ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Motion Correction and Training Effect Improvement System for sports Players driven by Digital Technology

Wenxian Fan[1,*] *and Jianfei Lu*[2]

[1] School of Physical Education, Yuzhang Normal College, Nanchang, 330103, Jiangxi, China
[2] School of Primary Education, Yuzhang Normal College, Nanchang, 330103, Jiangxi, China

SUMMARY: *This paper constructs a motion correction and training effect improvement system for sports athletes driven by digital technology. High-speed video acquisition, inertial sensing, plantar pressure and heart rate load data are used to synchronize modeling, and a multi-source representation for special training is formed. At the recognition end, the system introduced the key point timing correlation and deviation quantification mechanism to distinguish the joint trajectory, the center of gravity transfer, the force rhythm and the action stability. At the control end, the hierarchical feedback of training phase, load status and completion quality output was combined to realize the linkage update of correction tips, rhythm correction and intensity adjustment. Experimental results based on 5504 training samples and 200 athletes show that the deviation detection rate of the system in six types of training movements is above 95.9%, the correction accuracy is between 95.4%-97.0%, and the average response delay is maintained at about 78 ms. After 6 weeks of continuous training, the improvement rate of movement completion in the system assisted training group reached 11.7%-13.4%, and the average score of special training increased from 82.4 to 91.3. In the continuous training scenario, the success rate of the system is kept between 94.8% and 97.2%, which reflects good field adaptability and real-time operation performance.*
Povzetek: In the experiment of 5504 groups of training samples and 200 athletes, the deviation detection rate of more than 95.9% was achieved, the average response delay was maintained at about 78 ms, and the movement completion improvement rate reached 11.7%-13.4% after 6 weeks of training. It can provide stable technical support for improving the training effect of athletes.

KEYWORDS: *Sports players; Motion correction; Computer vision; Improved training effect*

1 Introduction

With the introduction of visual perception, wearable acquisition, edge computing and intelligent decision-making technologies into sports training scenarios, action correction systems have shifted from a single video review tool to a digital platform that integrates state perception, deviation recognition, feedback generation and effect evaluation. Competitive training does not only look at whether the movement is completed, but also pays attention to whether the force chain is coherent, whether the joint Angle is reasonable, whether the torso control is stable, and whether the rhythm meets the special requirements. Manual observation can provide empirical judgment, but it is difficult to stably characterize fine-grained deviations in movements, and it is also difficult to analyze movement changes, load response and training effectiveness in the

*fanwei7211@126.com
https://doi.org/10.65102/is2026080
Published: 30 April, 2026

same framework. When digital technology is introduced into the training field, image sequences, skeleton trajectories, inertial data and physiological responses can be recorded synchronously, thus providing a basis for structured analysis of the training process. Around this trend, the action correction system begins to show the characteristics of the collaborative development of multi-source perception, online reasoning and closed-loop feedback.

In the aspect of digital acquisition of training actions, the existing research has provided a relatively complete acquisition basis for the construction of the system. Wang et al. studied wearable feedback mechanism in hammerball training and proposed a biomechanical feedback system [1]. Suo et al. studied sports motion capture technology and proposed a collection link classification framework [2]. Xi et al. studied human pose estimation methods in motion training and proposed a spatio-temporal Transformer enhanced model [3]. Fukushima et al. studied the usability of pose estimation in sports motion capture and proposed a scene verification path [4]. Mercadal-Baudart et al. studied the markerless 3D pose quantification method under the condition of a single camera, and proposed the motion measurement process [5]. D 'Haene et al. studied a 3D markerless capture tool combining multi-pose and depth estimation, and proposed a quantitative analysis and verification scheme [6]. Dill et al. studied the 3D pose reconstruction algorithm based on binocular camera information fusion, and proposed a reconstruction accuracy evaluation method [7]. Aleksic et al. studied the automatic vertical jump analysis process, and proposed a comparison verification scheme between unlabeled and labeled systems [8]. This kind of research shows that the training action can be transformed from two-dimensional observation to three-dimensional reconstruction and continuous quantification, and the action correction system has a more refined state input basis.

In terms of motion monitoring and multimodal fusion, the focus is starting to shift from static recognition to continuous measurements under real training conditions. Ino et al. studied the joint measurement method based on OpenPose and proposed the jump test analysis process [9]. Cronin et al. studied the feasibility of markerless action analysis in real track and field events, and proposed an application verification idea [10]. Van Hooren et al. studied the accuracy of visual motion capture in measuring running kinematic parameters and proposed a markerless measurement evaluation framework [11]. Barzyk et al. studied the method of markerless capture of lower limb joint kinematics by smartphone, and proposed a portable training measurement scheme [12]. White et al. studied wearable sensor feature extraction in sports performance prediction and proposed a parallel modeling idea [13]. Kovoor et al. studied the automatic analysis mechanism of sensor-enhanced wearable devices in injury prevention, and proposed the linkage mode of monitoring and risk identification [14]. Zhao et al. studied the application of smart sensors in sports science and proposed a training monitoring deployment framework [15]. These studies show that the source of data in the training field is no longer limited to video, inertia, angular velocity, pressure and physiological parameters jointly enter the analysis link, and multimodal perception has gradually become an important basis for sports training systems.

In terms of action recognition modeling with training feedback, research further turns to joint representation and online decision making. Rehman et al. studied the multimodal joint analysis of RGB images, skeleton tracking and pose estimation, and proposed a comprehensive activity recognition scheme [16]. Mekruksavanich et al. studied the attention-enhanced motion recognition method and proposed a deep classification model [17]. Gilmore et al. studied the fusion recognition algorithm of physiological signals and inertial signals, and proposed a cross-modal joint representation path [18]. These results show that action recognition has gradually shifted from single-modal classification to cross-modal coupling modeling. However, for the action correction task in special training, the recognition results alone are not enough to support

the improvement of training performance. The system also needs to connect the deviation location, feedback timing, regulation strength and stage performance evaluation to form a continuous update mechanism oriented to the training process.

Based on the above research basis, this paper constructs a motion correction and training effect improvement system for sports athletes driven by digital technology, which integrates training action data acquisition, action deviation detection and correction, feedback regulation for training effect improvement, and multimodal real-time linkage mechanism into a unified link. The system does not regard action recognition as the end point, but continues to map the recognition results to corrective prompts, load adjustment and effect evaluation, so that the change of action quality and the change of training gain can be continuously recorded and dynamically analyzed under the same framework.

2 System core design

2.1 Data collection of athletes' training movements

The data acquisition of athlete training action determines the input quality of the whole system. The action correction task does not only rely on single frame image recognition results, but also involves continuous changes in joint trajectories, local force timing, body center of gravity transfer and training load state. Therefore, the acquisition end should not only save the video picture, but unify the visual information, inertial information, plantar force information and physiological response information on the same time axis to form a multi-modal sample that can be used for subsequent deviation detection and feedback control [19]. Aiming at this goal, this paper constructs a data acquisition link composed of dual-camera video, depth perception, inertial measurement and pressure sampling, and completes the unified deployment in the training hall scene.

The video part is captured synchronously by two 4K high-speed cameras and a depth camera. The main camera was located 4 m directly in front of the athlete to record the action frontal texture, torso orientation, and upper and lower limb expansion state. The auxiliary camera was located at 3.5 m in the direction of 45° in oblique front, which was used to supplement the lateral action, turning action and occlusion area. The depth camera is fixed in front of the action area to supplement the relationship between the human body and the spatial position. In order to ensure contour integrity under high-speed actions, the video sampling rate is set to 120 fps and the shutter time is controlled within 1/1000 s. The depth flow resolution is set to 640×480 and the sampling rate is 60 fps. Inertial data were collected by six nine-axis IMU modules fixed on both forearms, both calves, thoracoback, and lumbosacral regions with a sampling frequency of 200 Hz. The plantar pressure pads were placed in the action starting and falling areas, and the sampling frequency was 100 Hz. The heart rate band records the heart rate changes during training and is used to describe the correspondence between action execution and physical load.

Table 1 presents the acquisition hardware and environmental parameter configuration.

Table 1: Acquisition hardware and environmental parameters.

Acquisition Item	Parameter Configuration	Functional Description
Main video stream	4K, 120 fps	Records the main-view texture and motion trajectory
Auxiliary video stream	4K, 120 fps	Complements lateral movements and occluded regions
Depth stream	640 × 480, 60 fps	Provides spatial distance and positional hierarchy
IMU modules	6 units, 200 Hz	Captures angular velocity, acceleration, and posture changes
Plantar pressure	100 Hz	Describes support timing and center-of-gravity transfer
Heart rate acquisition	1 Hz	Characterizes training load and recovery state
Illumination range	350–1200 lx	Ensures clear contours and stable brightness

The collection objects cover six training units: starting acceleration, stride changing direction, vertical jump landing, upper limb swing, throwing force and special comprehensive action. Each athlete completed the action collection in the standard state, the intensified state and the fatigue state, so that the sample included the action completion quality change and the training load change at the same time. The background of the field is controlled as a blue-gray low-texture screen, and the ground marking line is kept uniform to avoid the model misjudging the background difference as the action difference. After all raw data enter the buffer, the system first completes time synchronization, and then performs human detection, skeleton extraction, depth alignment, and signal sectioning. Figure 1 shows the data acquisition and synchronous processing flow.

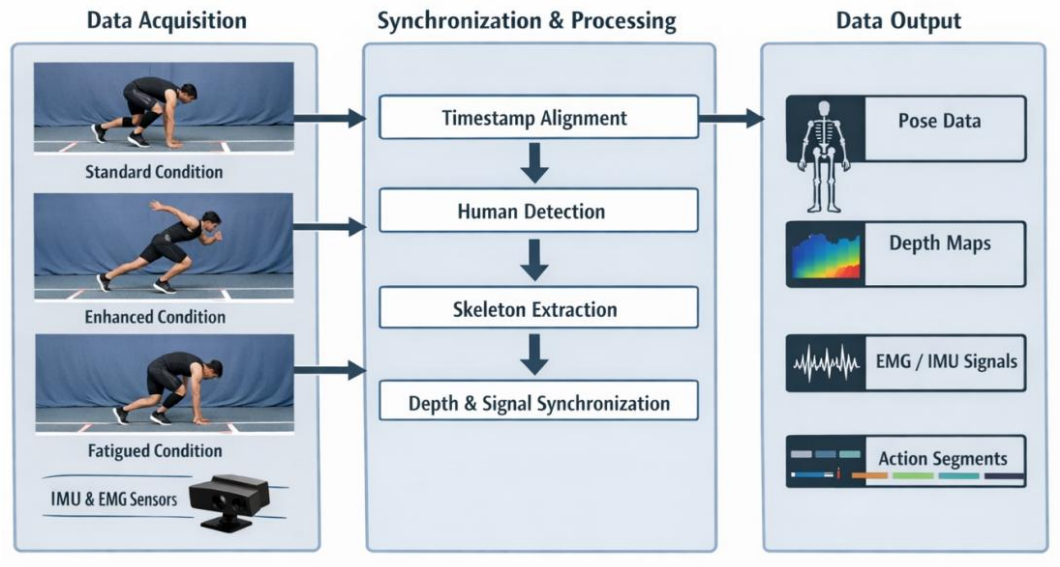

Figure 1: Data collection and synchronization process of athletes' training movements.

In order to make multi-source data correspond to the same action segment, the system uses the principle of minimum time difference to complete the temporal alignment. The synchronization error is defined as follows:

$$d_t = \min|t_v - t_s| \tag{1}$$

Here, t_v represents the video frame timestamp, t_s represents the sensor sampling timestamp, and d_t represents the minimum synchronization error between video and sensor. This formula is used to constrain the time offset when the data of each modality enters the unified fragment.

After keypoint extraction, the position of the same joint will drift slightly under different views. In order to reduce the impact of occlusion and jitter, the system uses confidence weighted fusion of key point coordinates:

$$p_t = \frac{\sum_{i=1}^{n} w_i \, p_i}{\sum_{i=1}^{n} w_i} \tag{2}$$

Here, p_i represents the keypoint coordinates in the i-th observation view, w_i represents the confidence weight of this observation, and p_t represents the position of the target keypoint after fusion. This formula is used to generate more stable skeleton inputs.

Video frames need to be standardized before entering the subsequent network to compress the distribution differences under different lighting conditions. The standardization process is expressed as follows:

$$x^* = \frac{x - \mu}{\sigma} \tag{3}$$

Here, x represents the original pixel value, μ represents the mean value of the corresponding channel, σ represents the standard deviation of the corresponding channel, and x^* represents the normalized pixel value. This formula is used to improve the consistency of the input distribution and reduce the influence caused by brightness fluctuations.

In order to screen out missing frames, interrupted and serious drift samples, the system constructs a sample quality scoring function:

$$q = \alpha c + \beta s + \gamma m \tag{4}$$

Here, c represents the keypoint continuity rate, s represents the multi-modal synchronization completion rate, m represents the action segment integrity rate, α, β, γ represent the weight coefficients, and q represents the sample quality score. This formula is used to control the validity of the data that eventually enters the training set.

For sample collation, all action clips were uniformly cropped to 64 frames of video length, corresponding to 128 steps of IMU sequence and 64 steps of pressure sequence. The dataset was divided into training set and test set by 3:1, and stratified sampling was performed according to action category, training phase and load level before division to maintain the same sample proportion in different states. The final dataset distribution is shown in Table 2.

Table 2: Distribution of training action dataset.

Action Type	Training Set	Test Set	Total
Sprint start acceleration	720	240	960
Crossover direction change	684	228	912
Vertical jump landing	708	236	944
Upper-limb swing	666	222	888
Throwing force generation	648	216	864

Specialized composite actions	702	234	936

After the above processing, the original video, depth map, skeleton sequence, inertia sequence, pressure phase and heart rate load are organized as multimodal training samples with a unified structure, which provides a stable input basis for subsequent action deviation detection and training effect evaluation.

2.2 Motion deviation detection and correction model

The action deviation detection and correction model is located in the core layer of the system, which is responsible for action recognition, template comparison, deviation quantification and correction instruction generation. Training misalignment is not only manifested as a single joint abnormality, but also reflected by trajectory deviation, rhythm misalignment and support imbalance. To this end, the video stream, skeleton sequence, IMU sequence and pressure sequence are uniformly mapped into the action structure space, and the continuous modeling of "identification-alignment-location-correction" is completed in the same time window. The video stream represents the appearance and displacement of the action, the skeleton sequence represents the joint connection relationship, the IMU sequence complements the angular velocity change in the rotation and force generation stage, and the pressure sequence reflects the support phase and the center of gravity migration. Multi-source features are sent to the deviation detection branch after time sequence alignment, and the output can be directly used for training and correction of structured results [20].

To illustrate the processing link of this model, the relationship between multimodal input, deviation quantization, and corrected output is organized in Figure 2. Figure 2 shows the whole process of timing alignment, joint Angle parsing, trajectory alignment, rhythm detection and support evaluation in turn after the training action enters the model, and outputs the comprehensive deviation score and correction vector.

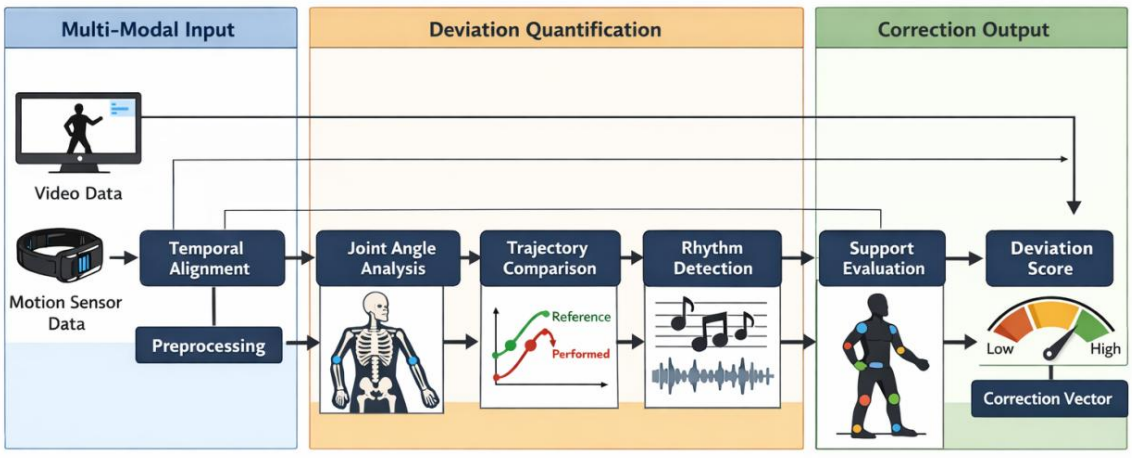

Figure 2: Model structure of action deviation detection and correction.

In order to reduce the influence of differences in height, segment length and sampling scale, the system first calculates key joint angles based on adjacent bone segment vectors. Let the bone segment vectors on both sides of the j-th joint be $u_{t,j}^{(1)}$ and $u_{t,j}^{(2)}$, respectively, then the joint Angle in frame t is expressed as follows:

$$\theta_{t,j} = \arccos\left(\frac{u_{t,j}^{(1)} \cdot u_{t,j}^{(2)}}{\|u_{t,j}^{(1)}\|_2 \|u_{t,j}^{(2)}\|_2 + \varepsilon}\right) \tag{5}$$

Here, $\theta_{t,j}$ represents the observation Angle of the j-th joint in frame t, and ε is a tiny constant to prevent the denominator from being zero. This formula transforms the skeleton coordinates into angular features with kinematic significance.

After obtaining the joint angles, the system aligns the observed action with the standard template, and the joint Angle deviation is defined as follows:

$$\delta_{t,j}^{ang} = \frac{\left|\theta_{t,j} - \theta_{\phi(t),j}^{*}\right|}{\theta_j^{max} - \theta_j^{min} + \varepsilon} \tag{6}$$

Here, $\theta_{\phi(t),j}^{*}$ denotes the template Angle after temporal alignment, and $\phi(t)$ denotes the template time index. This formula is used to describe joint level deviations such as insufficient flexion and extension, excessive swing and rotational hysteresis.

Only relying on Angle information is not enough to characterize the action path change, and the system further calculates the trajectory deviation of key points. Let the position of the j-th keypoint in frame t be $p_{t,j}$, and the template position be $p_{\phi(t),j}^{*}$. Then:

$$\delta_t^{trj} = \frac{1}{J\, l_{body}} \sum_{j=1}^{J} \left\| p_{t,j} - p_{\phi(t),j}^{*} \right\|_2 \tag{7}$$

Here, J is the total number of keypoints, and l_{body} is the human scale based on shoulder width and hip width. This formula is used to eliminate the influence of body size difference and reflect the overall trajectory deviation degree.

The training actions contain not only spatial structure but also distinct rhythmic features. The system establishes the event time set according to the velocity peak sequence, and the rhythm deviation is expressed as follows:

$$\delta^{tmp} = \frac{1}{K} \sum_{k=1}^{K} \frac{|\tau_k - \tau_k^{*}|}{T} \tag{8}$$

Here, τ_k denotes the observation time of the k-th action event, τ_k^{*} denotes the corresponding event time of the template, K denotes the number of events, and T denotes the total duration of the action segment. This formula is used to describe the timing dislocation in the process of starting, braking, taking off and landing.

In order to enhance the detection ability of lower limb stability and center of gravity control, the system introduces the imbalance degree of left and right foot support. If the left and right plantar pressures in frame t are F_t^L and F_t^R respectively, then:

$$\delta_t^{sup} = \frac{\left|F_t^L - F_t^R\right|}{F_t^L + F_t^R + \varepsilon} \tag{9}$$

Here, δ_t^{sup} denotes the support difference at that instant. This formula is used to reflect the landing imbalance, unilateral compensation, and center of gravity offset state.

After obtaining four types of deviations: Angle, trajectory, rhythm and support, the system constructs a comprehensive deviation scoring function:

$$S_{dev} = \frac{1}{T}\sum_{t=1}^{T}\left(\alpha\sum_{j=1}^{J}\omega_j\,\delta_{t,j}^{ang} + \beta\delta_t^{trj} + \gamma\delta^{tmp} + \eta\delta_t^{sup}\right) \tag{10}$$

Here, ω_j represents the joint importance weights, and α,β,γ and η represent the fusion coefficients of each bias term. This formula is used to unify the local error and the global error into the same evaluation scale.

Instead of a single score, the model outputs an executable correction vector. Let the correction instruction of the j-th key joint be c_j, then we have:

$$c_j = \lambda_1\left(\theta_{\phi(t),j}^{*} - \theta_{t,j}\right)n_j^{ang} + \lambda_2\left(p_{\phi(t),j}^{*} - p_{t,j}\right) + \lambda_3(\tau_k^{*} - \tau_k)n^{tmp} \tag{11}$$

Here, n_j^{ang} represents the joint correction direction unit vector, n^{tmp} represents the rhythm correction direction marker, and λ_1,λ_2, and λ_3 represent the three types of correction term weights. This formula is used to convert the deviation analysis results into corrective information that can be directly sent to the display and voice terminals.

In order to ensure that the model maintains consistency between action category determination and deviation regression, a joint objective function is used in the training stage:

$$L = \lambda_c L_{cls} + \lambda_d L_{dev} + \lambda_r \sum_{j=1}^{J}\left\|c_j - c_j^{*}\right\|_2^2 \tag{12}$$

Here, L_{cls} represents the action classification loss,L_{dev} represents the bias score loss, the last term represents the corrected vector regression loss, and λ_c,λ_d, and λ_r represent the corresponding weights. This formula enables the model to recognize the action category while maintaining the stability of the deviation localization and correction output.

During training, the video branch uses lightweight spatio-temporal convolution, the skeleton branch maintains topological consistency according to the joint connection matrix, the IMU branch uses one-dimensional convolution to compress the high-frequency jitter components, and the pressure branch emphasizes the force changes in the landing and commutation stages. The four branches share a unified time index in the fusion layer, the input segment length is set to 64 frames, the batch size is set to 32, the initial learning rate is set to 0.0003, the optimizer uses AdamW, and the learning rate is decreased by cosine annealing after 20 rounds. The system finally outputs the action category, deviation position, deviation degree and correction direction, which provide continuous input for the subsequent feedback module.

2.3 Feedback control module for training effect improvement

The feedback control module for training effect improvement adjusts prompt content, prompt intensity, rhythm amplitude and load level in real time according to action deviation, training load and stage completion quality [21]. The feedback should not only fall to the most obvious part of the deviation, but also maintain the continuity of the action execution. Based on this requirement, the system constructs a closed-loop regulation module, which encodes the

comprehensive deviation score, joint deviation distribution, rhythm misalignment, support stability, heart rate load and recent training improvement rate into a state vector, and drives the feedback decision network to output control parameters such as correction position, correction direction, rhythm adjustment, load increase or decrease, and recovery time. Figure 3 illustrates the complete closed loop of this module from action deviation input to reward update.

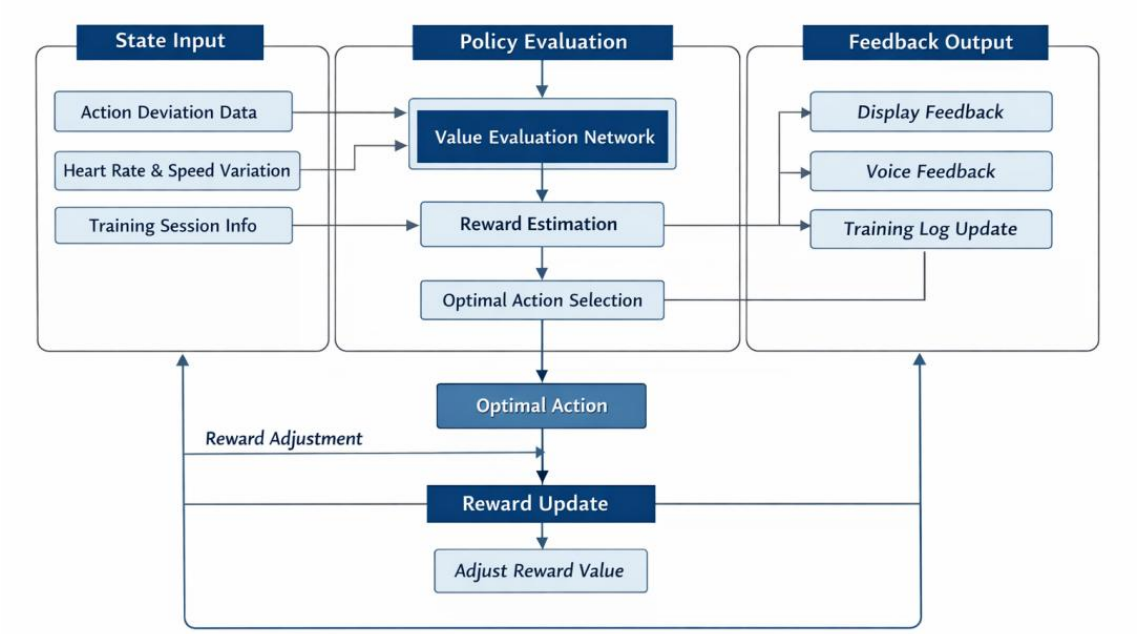

Figure 3: Structure of feedback regulation module for training effect improvement.

The module receives the action deviation detection results, and combines heart rate, speed fluctuation and training rounds to form the current state. The value evaluation network calculated the long-term benefit of each feedback action, and then sent the optimal action to the display end, the voice end and the training record end, and wrote back the reward value according to the improvement of the next round of action. This structure enables the feedback to be continuously updated with the training process.

In the process of feedback control, the system does not make judgments based on a single error term, but integrates action quality, rhythm state, support stability and physiological load into a unified decision space. The current training round state is defined as follows:

$$s_t = \left[S_{dev,t}, \bar{\delta}_t^{ang}, \delta_t^{tmp}, \delta_t^{sup}, \rho_t, \kappa_t\right] \tag{13}$$

Here, $S_{dev,t}$ represents the comprehensive deviation score, $\bar{\delta}_t^{ang}$ represents the average deviation of joint angles, δ_t^{tmp} represents the rhythm deviation, δ_t^{sup} represents the support imbalance, ρ_t represents the normalized heart rate load, and κ_t represents the improvement rate of the last three rounds of training. This formula is used to uniformly encode action quality and load state into feedback decision space.

To map abstract decisions to field executable instructions, a single-round feedback action is defined as follows:

$$a_t = [g_t, d_t, l_t, \tau_t, r_t] \tag{14}$$

Here, g_t represents the label of correction position, d_t represents the label of correction direction, l_t represents the level of load increase or decrease, τ_t represents the amplitude of

rhythm adjustment, and r_t represents the level of recovery duration. This equation is used to turn training feedback into an executable control.

To measure whether the feedback action actually improves the quality of the action, the system constructs an action quality reward function:

$$R_t^q = \tanh\left(\frac{S_{dev,t} - S_{dev,t+1}}{S_{dev,t} + \varepsilon}\right) + \lambda_1(\kappa_{t+1} - \kappa_t) \tag{15}$$

Here, R_t^q represents the quality improvement reward, ε is a tiny constant, and λ_1 is the improvement rate gain coefficient. This formula makes the system more inclined to choose the feedback strategy that can continuously reduce the deviation score and improve the training improvement rate.

In order to ensure the balance between training intensity and action stability, the system further introduces load and stability constraints:

$$R_t^l = \exp(-|\rho_{t+1} - \rho_t^*|) - \lambda_2 \max(0, \rho_{t+1} - \rho_{max}) - \lambda_3 \delta_{t+1}^{sup} \tag{16}$$

Here, ρ_t^* represents the target load level, ρ_{max} represents the safety threshold, and λ_2 and λ_3 are penalty weights. This formula is used to ensure that the feedback action maintains the effective training intensity and does not push the load to a state of imbalance.

After defining the quality improvement term and the load constraint term, the feedback control module also needs to incorporate the rhythm recovery effect and the smoothness of the adjacent rounds of action switching into the same evaluation framework. For this purpose, the single-round total reward function is written as a multinomial coupling form in this paper, so that the policy network not only focuses on the training gain, but also maintains the continuity and enforceability of the feedback output:

$$R_t = \omega_1 R_t^q + \omega_2 R_t^l + \omega_3 \exp\left(-\left|\delta_{t+1}^{tmp}\right|\right) - \omega_4 \|a_t - a_{t-1}\|_2 \tag{17}$$

Here, $\omega_1, \omega_2, \omega_3 3$ and ω_4 denote the weight coefficients. The third term is used to encourage rhythm recovery, and the last term is used to suppress drastic jumps in the two feedback actions before and after, thus maintaining the stability of the cue style and training rhythm.

In order to suppress the fluctuation of value estimation and enhance the policy stability under high noise conditions, this paper uses the dual network objective to construct the Q-value update term:

$$y_t = R_t + \gamma Q_{\tilde{\theta}}\left(s_{t+1}, \arg\max_a Q_\theta(s_{t+1}, a)\right) \tag{18}$$

Here, Q_θ denotes the online network, $Q_{\tilde{\theta}}$ denotes the target network, γ denotes the discount factor, and y_t denotes the target Q-value. This formula can reduce the overestimation bias in feedback action evaluation and improve the stability of online regulation.

In the training phase, the input dimension of the state encoder is 6, the value network contains a three-layer fully connected structure with hidden dimensions of 256, 128 and 64, respectively, and the output dimension corresponds to the feedback action combination space. The experience replay buffer capacity is set to 12000, the batch size is set to 64, the initial learning rate is set to 1×10−4, the discount factor is set to 0.95, and the target network is updated every 120 steps. The exploration phase adopts an ε greedy strategy with ε linearly decaying from 0.9 to 0.1. The system performs sliding statistics on the average reward of 200 consecutive

steps, and records the reward curve, load curve and deviation curve synchronously. All the training was completed on the RTX 3090 platform, and a single round of complete regulation training took about 11.6 hours. Through the above design, the feedback module can transform the deviation detection results into a dynamic, continuous and executable training control strategy, so that the action correction and training effect improvement can be linked in the same closed loop.

3 System integration and operation implementation

3.1 Multimodal real-time feedback linkage mechanism

The multimodal feedback linkage mechanism takes the role of connecting the recognition results with the training execution. A single video stream can describe the action form, but it is difficult to reflect the force sequence, support changes and load fluctuations. Inertia, pressure and heart rate data can complement dynamic information, but lack spatial posture semantics [22]. Therefore, the system connects visual features, deviation, IMU timing, plantar pressure and load indicators into the linkage module at the running layer, and synchronizes them with frame-level timestamps, so that correction prompts, rhythm adjustment and load feedback respond synchronously within the same round of action. In order to maintain the consistency of the feedback side, the weighted gated fusion expression is constructed as follows:

$$F_t = \sigma(W_1 v_t + W_2 k_t + W_3 m_t + W_4 p_t + b) \odot h_t \tag{19}$$

Here, v_t represents the visual state vector,k_t represents the skeleton deviation vector, m_t represents the inertia feature, p_t represents the pressure and load feature,h_t represents the training phase context, and σ is the gating function. This equation is used to suppress transient noise and highlight the most sensitive modal information to correction. After the fusion results enter the feedback control terminal, the system generates voice prompts, screen markers and rhythm control instructions according to the deviation position, deviation level and training intensity. At the same time, the execution results are written back to the log terminal for the next round of status update. Therefore, action recognition, deviation judgment and training intervention form a continuous closed loop, and the system can still maintain feedback coherence and execution stability under high-speed training conditions.

As shown in Figure 4, after the video stream, skeleton stream, IMU stream, pressure stream and heart rate stream enter the processing chain, time synchronization and state alignment are completed first, and then feedback trigger signals are generated by gated fusion.

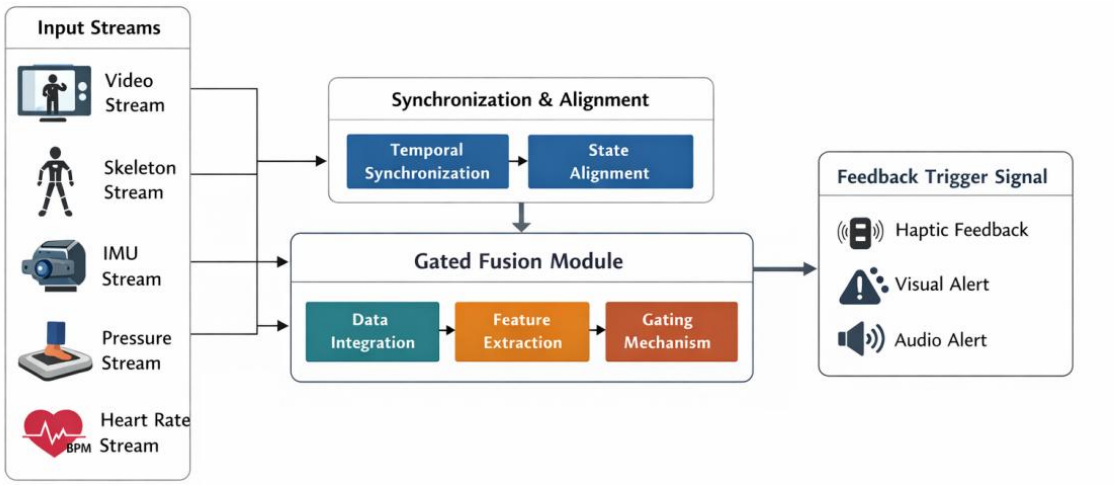

Figure 4: Multimodal real-time feedback linkage mechanism.

3.2 Experimental setup and evaluation index

In order to test the comprehensive performance of the system in action correction and training effect improvement, the experiment was evaluated from three levels: action recognition and correction performance, training gain effect and operation stability. In the recognition and correction layer, deviation detection rate, false trigger rate, correction accuracy, F1 value and average positioning error were used to measure the system's ability to identify motion deviation, the consistency of correction judgment and the positioning accuracy of deviation parts. In the training effect layer, the improvement rate of action completion, the improvement rate of rhythm consistency and the change of stage performance were used to reflect the changes of action quality, rhythm control and special performance after feedback regulation. The average response delay, delay standard deviation, triggering success rate and stability coefficient are used in the running layer to evaluate the real-time feedback ability and operation stability of the system under continuous training conditions.

The data set includes 5504 groups of training samples, and the subjects are 200 sports athletes. They are divided into regular training group and system assisted training group according to age and special basis. All experiments lasted for 6 weeks, with training sessions of 45 min four times a week. At the system end, a dual-camera RGB-D camera and a six-point IMU are used to synchronously collect action data. The video resolution is 1280×720, the frame rate is 60 fps, and the IMU sampling frequency is 200 Hz.

In the preprocessing stage, background suppression and normalization are performed on the image, and low-pass filtering and time alignment are performed on the inertia and pressure signals. Finally, the action completion degree, correction times and stage performance changes in the test stage are used as the comprehensive evaluation basis [23]. In order to reduce the influence of individual differences on the evaluation results, all indicators were calculated by using the baseline normalization method within the group. The training results were independently reviewed by two specialized coaches, and the average value was taken as the final record value. The subjective evaluation results and objective measurement indicators are synchronously included in the evaluation system for cross-validation.

4 System application results

4.1 Deviation detection and trigger stability analysis in action correction performance

The deviation detection rate versus false trigger rate for different training actions is shown in Figure 5. In this figure, the bias recognition effect and trigger stability of the system in six classes of training actions are compared. The deviation detection rate of special comprehensive action was the highest, reaching 97.6%, and the false trigger rate was the lowest, reaching 1.9%. Starting acceleration followed closely, with a deviation detection rate of 97.3% and a false trigger rate of 2.1%. This indicates that the system has a strong deviation analysis ability and a stable feedback trigger ability for the training task with a relatively complete action chain and a clear structure hierarchy. The deviation detection rate of stride and upper limb swing is 96.8%, the false trigger rate is 2.4%, and the deviation detection rate of upper limb swing is 96.5%, and the false trigger rate is 2.3%, which indicates that the system can still maintain good recognition consistency in the direction switching and local swing scenes.

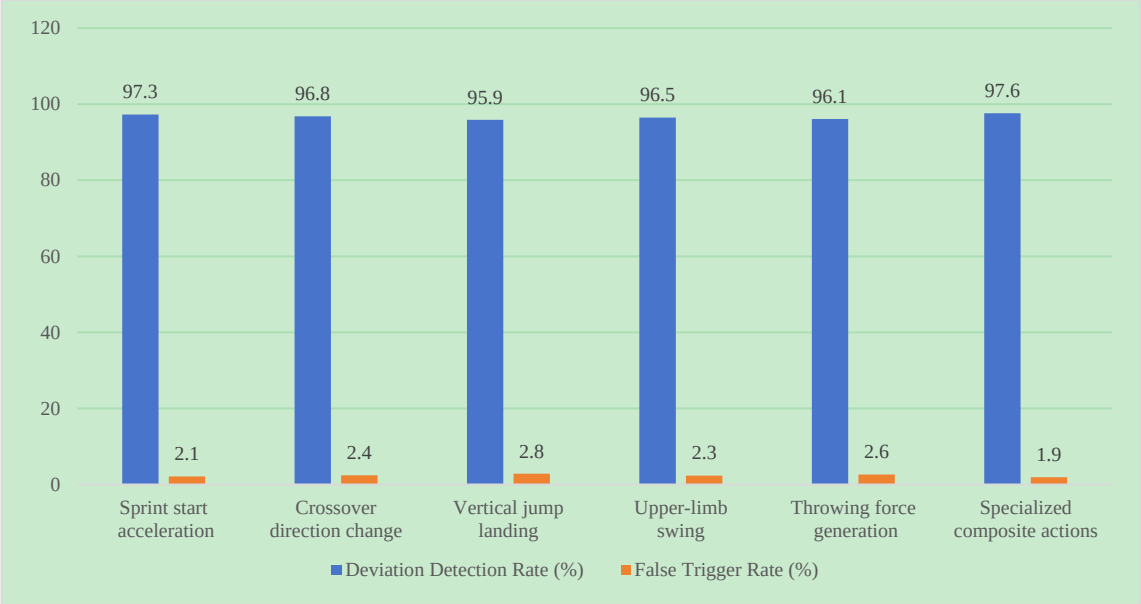

Figure 5: Comparison of deviation detection rate and false trigger rate for different training actions.

The deviation detection rate of vertical jump landing and throwing force was 95.9% and 96.1%, respectively, and the false trigger rate was 2.8% and 2.6%, which was slightly lower than that of other action types. This result shows that in the training phase when the characteristics of takeoff, landing and fast swing are more concentrated, the system's decision of deviation boundary and feedback trigger control are more dependent on the synchronization accuracy between multimodal information. Although the indicators of the two types of actions fluctuated slightly, the deviation detection rate of each action remained above 95.9%, and the false trigger rate was controlled within 2.8%, indicating that the system still had high deviation recognition ability and good trigger stability under different training action conditions.

Table 3 further presents the specific results of each action in terms of correction accuracy, F1 value, and average localization error. The system maintains high consistency in joint deviation location and correction direction judgment. The average positioning errors of the

starting acceleration and the special comprehensive action are controlled at 4.2° and 4.1°, respectively, which indicates that the system has a strong ability to capture the key deviation points in the continuous action. The positioning error of vertical jump landing is 4.9°, which is slightly higher than that of other actions, which is mainly related to the concentrated change of action and the faster conversion speed of support during the stage from takeoff to landing. However, the F1 value still reaches 0.949, indicating that the overall missed judgment and misjudgment are at a low level.

Table 3: Results of action correction performance analysis.

Action Type	Correction Accuracy (%)	F1 Score	Average Positioning Error (°)
Sprint start acceleration	96.8	0.965	4.2
Crossover direction change	96.1	0.958	4.6
Vertical jump landing	95.4	0.949	4.9
Upper-limb swing	95.9	0.955	4.7
Throwing force generation	95.7	0.952	4.8
Specialized composite actions	97.0	0.968	4.1

From the overall results, the system can not only locate the deviation part, but also maintain a stable correction output quality. The F1 value of each action is higher than 0.949, indicating that the model achieves a good balance between precision and recall. Combined with the training records, it can be seen that the average improvement rate of the next round of action after the system output prompts reaches 12.6%, and the improvement of lower limb control related actions is more obvious, indicating that there is a strong correspondence between the action correction results and the subsequent training improvement.

4.2 Analysis of training effect improvement

The training effect improvement analysis is shown in Figure 6. In order to verify the gain effect of the system in continuous training, this paper compared the conventional training group with the system assisted training group based on the results of six weeks of training. The statistical indicators include the improvement rate of action completion, the improvement rate of rhythm consistency and the change of stage performance, and the comprehensive judgment is carried out by combining the training log. The results showed that the overall performance of the system-assisted training group was significantly better than that of the conventional training group, including acceleration from 7.4% to 12.1%, stride change direction from 8.0% to 12.8%, vertical jump landing from 8.5% to 13.2%, and upper limb swing from 7.1% to 11.7%. Throw power increased from 7.8% to 12.4%, and SPECIAL COMBINED ACTION increased from 8.0% to 13.4%. This indicates that multimodal feedback can act on the continuous action chain faster and reduce the phenomenon of local imbalance spreading to subsequent actions.

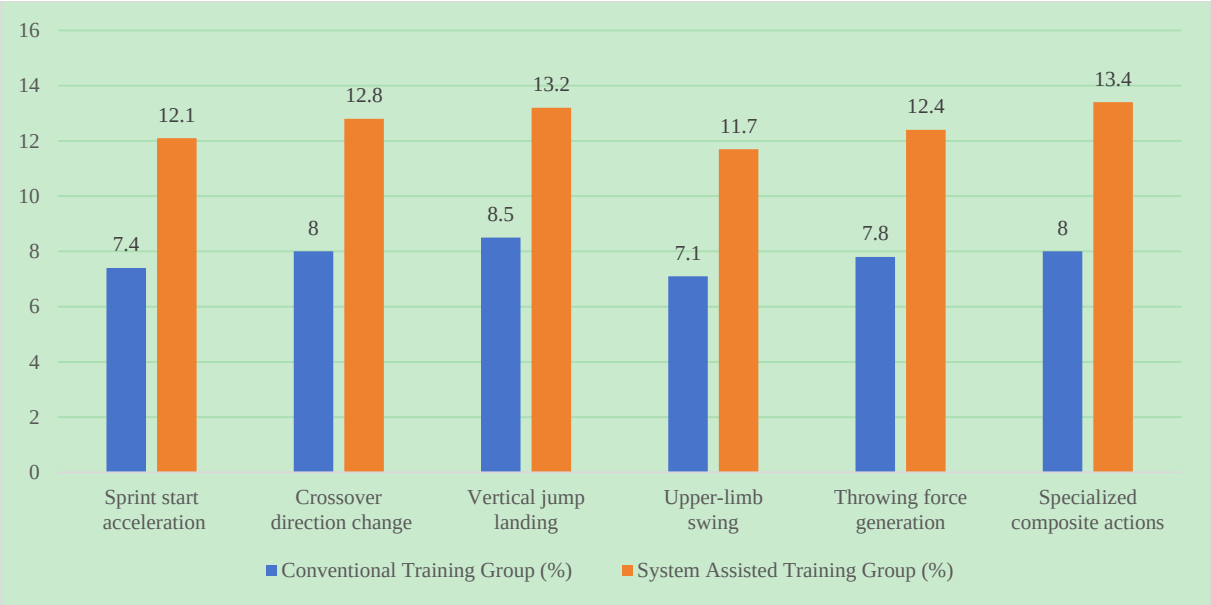

Figure 6: Improvement rate of action completion for two groups of different training actions.

It can be seen from Figure 6 that the increase of the system group in the six types of actions is higher than that of the conventional group, and the increase of the high-burst action is more concentrated, indicating that the real-time correction prompt and load regulation have a direct role in promoting the posture maintenance and force control in the rapid action. The improvement of special comprehensive action and vertical jump landing is the most obvious, indicating that the closed-loop feedback mechanism is easier to form stable gains in scenes with more complex action cohesion and more intensive rhythm changes.

Figure 7 further presents the difference in rhythm consistency improvement rate between the two groups on different training movements. The improvement of rhythm consistency in the six types of movements in the system assisted training group was higher than that in the conventional training group, and the improvement of vertical jump landing and special comprehensive movement was the most obvious, reaching 11.6% and 11.5%, which were 4.2 and 4.3 percentage points higher than that in the conventional training group. The starting acceleration, stride change direction and throwing force also maintained a relatively stable increase, indicating that the system can better improve the quality of action cohesion in the scene of continuous action chain and high-frequency rhythm switching.

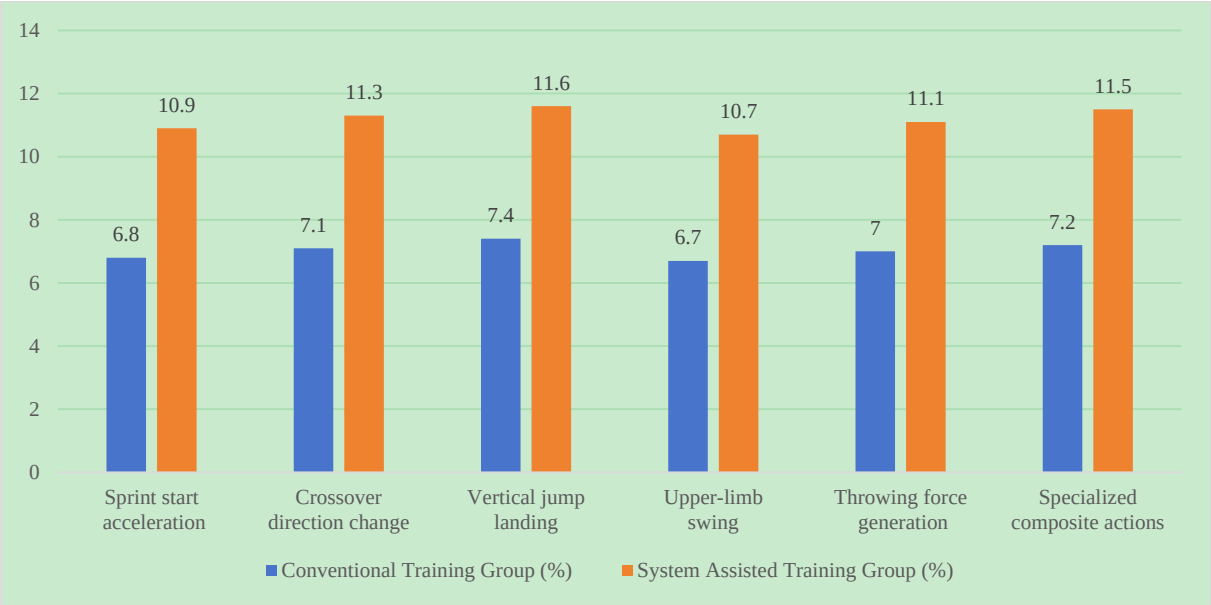

Figure 7: Improvement rate of rhythm consistency for two different groups of training actions.

Overall, the system group showed a higher increase in both movement completion and rhythm consistency, indicating that the system did not only improve the quality of a single movement, but could synchronously adjust the movement cohesion and force rhythm during the training process, so that the correction cue and the training execution formed a continuous linkage. In the phase test, the average score of the system group increased from 82.4 points to 91.3 points, and that of the conventional group increased from 82.7 points to 87.6 points. The gap between the two groups widened further in the last two weeks, indicating that the constructed system had strong application value in improving the training effect.

4.3 Real-time response and stability analysis of the system

The real-time response and stability of the system are shown in Figure 8. The figure also shows the response delay variation of the proposed system and the conventional video feedback scheme in ten consecutive rounds of training. It can be seen that the delay of the proposed system is lower than that of the conventional scheme as a whole, and the fluctuation range is smaller. In the first five rounds, the proposed system reduced from 92 ms to 76 ms, while the conventional scheme reduced from 126 ms to 113 ms. In the sixth round, both of them rebounded, but the proposed system rose to 83 ms and the conventional scheme rose to 121 ms. In the subsequent rounds, the proposed system stabilizes between 75 and 80 ms, while the conventional scheme maintains between 112 and 119 ms. This result indicates that the proposed system not only responds faster, but also operates with smaller fluctuations in the continuous training environment.

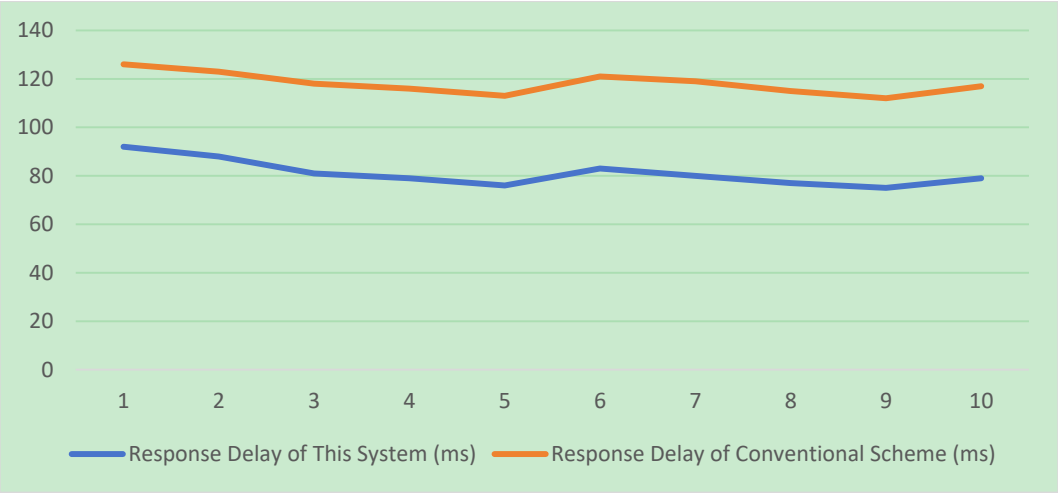

Figure 8: Comparison of response delay between the proposed system and the conventional scheme for ten consecutive rounds of training.

It can be seen from Figure 8 that the response delay of the two schemes decreases in the first five rounds, and there is a partial rebound in the sixth round, and then it falls back again and enters a relatively stable interval. Among them, the proposed system maintained between 75 and 80 ms in the subsequent rounds, and the conventional scheme maintained between 112 and 119 ms. This indicates that although the continuous training process is affected by the action complexity and signal peak changes, the overall fluctuation range of the proposed system is smaller, and there is no obvious continuous stacking phenomenon in the running links.

Table 4 presents the response results under different training actions. The average response delay of starting acceleration is the lowest, 74 ms, and the trigger success rate is the highest, reaching 97.2%. The overall performance of stride direction change is also stable, the standard deviation of delay is 3.9 ms, and the stability coefficient is 0.946. The average delay of vertical jump landing and special comprehensive action is relatively high, which is 82 ms and 84 ms, respectively, and the stability coefficient is reduced to 0.934 and 0.931, which indicates that the system computational load will increase under the conditions of air takeoff and landing, continuous connection and multimodal intensive interaction. The upper limb swing and throw force are in the middle level, and the triggering success rate is 96.0% and 95.4%, respectively, which indicates that the system can still maintain good feedback consistency in local fast swing and upper limb dominant action.

Table 4: System real-time response and stability results.

Action Type	Average Response Delay (ms)	Delay Standard Deviation (ms)	Trigger Success Rate (%)	Stability Coefficient
Sprint start acceleration	74	3.6	97.2	0.952
Crossover direction change	76	3.9	96.8	0.946
Vertical jump landing	82	4.3	95.1	0.934
Upper-limb swing	77	3.8	96.0	0.943
Throwing force generation	79	4.1	95.4	0.937
Specialized composite actions	84	4.2	94.8	0.931

In summary, the proposed system maintains a short response delay and a high triggering success rate under different training actions. When the action complexity increases, it will bring

some fluctuations, but the overall operation state is not changed. The abnormal records show that the conventional scheme has three feedback losses after continuous training, and the proposed system has only one delayed recovery, which indicates that its scheduling control is more stable in long-term operation.

5 Discussion

The digital technology-driven movement correction and training effect improvement system for sports athletes constructed in this paper shows strong comprehensive advantages in movement deviation detection, training gain maintenance and real-time feedback stability. The above results show that the deviation detection rate of the system on six types of training actions is higher than 95.9%, the false trigger rate is controlled within 2.8%, the action completion is increased by 12.6% on average, and the average response delay under continuous training conditions is maintained at about 78 ms. This indicates that multimodal perception, deviation modeling and feedback regulation do not act in isolation, but form a relatively stable synergistic relationship in the same closed loop.

This advantage mainly comes from three aspects. One is that video, skeleton, inertia, pressure, and load information are unified on the same timeline, enabling the system to simultaneously grasp spatial posture, action rhythm, and support changes. Secondly, the deviation detection model not only gives the category judgment, but also outputs the deviation position, degree and correction direction, so that the feedback results can be directly entered into the training execution link. Thirdly, the feedback control module integrates the training status and improvement trend into the decision space, so that the system can dynamically adjust the cue intensity and load rhythm according to the continuous training performance.

The system also puts forward higher requirements for operating conditions. Multi-modal synchronization, continuous log writing, and feedback-side scheduling all rely on stable hardware support, and the response fluctuations will increase in high-density training or device constrained scenarios. Complex combined actions and higher intensity continuous confrontation scenarios still need to further expand the sample coverage [24]. On the whole, the system has shown good engineering implementation ability and training application value, which can provide strong support for intelligent sports training equipment and digital training platform.

Further analysis indicates that the system maintains reliable correction performance not only at the level of isolated action segments, but also across repeated training cycles with varying intensity. During consecutive sessions, the adaptive calibration mechanism continuously updates the baseline distribution of posture deviation, joint angle fluctuation, pressure transfer characteristics, and inertial response, preventing the feedback model from being overly sensitive to temporary fatigue or short-term rhythm disorder. This dynamic calibration process improves the consistency of correction decisions and reduces the probability of unnecessary intervention in the middle stage of long-duration training.

Another notable characteristic is the interpretability of the correction process. Instead of returning a simple alarm signal, the system decomposes each detected abnormal movement into several measurable components, including deviation source, affected body region, temporal occurrence, and expected correction trajectory. Such structured outputs make it possible to transform digital detection results into human-readable guidance, which is particularly important for high-frequency training scenarios where decision time is limited. In practical deployment, this interpretability also supports rapid review of training logs and more accurate retrospective evaluation of technical adjustment paths.

From the perspective of system expansion, the current framework already provides a feasible foundation for integrating individualized athlete profiles, cross-session comparison, and online optimization of corrective thresholds. Once combined with larger longitudinal datasets, the platform can further model the relationship between deviation accumulation and performance gain, allowing the correction mechanism to move from short-term response toward medium-term training planning. This extension would strengthen the role of digital systems in athlete preparation, recovery supervision, and precision load management. In addition, the modular deployment strategy allows sensing nodes, computing units, and feedback terminals to be independently upgraded without rebuilding the entire pipeline. This reduces maintenance cost, shortens adaptation time for new training items, and significantly improves the practical sustainability of large-scale field deployment.

6 Conclusions

Aiming at the actual needs that movement deviation is difficult to quantify continuously, feedback is difficult to arrive in time and training gain is difficult to maintain stably in sports training, a movement correction and training effect improvement system for sports athletes driven by digital technology is constructed. Based on multi-modal acquisition, the system unifies video, skeleton, inertia, pressure and load information on the same time axis to form a continuous state representation for the training process. On this basis, the motion deviation detection and correction model is used to realize position positioning, degree discrimination and correction direction output, and the feedback control module is used to complete the dynamic linkage of rhythm correction, load adjustment and recovery arrangement. Experimental results show that the system performs stably in deviation detection rate, action completion improvement rate and real-time response ability, and can maintain high feedback consistency under continuous training conditions. The value of the system is not only reflected in the improvement of single action recognition accuracy, but also in the organization of action perception, deviation judgment and training intervention as a closed-loop link, so that the training process is transformed from experience-led to continuous correction and gradual improvement supported by data, which can provide practical technical reference for the construction of intelligent sports training platform and digital training equipment. In six types of special actions and continuous training scenarios, the system shows good adaptation ability, indicating that the linkage path of digital perception and feedback has certain engineering promotion value. The system has strong practicability in the actual deployment level, and has a wide range of application adaptation.

About the Author

Fan Wenxian: Associate Professor, Master's Degree, Faculty Member of the School of Physical Education, Senior Rope Skipping Coach of China, National-Level Referee. Teaches the courses “Gymnastics” and “Artistic Rope Skipping” for the Physical Education major. During her teaching career, she has repeatedly led students to participate in national championships and provincial rhythmic jump rope competitions. Her teams achieved third place in the national rhythmic jump rope championship and second place in the provincial rhythmic jump rope team total score competition. She has been honored with titles including “Outstanding Coach,” “Outstanding Referee,” and “Outstanding Organizer.” In individual professional achievements, she has won first and second prizes at the National Championships (Nanchang Station), been honored multiple times as “Outstanding Homeroom Teacher,” and received third prize in the

university's Fourth Innovative Teaching Competition. She has led three provincial-level research projects and two university-level projects, published two teaching monographs, authored three core journal articles, and contributed twelve provincial-level papers.
fanwei7211@126.com.

References

[1] Wang Y, Shan G, Li H, et al. A wearable-sensor system with AI technology for real-time biomechanical feedback training in hammer throw[J]. Sensors, 2022, 23(1): 425. https://doi.org/10.3390/s23010425

[2] Suo X, Tang W, Li Z. Motion capture technology in sports scenarios: a survey[J]. Sensors, 2024, 24(9): 2947.https://doi.org/10.3390/s24092947

[3] Xi X, Zhang C, Jia W, et al. Enhancing human pose estimation in sports training: Integrating spatiotemporal transformer for improved accuracy and real-time performance[J]. Alexandria Engineering Journal, 2024, 109: 144-156. https://doi.org/10.1016/j.aej.2024.08.072

[4] Fukushima T, Blauberger P, Guedes Russomanno T, et al. The potential of human pose estimation for motion capture in sports: a validation study[J]. Sports Engineering, 2024, 27(1): 19.https://doi.org/10.1007/s12283-024-00460-w

[5] Mercadal-Baudart C, Liu C J, Farrell G, et al. Exercise quantification from single camera view markerless 3D pose estimation[J]. Heliyon, 2024, 10(6): e27596. https://doi.org/10.1016/j.heliyon.2024.e27596.

[6] D'haene M, Chorin F, Colson S S, et al. Validation of a 3D markerless motion capture tool using multiple pose and depth estimations for quantitative gait analysis[J]. Sensors, 2024, 24(22): 7105.https://doi.org/10.3390/s24227105

[7] Dill S, Ahmadi A, Grimmer M, et al. Accuracy evaluation of 3d pose reconstruction algorithms through stereo camera information fusion for physical exercises with mediapipe pose[J]. Sensors, 2024, 24(23): 7772. https://doi.org/10.3390/s24237772

[8] Aleksic J, Kanevsky D, Mesaroš D, et al. Validation of automated countermovement vertical jump analysis: Markerless pose estimation vs. 3D marker-based motion capture system[J]. Sensors, 2024, 24(20): 6624. https://doi.org/10.3390/s24206624

[9] Ino T, Samukawa M, Ishida T, et al. Validity and reliability of OpenPose-based motion analysis in measuring knee valgus during drop vertical jump test[J]. Journal of sports science & medicine, 2024, 23(3): 515.https://doi.org/10.52082/jssm.2024.515

[10] Cronin N J, Walker J, Tucker C B, et al. Feasibility of OpenPose markerless motion analysis in a real athletics competition[J]. Frontiers in Sports and Active Living, 2024, 5: 1298003.https://doi.org/10.3389/fspor.2023.1298003

[11] Van Hooren B, Pecasse N, Meijer K, et al. The accuracy of markerless motion capture combined with computer vision techniques for measuring running kinematics[J].

Scandinavian Journal of Medicine & Science in Sports, 2023, 33(6): 966-978. https://doi.org/10.1111/sms.14319

[12] Barzyk P, Zimmermann P, Stein M, et al. AI-smartphone markerless motion capturing of hip, knee, and ankle joint kinematics during countermovement jumps[J]. European Journal of Sport Science, 2024, 24(10): 1452-1462.https://doi.org/10.1002/ejsc.12186

[13] White M, De Lazzari B, Bezodis N, et al. Wearable sensors for athletic performance: A comparison of discrete and continuous feature-extraction methods for prediction models[J]. Mathematics, 2024, 12(12): 1853.https://doi.org/10.3390/math12121853

[14] Kovoor M, Durairaj M, Karyakarte M S, et al. Sensor-enhanced wearables and automated analytics for injury prevention in sports[J]. Measurement: Sensors, 2024, 32: 101054. https://doi.org/10.1016/j.measen.2024.101054

[15] Zhao J, Yang Y, Bo L, et al. Research progress on applying intelligent sensors in sports science[J]. Sensors, 2024, 24(22): 7338. https://doi.org/10.3390/s24227338

[16] Rehman S U, Yasin A U, Ul Haq E, et al. Enhancing human activity recognition through integrated multimodal analysis: a focus on rgb imaging, skeletal tracking, and pose estimation[J]. Sensors, 2024, 24(14): 4646.https://doi.org/10.3390/s24144646

[17] Mekruksavanich S, Phaphan W, Hnoohom N, et al. Recognition of sports and daily activities through deep learning and convolutional block attention[J]. PeerJ Computer Science, 2024, 10: e2100.https://doi.org/10.7717/peerj-cs.2100

[18] Gilmore J, Nasseri M. Human activity recognition algorithm with physiological and inertial signals fusion: photoplethysmography, electrodermal activity, and accelerometry[J]. Sensors, 2024, 24(10): 3005.https://doi.org/10.3390/s24103005

[19] Pellano K N, Strümke I, Ihlen E A F. From movements to metrics: Evaluating explainable AI methods in skeleton-based human activity recognition[J]. Sensors, 2024, 24(6): 1940. https://doi.org/10.3390/s24061940

[20] Roggio F, Trovato B, Sortino M, et al. A comprehensive analysis of the machine learning pose estimation models used in human movement and posture analyses: A narrative review[J]. Heliyon, 2024, 10(21): e39977.https://doi.org/10.1016/j.heliyon.2024.e39977

[21] Tharatipyakul A, Srikaewsiew T, Pongnumkul S. Deep learning-based human body pose estimation in providing feedback for physical movement: A review[J]. Heliyon, 2024, 10(17): e36589.https://doi.org/10.1016/j.heliyon.2024.e36589

[22] Davis J, Bransen L, Devos L, et al. Methodology and evaluation in sports analytics: challenges, approaches, and lessons learned[J]. Machine Learning, 2024, 113(9): 6977-7010. https://doi.org/10.1007/s10994-024-06585-0

[23] Zheng C, Wu W, Chen C, et al. Deep learning-based human pose estimation: A survey[J]. ACM computing surveys, 2023, 56(1): 1-37.https://doi.org/10.1145/3603618

[24] Contoli C, Freschi V, Lattanzi E. Energy-aware human activity recognition for wearable devices: A comprehensive review[J]. Pervasive and Mobile Computing, 2024, 104: 101976. https://doi.org/10.1016/j.pmcj.2024.101976