 81, Feb. 2020, Art. no. 115686.

[13] Y. Kuang, H. Cheng, Y. Zheng, F. Cui, and R. Huang, "One-shot gesture recognition with attention-based DTW for human–robot collaboration," *Assem. Autom.*, vol. 40, no. 1, pp. 40–47, Aug. 2019.

[14] M. Masnad, G. M. MukitHasan, K. M. Iftekhar, and M. S. Rahman, "Human activity recognition using DTW algorithm," in *Proc. IEEE Region 10 Symp. (TENSYMP)*, Jun. 2019, pp. 39–43.

[15] J. Zhao and L. Itti, "ShapeDTW: Shape dynamic time warping," *Pattern Recognit.*, vol. 74, pp. 171–184, Feb. 2018.

[16] M. Afrasiabi, H. Khotanlou, and M. Mansoorizadeh, "DTW-CNN: Time series-based human interaction prediction in videos using CNN-extracted features," *Vis. Comput.*, vol. 36, no. 6, pp. 1127–1139, Jun. 2020.

[17] D. Maji, S. Nagori, M. Mathew, and D. Poddar, "YOLO-pose: Enhancing YOLO for multi person pose estimation using object keypoint similarity loss," in *Proc. Comput. Vis. Pattern Recognit.*, 2022, pp. 2637–2646.

[18] S. Li, H. Zhang, H. Ma, J. Feng, and M. Jiang, "CSIT: Channel spatial integrated transformer for human pose estimation," *IET Image Process.*, vol. 17, no. 10, pp. 3002–3011, Aug. 2023.

[19] Y. Li, R. Liu, X. Wang, and R. Wang, "Human pose estimation based on lightweight basicblock," *Mach. Vis. Appl.*, vol. 34, no. 1, p. 3, Jan. 2023.

[20] Y. Zeng, H. Lv, M. Jiang, J. Zhang, L. Xia, Y. Wang, and Z. Wang, "Deep arrhythmia classification based on SENet and lightweight context transform," *Math. Biosci. Eng.*, vol. 20, no. 1, pp. 1–17, 2022.

[21] T. Chen, C. Fang, X. Shen, Y. Zhu, Z. Chen, and J. Luo, "Anatomy-aware 3D human pose estimation with bone-based pose decomposition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 1, pp. 198–209, Jan. 2022.

[22] S.-E. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, "Convolutional pose machines," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 4724–4732.

[23] H.-S. Fang, S. Xie, Y.-W. Tai, and C. Lu, "RMPE: Regional multi-person pose estimation," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2353–2362.

[24] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime multi-person 2D pose estimation using part affinity fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 1, pp. 172–186, Jan. 2021.

[25] G. Papandreou, T. Zhu, L. C. Chen, S. Gidaris, J. Tompson, and K. Murphy, "PersonLab: Person pose estimation and instance segmentation with a bottom-up, part-based, geometric embedding model," in *Proc. Comput. Vis. Pattern Recognit.*, vol. 11218, 2018, pp. 282–299.

[26] M. X. Tan, R. M. Pang, and V. L. Quoc, "EfficientDet: Scalable and efficient object detection," in *Proc. Comput. Vis. Pattern Recognit.*, 2020, pp. 10778–10787.

[27] Y. Deng, J. X. Luo, and F. L. Jin, "Overview of human pose estimation methods based on deep learning," *Comput. Eng. Appl.*, vol. 55, no. 19, pp. 22–42, 2019.

[28] A. Howard, M. Sandler, B. Chen, W. Wang, L.-C. Chen, M. Tan, G. Chu, V. Vasudevan, Y. Zhu, R. Pang, H. Adam, and Q. Le, "Searching for MobileNetV3," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 1314–1324.

[29] A. Kirillov, R. B. Girshick, K. M. He, and P. Dollàr, "Panoptic feature pyramid networks," in *Proc. Comput. Vis. Pattern Recognit.*, 2019, pp. 6399–6408.

[30] K. Mackala, J. Stodółka, A. Siemienski, and M. Ćoh, "Biomechanical analysis of standing long jump from varying starting positions," *J. Strength Conditioning Res.*, vol. 27, no. 10, pp. 2674–2684, 2013.

[31] Y. M. Li, "Review and research on the detailed teaching of standing long jump in the senior high school entrance examination," *Contemp. Sports Sci. Technol.*, vol. 7, no. 33, pp. 242–246, 2017.

[32] X. P. Zhang, "Research on teaching and training of standing long jump in junior high school," *Sports*, vol. 3, no. 2, pp. 129–130, 2015.

[33] W. M. Lu, "Discuss on correction and training guidance of key movements of standing long jump in junior middle school physical education," *Way Success*, vol. 12, pp. 92–94, 2022. [Online]. Available: https://xueshu.baidu.com/usercenter/journal/baseinfo?cmd=journal_page&entity_id=db7b4dd2c91fa349e731c33f5ddcbfdd

[34] A. Switonski, H. Josinski, and K. Wojciechowski, "Dynamic time warping in classification and selection of motion capture data," *Multidimensional Syst. Signal Process.*, vol. 30, no. 3, pp. 1437–1468, Jul. 2019.

[35] Y. Dang, J. Yin, and S. Zhang, "Relation-based associative joint location for human pose estimation in videos," *IEEE Trans. Image Process.*, vol. 31, pp. 3973–3986, 2022.

[36] B. Hassan, S. E. Abdelrahman, R. Bahgat, and I. Farag, "UESTS: An unsupervised ensemble semantic textual similarity method," *IEEE Access*, vol. 7, pp. 85462–85482, 2019.

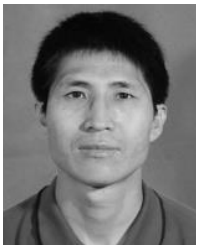
[37] Y. E. Song and X. Q. Wen, "TaiChi video registration method based on joint angle and DTW," *Comput. Technol. Autom.*, vol. 39, no. 1, pp. 117–122, 2020.

[38] J. Zhao, F. Hu, Y. Xu, W. Zuo, J. Zhong, and H. Li, "Structure-PoseNet for identification of dense dynamic displacement and three-dimensional poses of structures using a monocular camera," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 37, no. 6, pp. 704–725, May 2022.

[39] D. Osokin, "Real-time 2D multi-person pose estimation on CPU: Lightweight OpenPose," in *Proc. 8th Int. Conf. Pattern Recognit. Appl. Methods*, 2019, pp. 744–748.

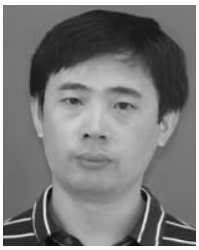
[40] K. Duan, S. Bai, L. Xie, H. Qi, Q. Huang, and Q. Tian, "CenterNet: Keypoint triplets for object detection," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 6568–6577.

[41] H. Wang, Q. J. Yang, Q. Wang, and L. Yu, "Lightweight gesture pose estimation based on CPM algorithm," in *Proc. Int. Conf. Modern Manage. Big Data (MMBD)*, 2022, pp. 361–369.
[42] H. Sun, S. Zhang, R. Ren, and L. Su, "Maturity classification of 'Huping-zao' jujubes with an imbalanced dataset based on improved MobileNet V2," *Agriculture*, vol. 12, no. 9, p. 1305, Aug. 2022.
[43] S. Mehdizadeh, H. Nabavi, A. Sabo, T. Arora, A. Iaboni, and B. Taati, "Concurrent validity of human pose tracking in video for measuring gait parameters in older adults: A preliminary analysis with multiple trackers, viewing angles, and walking directions," *J. Neuroeng. Rehabil.*, vol. 18, no. 1, p. 139, Dec. 2021.
[44] N. N. Ma, X. Y. Zhang, H. T. Zheng, and J. Sun, "ShuffleNet V2: Practical guidelines for efficient CNN architecture design," in *Proc. Eur. Conf. Comput. Vis.*, vol. 11218, 2018, pp. 122–138.



XIUGANG GONG is currently an Associate Professor with the Shandong University of Technology and the Executive Director of Education Branch of Shandong Internet of Things Association. His primary research interests include the realms of Internet of Things (IoT) technology and its applications, computer measurement and control, and embedded applications.

XINYUAN GENG is currently pursuing the Graduate degree with the Shandong University of Technology. His research interests include the Internet of Things (IoT) and computer vision.



GUANGJUN NIE received the Graduate degree from the Jining Medical College, in 2006, and the master's degree in neurology from the Bethune Medical Department, Jilin University, in July 2009. His research interests include the diagnosis and treatment of cerebrovascular disease and epilepsy.

TAO WANG received the Ph.D. degree in control science and engineering. His academic pursuits primarily revolve around the fields of computer software and computer applications.

JIAJUN ZHANG is currently pursuing the Graduate degree with the Shandong University of Technology. His research interests include the Internet of Things (IoT) and computer vision.

JIEBING YOU received the master's degree in clinical medicine science from Shandong University, China, in 2015. She is currently with Zibo Central Hospital. Her research interests include cerebrovascular disease, functional movement disorders, and peripheral neuropathy.