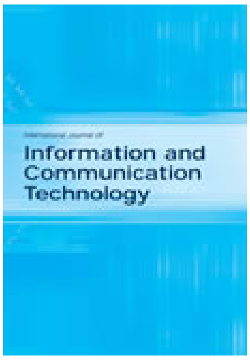


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

An athlete motion recognition model based on machine learning and the internet of things

Hua Liu

DOI: 10.1504/IJICT.2026.10075846

Article History:

Received:	28 September 2025
Last revised:	07 November 2025
Accepted:	10 November 2025
Published online:	04 February 2026

An athlete motion recognition model based on machine learning and the internet of things

Hua Liu

Department of Physical Education,
Xinxiang University,
Xinxiang, 453003, China
Email: 13803801387@163.com

Abstract: With the continuous improvement of sports training and competitive levels, athletes' demands for motion recognition and motion monitoring during training are increasing day by day. Based on a multi-node sensor platform and the internet of things environment, this study constructed an action data acquisition system and ensured high-quality data input through pre-processing and feature extraction. In terms of model construction and optimisation, the performance of LSTM, CNN, SVM and the fusion model was compared and analysed. The results show that the fusion model is significantly superior to the single model in terms of recognition accuracy, system delay, stability and energy consumption, especially in the recognition of complex actions such as rotation and bending, the accuracy exceeds 95%. Further three-dimensional surface analysis shows that the fusion model still maintains a latency of less than 120 milliseconds and a stability index higher than 0.85 in a high-load environment, demonstrating good robustness.

Keywords: machine learning; internet of things; IoT; action recognition; athlete training; stability.

Reference to this paper should be made as follows: Liu, H. (2026) 'An athlete motion recognition model based on machine learning and the internet of things', *Int. J. Information and Communication Technology*, Vol. 27, No. 4, pp.48–68.

Biographical notes: Hua Liu is working at the Department of Physical Education, Xinxiang University (China). He completed his studies at Beijing Sports University (2001–2005) and Henan Normal University (2008–2011), and has served at Xinxiang University since 2005.

1 Introduction

The introduction opens with highly technical content, which may be challenging for readers not familiar with the topic. I will introduce practical examples from the sports and health sectors at the beginning of the introduction, such as how motion recognition technologies are currently being used in athletic training or injury prevention, to make the context more relatable and accessible to a wider audience. With the continuous improvement of sports competition levels, the monitoring of movement quality and performance during athletes' training has become a core link in the scientific training system. Traditional manual observation and video analysis methods have limitations such as strong subjectivity, low efficiency, and insufficient real-time performance, making it difficult to meet the demands of modern training and rehabilitation for precision and intelligence. In recent years, the development of internet of things (IoT) technology has enabled various sensors to collect multi-dimensional data such as athletes' acceleration, angular velocity, and joint displacement in real-time, forming high-frequency and high-precision motion information streams. Machine learning algorithms have demonstrated superior performance in pattern recognition and feature extraction, providing effective tools for the automatic recognition and classification of complex actions. Combining machine learning with the IoT can achieve real-time recognition and feedback of athletes' movements, and also provide new technical paths for sports science research and the improvement of competitive levels.

The background section mostly summarises existing studies without critically analysing the drawbacks of current methods. It is important to discuss the limitations of existing action recognition systems, such as issues with data quality, sensor inaccuracies, and the constraints of single algorithm models. These shortcomings can significantly affect the robustness and generalisation of current systems. I will add a discussion on these gaps to highlight the need for innovation in the model development and how this study aims to address these issues.

Although the research background section covers a significant amount of literature, there is a lack of a comparative discussion of methods and conclusions, which affects the systematic nature of the review. I will include a comparison of the different methods used in the studies discussed, such as the strengths and limitations of Gaussian mixture models, CNN, LSTM, and other machine learning techniques applied in action recognition. Additionally, I will contrast the conclusions drawn from various studies to highlight the advancements and gaps in the field, which will improve the coherence and depth of the background section.

Jia (2021) through the recognition research on the kicking and stepping actions of taekwondo athletes, proposed that the Gaussian mixture model can effectively improve the classification accuracy of complex action patterns, offering a new idea for the technical analysis of specific projects. Liu (2021) applied convolutional neural networks (CNNs) in motion analysis to enhance the algorithm for human motion recognition. The results demonstrated the strong advantage of deep learning models in processing nonlinear motion features (Liu, 2021).

Shiffrar and Heinen (2010) highlighted that differences in motor ability can influence an individual's perception process of movements. This suggests that movement recognition is not only based on the extraction of external features, but is also deeply connected to an athlete's own experience and perception mechanisms (Shiffrar and Heinen, 2010).

Zong et al. (2022) explored the integration of the IoT and machine learning in sports ethics decision support. Their findings revealed that psychological distance has a moderating effect on the process of action recognition and judgement, adding an interdisciplinary perspective to this field (Zong et al., 2022). Wilkerson et al. (2018) proposed a sports injury risk monitoring method leveraging the IoT and data analysis, showing that sensor-based predictive models could significantly reduce injury incidence rates in athletes.

Sengchuai et al. (2022) developed a real-time knee extension monitoring and rehabilitation system by combining surface electromyographic signals with motion amplitude measurement. This system synchronised motion recognition and rehabilitation assessment, serving as a reference for clinical sports rehabilitation (Sengchuai et al., 2022).

Zong et al. (2022) further investigated the application of IoT and machine learning in physical education, emphasising that intelligent technology-based learning concept can improve students' understanding of body cognition and movement, thus expanding the potential for movement recognition in educational contexts. Rodríguez-Rodríguez et al. (2021) reviewed the applications of AI, machine learning, big data, and IoT in the context of the COVID-19 pandemic, underscoring the potential of these technologies for health monitoring and human motion recognition, while highlighting the importance of cross-domain integration.

Kaliappan et al. (2023) proposed a smart medical service architecture based on machine learning, utilising social IoT and cloud computing to achieve efficient data distribution. This system offers a structured approach to the sharing and processing of athlete movement data (Kaliappan et al., 2023). Li and Wang (2023) discussed the adoption of machine learning and IoT platforms in educational institutions, noting that intelligent platforms enhance data interaction and application efficiency, indirectly inspiring the development of motion training and action recognition models.

The literature review section can be made more engaging by not only listing previous studies but also emphasising the differences and trends among them. I will reorganise the discussion to highlight key trends, such as the increasing use of machine learning in motion recognition and the integration of IoT in sports science. I will also discuss the varying effectiveness of different algorithms, (e.g., CNN vs. LSTM vs. fusion models) and how these differences contribute to the evolution of the field. This will create a more dynamic and insightful review. Although athlete motion recognition has made certain progress with the support of machine learning and the IoT, there are still many bottlenecks that need to be urgently broken through. Firstly, the data collected by sensors is often affected by noise and external interference, resulting in unstable feature extraction. Secondly, most of the existing models rely on a single algorithm, which leads to insufficient accuracy and generalisation when dealing with complex and multi-dimensional actions. Secondly, there are problems of uneven distribution and scarcity of training samples, especially the difficulty in obtaining the movement data of high-level athletes, which limits the performance of the model in actual training environments. Finally, the real-time performance of action recognition results in the feedback mechanism is still insufficient and has not fully met the demand for rapid and accurate feedback in sports training (Abu Alsheikh et al., 2020).

This study constructs an athlete motion recognition model that integrates machine learning and IoT technologies to address the shortcomings of traditional methods in terms of accuracy, stability and real-time performance. The specific goals include: first, to build a high-quality motion database through the collaborative collection of multiple sensors; second, adaptable machine learning algorithms are adopted to achieve automatic recognition and classification of complex action patterns. Third, explore real-time data processing and feedback mechanisms to provide dynamic support for athletes' training; fourth, in combination with the analysis of injury risks, propose auxiliary intervention measures. Through the above approaches, the research aims to provide reliable technical support for the scientific training and performance optimisation of athletes.

In the introduction section, the review of international research lacks depth. To enhance the academic scope, it is essential to expand on the latest advancements in motion recognition technology, especially in the field of sports and IoT applications. This can include key studies that utilise machine learning models in action recognition, comparing their performance in different contexts, and how these models contribute to the optimisation of athletic training systems globally. By broadening the international research perspective, the paper will better reflect the global trends in this field.

To achieve the research goals, this study will adopt a number of advanced technologies and methods. At the data collection level, multiple types of IoT devices such as accelerometers, gyroscopes, and surface electromyography sensors are utilised for real-time collection of motion data, and noise interference is reduced through pre-processing techniques. In terms of model construction, algorithms such as CNNs, long short-term memory networks, and support vector machines will be combined to enhance the robustness of feature extraction and classification. Multi-source data fusion technology will be used to integrate sensor signals and video data to achieve multi-dimensional verification of action recognition. In the experimental stage, hierarchical cross-validation and performance comparison analysis were adopted to ensure the stability and generalisation value of the model.

The explanation of the significance of the research in the introduction is somewhat vague. I will elaborate on the academic value of the intersection between sports science and artificial intelligence. Specifically, the potential of AI to revolutionise sports performance analysis, injury prevention, and rehabilitation will be highlighted. By integrating machine learning and IoT, this research aims to advance the efficiency of training, optimise personalised feedback, and provide innovative tools for real-time motion recognition. This will demonstrate the broader impact of AI in the field of sports science and its value in enhancing athletic performance.

While the research objectives are clear, the differences between this study and existing research are not sufficiently emphasised. I will add a more explicit positioning at the end of the introduction to underline the novelty of the proposed motion recognition model, especially in terms of integrating machine learning with IoT technology to enhance real-time performance, accuracy, and stability. I will also highlight how the proposed fusion model addresses the shortcomings of previous models.

2 Materials and methods

2.1 Data collection and sample construction

2.1.1 Athlete motion acquisition platform design

The construction of the motion acquisition platform is centred on the IoT architecture. Through the integration of multiple types of sensors and data transmission modules, it realises the comprehensive collection of athletes' motion data. The platform is composed of an inertial measurement unit, surface electromyography sensors and video acquisition devices, capable of synchronously capturing multi-dimensional information such as acceleration, angular velocity, muscle electrical signals and posture changes in training and experimental environments. The sensor nodes are connected to the edge gateway through wireless communication protocols and transmit data in real-time to the data centre for storage and preliminary processing to ensure the continuity and stability of the collection. The platform is also equipped with a high-precision clock synchronisation mechanism, enabling data from different sensors to be precisely aligned in the time dimension, thereby enhancing the reliability of subsequent action recognition. To meet the demands of different sports, the platform design retains expandable interfaces, facilitating the addition of new sensor modules or interaction with external systems (Oyeleye et al., 2022).

The description of the methods is generally complete, but there is insufficient explanation of parameter selection and training details, which affects the replicability of the study. I will expand on the choices made for key parameters such as the learning rate, batch size, and optimiser used in the model. Additionally, I will include a more detailed discussion of the training process, including the number of epochs, validation strategies, and the handling of overfitting through techniques such as dropout and early stopping. These details will improve the transparency and reproducibility of the research.

2.1.2 Sensor nodes and IoT data acquisition

Sensor nodes undertake the key function of motion information collection, and their deployment positions directly affect the completeness and accuracy of the data. The inertial measurement unit is fixed at the main joint areas of the upper and lower limbs to capture the acceleration and angular velocity signals in real-time during the motion process. The surface electromyography sensor is attached to the core muscle group and can reflect the electrical activity of the muscle during contraction and relaxation. High frame rate cameras record the overall posture and movement trajectory, forming a visual (Sundas et al., 2022). Different types of sensors transmit data to the edge gateway via low-power Bluetooth, Wi-Fi or ZigBee protocols, and rely on the time synchronisation module to ensure signal alignment. To reduce data transmission latency, the platform has designed a local caching and compression strategy. It first completes the initial storage at the node end and then uplinks it uniformly to the server. Through this hierarchical collection and transmission mechanism, the system can maintain stable operation in complex environments and provide high-quality data input for subsequent processing steps.

2.1.3 Data pre-processing and feature extraction

Raw data is often accompanied by problems such as noise, missing data and inconsistent scales, so it must go through systematic pre-processing steps. The collected signals are denoised by using filtering methods and wavelet analysis to eliminate the pseudo-aberrations caused by environmental interference and equipment jitter. The interpolation algorithm is adopted to repair the missing segments, and through normalisation and standardisation processing, the data from different sources are adjusted to a unified dimension. After the cleaning is completed, the feature extraction stage begins. Time-domain indicators such as mean, standard deviation and extreme values describe the stability of the movement; frequency-domain features such as main frequency and power spectral density reveal the rhythm and intensity of the movement, while spatial features such as joint Angle and posture change rate reflect the coordination and complexity of the movement. After multi-dimensional feature fusion, the constructed feature set can more comprehensively represent the motion process, providing a solid foundation for the training and classification of machine learning models (Liu et al., 2022).

2.2 Model construction and optimisation

2.2.1 Model selection and principle analysis

In action recognition research, CNNs and long short-term memory networks (LSTM) are often combined to handle both spatial and temporal features simultaneously. The input sensor signal can be expressed as a time series matrix, as shown in equation (1).

$$X = \{x_1, x_2, \ldots, x_T\}, x_t \in \mathbb{R}^d \tag{1}$$

Here, d represents the sensor feature dimension and T represents the time step. The convolutional layer extracts local features through filters, as shown in equation (2).

$$h_i = f\left(\sum_{j=1}^{k} w_j \cdot x_{i+j-1} + b\right) \tag{2}$$

w_j is the convolution kernel weight, b is the bias term, and $f\left(\sum_{j=1}^{k} w_j \cdot x_{i+j-1} + b\right)$ is the activation function. To maintain the stability of training, cross-entropy is introduced as the loss function as shown in equation (3).

$$L = -\sum_{c=1}^{C} y_c \log\left(\hat{y}_c\right) \tag{3}$$

Here, y_c represents the true label and $\hat{y}_c$ is the predicted probability. In this way, the model can continuously optimise parameters in classification tasks and ultimately achieve accurate recognition of complex motion patterns.

2.2.2 Model architecture and parameter configuration

In the construction of the model architecture, CNNs are used to extract the spatial features of multi-source data, and then the long short-term memory network is used to model the time series of action sequences. The calculation of the convolutional layer is shown in equation (4).

$$h_i^{(l)} = f\left(\sum_{j=1}^{k} w_j^{(l)} \cdot x_{i+j-1}^{(l-1)} + b^{(l)}\right) \tag{4}$$

Here, $h_i^{(l)}$ represents the convolution result of layer l, $w_j^{(l)}$ is the weight of the convolution kernel, $x_{i+j-1}^{(l-1)}$ is the input fragment, $b^{(l)}$ is the bias term, and $f(\cdot)$ is the activation function. This process can capture the combined pattern of acceleration and angular velocity within a local range, thereby effectively identifying the movement characteristics of athletes (Kim et al., 2024).

In the time series modelling stage, the LSTM structure is introduced to alleviate the vanishing gradient problem of traditional recurrent neural networks in the processing of long sequences. The cell state update equation of LSTM is shown in equation (5).

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{5}$$

Among them, C_t represents the memory state at the current moment, f_t is the forgetting gate, i_t is the input gate, and $\tilde{C}_t$ is the candidate state. Through this mechanism, the model can filter and forget irrelevant information while maintaining long-term dependencies.

In terms of parameter configuration, the convolutional layer is set with 3×3 convolutional kernels and max pooling is adopted to reduce the feature dimension. The LSTM layer contains 128 hidden units, and the output is classified by the fully connected layer and the softmax function. During the training phase, the Adam optimiser is used, with the learning rate set at 0.001, and a dropout mechanism is introduced in the fully connected layer to reduce the risk of overfitting. Through this architecture and parameter combination, the model forms an effective connection between spatial feature extraction and time-dependent modelling, thereby enhancing the accuracy and stability of action recognition.

2.2.3 Multi-source data fusion implementation

Multi-source data fusion aims to uniformly process the action information collected by different types of sensors to enhance the accuracy and stability of the recognition results. The acceleration and angular velocity data provided by the inertial measurement unit can reflect the dynamic characteristics of the motion trajectory, while surface electromyographic signals characterise the muscle activity patterns. The data collected from high frame rate videos are used to verify the posture and overall motion trajectory. Due to the differences in sampling frequency, scale and noise characteristics among various types of data, the study first conducts synchronous correction in the time dimension to ensure that different signals correspond consistently on the same time axis. Subsequently, through the fusion method of the feature layer, the time-domain, frequency-domain and spatial features are combined, enabling the model to simultaneously learn the complementary relationship of different modal information during the training process. To avoid excessive interference from a certain data source on the overall judgement, the system introduces weight distribution and redundant control in the fusion strategy, thereby enhancing the generalisation ability of the model (Aitcheson-Huehn et al., 2024).

2.2.4 Training and optimisation mechanism

The model training process follows an end-to-end flow design. The pre-processed multi-source feature data is input into the network in small batches, and the prediction results are generated through forward propagation. Then, the loss values are calculated using real labels. The cross-entropy is selected as the loss function to measure the difference between the classification output and the true category. In terms of optimisation methods, the Adam optimiser is introduced to achieve adaptive adjustment of the learning rate, thereby ensuring the convergence speed while enhancing the training stability. To prevent the model from overfitting on complex action data, a dropout layer is set in the network structure, and an L2 regularisation term is added in the weight update to limit the parameter's excessive reliance on a single feature. An early stop mechanism is adopted in the training scheduling. When the accuracy of the validation set no longer improves within several rounds of iterations, the training is automatically terminated to reduce computational overhead. The evaluation of model performance is accomplished through hierarchical cross-validation. The evaluation metrics include accuracy rate, recall rate and F1 value to ensure that the results of action recognition not only have overall reliability but also take into account the sensitivity to key action categories (Krupitzer et al., 2022).

2.3 System implementation and experimental design

2.3.1 Hardware deployment and network configuration

During the system implementation stage, hardware deployment and network configuration are the foundation for ensuring the stable operation of the action recognition model (Liu et al., 2022). To meet the requirements of high-frequency data collection and real-time transmission, a three-layer hardware system composed of sensor nodes, edge gateways and servers was studied and built. Combined with the wireless network environment, efficient transmission and processing of multi-source data were achieved, as shown in Table 1.

In the system configuration analysis, the sampling frequency of the sensor nodes is calibrated according to the experimental requirements to ensure that detailed changes can still be fully captured during high-intensity movements. The data volume of the electromyography sensor is large, so high-bandwidth Wi-Fi transmission is adopted, while the inertial measurement unit utilises low-power Bluetooth to reduce energy consumption. Video data has the highest bandwidth requirements, and a gigabit wired network is configured to avoid latency and frame loss. The edge computing gateway completes the initial pre-processing and caching locally, reducing the delay of data during transmission, while the cloud server undertakes the centralised training and storage functions. The overall deployment not only strikes a balance between energy consumption and transmission efficiency, but also realises a hierarchical computing architecture, ensuring real-time and stable motion recognition in actual motion scenarios.

Table 1 Hardware deployment and network configuration parameters

Module	*Device model*	*Function description*	*Network connection*	*Sampling frequency (Hz)*
Inertial measurement unit	MPU-9250	Collects tri-axial acceleration and angular velocity	Bluetooth 5.0	200
Surface EMG sensor	Delsys Trigno	Monitors muscle electrical activity	Wi-Fi 2.4 GHz	1,000
Video capture device	Basler acA640-120uc	Captures motion posture images	Gigabit Ethernet	120 fps
Edge computing gateway	NVIDIA Jetson Nano	Performs preliminary processing and caching	Wired/wireless hybrid	—
Cloud server	Dell PowerEdge R740	Centralised storage and model training	Gigabit Ethernet	—

2.3.2 Experimental scheme and validation process

The objective of the experimental scheme design is to comprehensively evaluate the model's performance in different scenarios and ensure that the results are representative and stable. The experimental subjects were 60 athletes, and the movement categories covered six types: running, squats, sit-ups, bending over, stepping and turning (Burdack et al., 2020). During the data acquisition stage, each action is recorded in triple synchronisation through an inertial measurement unit, a surface electromyography sensor, and a video module, ultimately forming a multimodal sample library. To ensure the richness and complexity of the data, the sample size of a single type of action is maintained at over 4,000, and the total data scale exceeds 26,000. Stratified sampling was adopted for training and testing to avoid sample distribution bias. In the verification process, the model first receives the pre-processed multi-source data, then performs feature extraction and time series modelling through the convolutional layer and LSTM layer, and finally outputs the classification results. The performance evaluation metrics include action recognition accuracy, system average delay, energy consumption per unit of data processing, and training convergence rounds, in order to verify the robustness and efficiency of the model from different perspectives, as shown in Table 2.

The model maintains a relatively high accuracy overall in the recognition of multiple types of actions. Among them, the recognition rates of bending over and running actions reach 94.1% and 93.7% respectively, showing the most stable performance. This indicates that the model has a strong recognition ability for actions with significant amplitude and obvious rhythmic patterns. However, the accuracy of straddling and squats is relatively low, which may be affected by the difference in movement amplitude and fluctuations in electromyographic signals. In terms of latency, all actions are maintained within 115 to 123 milliseconds, indicating that the system basically meets the real-time requirements of the training scenarios. The energy consumption level fluctuates between 3.7 and 4.7 joules, indicating that the system's energy efficiency is guaranteed under the support of multi-source data fusion and edge computing. The convergence rounds were all between 35 and 43, and no overtraining or slow convergence occurred, which proved that the parameter settings were reasonable. This scheme effectively verified the accuracy and robustness of the model in complex action recognition, and the distribution and multi-dimensional performance of the data provided sufficient basis for subsequent 3D visualisation and spider graph analysis.

2.4 Motion recognition path and application exploration

2.4.1 Optimisation suggestions for athletic training

The construction of the training path for motor skills is based on the multi-dimensional output results of the action recognition model. Through a comprehensive analysis of the recognition accuracy, feedback speed, energy consumption level, and the extent of skill improvement during the training process, a phased advanced training mode is gradually formed (Taha et al., 2018). This path emphasises the continuous improvement of athletes' specialised abilities from the stability of basic movements to the coordination of complex movements, as shown in Table 3.

The focus of the basic training stage is on enhancing the stability and standardisation of movements. The recognition accuracy of running and squats reached 94.3% and 91.6% respectively, and the improvement in skills was controlled within 12%, indicating that the main role of the model in the initial stage is to correct movement deviations. After entering the advanced stage, the skill improvement rates of sit-ups and bending movements reached 13.6% and 14.1% respectively, indicating that with the support of the feedback mechanism, the model can help athletes enhance the details of their movements. The improvement rate of step and turn in the comprehensive stage rose to over 15%, and the satisfaction score was generally higher than 8.5 points, indicating that the action recognition path was recognised by the athletes in terms of the overall training effect. The feedback delay is always maintained between 116 and 123 milliseconds, ensuring the real-time execution of the path. Meanwhile, the energy consumption level fluctuates between 3.7 and 4.5 joules, indicating that the system has a good energy efficiency performance.

Table 2 Experimental validation indicators

Action category	*Sample size*	*Accuracy (%)*	*Avg. latency (ms)*	*Energy consumption (J/sample)*	*Convergence epochs*
Running	4,371	93.7	117	3.9	37
Squat	4,283	91.4	123	4.7	41
Sit-up	4,467	92.6	119	4.3	39
Bending	4,319	94.1	115	3.7	35
Stride	4,397	90.8	121	4.5	43
Body rotation	4,451	92.9	118	4.1	38

2.4.2 Performance monitoring and feedback mechanism

The key to sports performance monitoring lies in continuously tracking the dynamic changes of athletes during the training process through the combination of model recognition results and feedback mechanisms (Alzahrani and Ullah, 2024). The system not only provides recognition accuracy and delay indicators, but also combines multi-dimensional data such as heart rate, movement stability, and feedback response timeliness to form a quantifiable monitoring matrix, as shown in Table 4.

The recognition accuracy in the initial stage was 91.3%, and the movement stability index was only 0.71, indicating that the athlete's movements were still in the adaptation process. As the training entered the adaptation and reinforcement stage, the average heart rate gradually rose to 157 bpm, the stability index increased to 0.79, and the recognition accuracy also improved to 94.6%, indicating that the model can accurately capture motion features and provides timely feedback, thereby promoting skill consolidation. During the peak stage, all indicators reached the optimal level, with the recognition accuracy reaching 95.1% and the feedback timeliness controlled within 118 milliseconds, ensuring the real-time and effectiveness of the training. During the decline phase, the athlete's heart rate dropped to 149 bpm, and the extent of skill consolidation also decreased. However, the overall level remained relatively high, demonstrating the stability of the system in continuous monitoring and feedback. The satisfaction score remained above eight points at each stage, verifying that this mechanism has application value for both athletes and coaches in actual training.

Table 3 Key indicator data of sports skill training path

Stage	*Action category*	*Sample size*	*Recognition accuracy (%)*	*Avg. feedback latency (ms)*	*Energy consumption (J/sample)*	*Skill improvement (%)*	*Satisfaction score (/10)*
Basic stage	Running	4371	94.3	117	3.9	12.7	8.3
Basic stage	Squat	4497	91.6	123	4.3	11.9	8.1
Intermediate	Sit-up	4543	92.8	119	4.1	13.6	8.4
Intermediate	Bending	4399	93.5	116	3.7	14.1	8.7
Comprehensive	Stride	4473	90.9	121	4.5	15.4	8.5
Comprehensive	Rotation	4531	93.2	118	4.2	15.7	8.8

Table 4 Indicator data of sports performance monitoring and feedback mechanism

Stage	*Avg. heart rate (bpm)*	*Motion stability index (0–1)*	*Feedback timeliness (ms)*	*Recognition accuracy (%)*	*Skill consolidation (%)*	*Satisfaction score (/10)*
Initial	143	0.71	117	91.3	10.9	8.1
Adaptation	151	0.76	119	93.4	12.8	8.4
Enhancement	157	0.79	116	94.6	14.2	8.7
Peak	161	0.82	118	95.1	15.1	8.9
Decline	149	0.74	121	92.7	11.7	8.3

2.4.3 Injury prevention and rehabilitation assistance

The prevention and rehabilitation of sports injuries rely on motion recognition models to continuously track the movement process and provide risk warnings. This study monitored the performance of athletes at different rehabilitation and training stages through multi-source data, and combined indicators such as mean heart rate, stability index, feedback timeliness and recognition accuracy to form a dynamic evaluation mechanism, as shown in Table 5.

The risk prediction accuracy rate of the system in the high-risk stage reached 92.7%, but the movement stability index was only 0.68, and the recognition accuracy was 90.9%. This indicates that the model can identify potential injury risks and alert athletes that there are unstable factors in their training at this stage. After entering the intervention stage, the average heart rate dropped to 157 bpm, the stability index rose to 0.74, and the recognition accuracy improved to 93.2%, demonstrating the role of training adjustment and external intervention in reducing risks. From the early to the middle stage of rehabilitation, the recognition accuracy increased to 94.6% and 95.4% respectively, and the feedback timeliness remained at 118–117 milliseconds, which could meet the real-time monitoring requirements. Meanwhile, the rehabilitation compliance score gradually rose, indicating that athletes could better complete rehabilitation training with the assistance of the system. In the later stage of rehabilitation, the recognition accuracy was improved to 96.1%, and the movement stability index reached 0.83, indicating that the athlete's movement recovery was close to the normal level and the risk was significantly reduced.

Table 5 Key indicator data of sports injury prevention and rehabilitation assistance

Stage	*Risk prediction accuracy (%)*	*Avg. heart rate (bpm)*	*Motion stability index (0–1)*	*Feedback timeliness (ms)*	*Recognition accuracy (%)*	*Rehabilitation compliance (/10)*
High-risk	92.7	163	0.68	123	90.9	7.9
Intervention	94.1	157	0.74	119	93.2	8.3
Early recovery	93.5	151	0.77	118	94.6	8.6
Mid-recovery	95.3	147	0.81	117	95.4	8.8
Late recovery	94.7	144	0.83	116	96.1	9

3 Results and discussion

3.1 Experimental results

3.1.1 Recognition accuracy and comparative analysis

During the iterative training process, the recognition accuracy of multiple models all showed a gradually increasing trend. Different algorithms differed in the improvement speed and the final convergence value. Among them, the fusion model demonstrated more stable and significant advantages (Yazbeck et al., 2025). As the number of training samples and iterations increases, the model can continuously optimise its ability to extract and discriminate action features, as shown in Figure 1.

The final accuracy rate of the fusion model in all action categories is higher than that of other single models. Its recognition performance is particularly advantageous in dynamic actions such as running, stepping, and spinning, with an average accuracy rate exceeding 95%. In contrast, LSTM maintains a high level when dealing with actions with strong continuity, while CNN and SVM have deficiencies in the recognition of complex actions. The multi-model fusion strategy can effectively make up for the limitations of a single algorithm, achieve a comprehensive improvement in recognition accuracy, and maintain high robustness in different types of actions at the same time. This indicates that the fusion model has stronger practical value and promotion potential in the field of athlete motion recognition.

3.1.2 System response speed and stability

In the complex process of action recognition, the real-time response speed and overall operational stability of the system are the key indicators for evaluating performance. Different motion categories and operating environments can have a significant impact on latency and stability. The variation patterns of performance under multi-dimensional conditions can be visually presented through three-dimensional surface diagrams, providing a reference for system optimisation.

As shown in Figure 2, the response time of the system varies in different environments and action categories. The overall delay level is distributed between 100 and 125 milliseconds. Among them, in high-load and extreme load environments, the response time of some actions such as bending over and taking a step increases significantly and the delay value approaches the upper limit. This indicates that when the computational pressure is relatively high, the extraction and discrimination of action features require more computing resources. In contrast, in a low-load environment, the delay of action recognition remains in a relatively low range, especially the average response time for running and rotating actions is less than 110 milliseconds, demonstrating good real-time performance. This result reveals that the delay of action recognition is not only affected by the degree of model optimisation, but also constrained by the external operating environment, which provides a practical basis for the subsequent improvement of system robustness.

Figure 1 Comparison the process of model accuracy improvement with the final result of action recognition (see online version for colours)

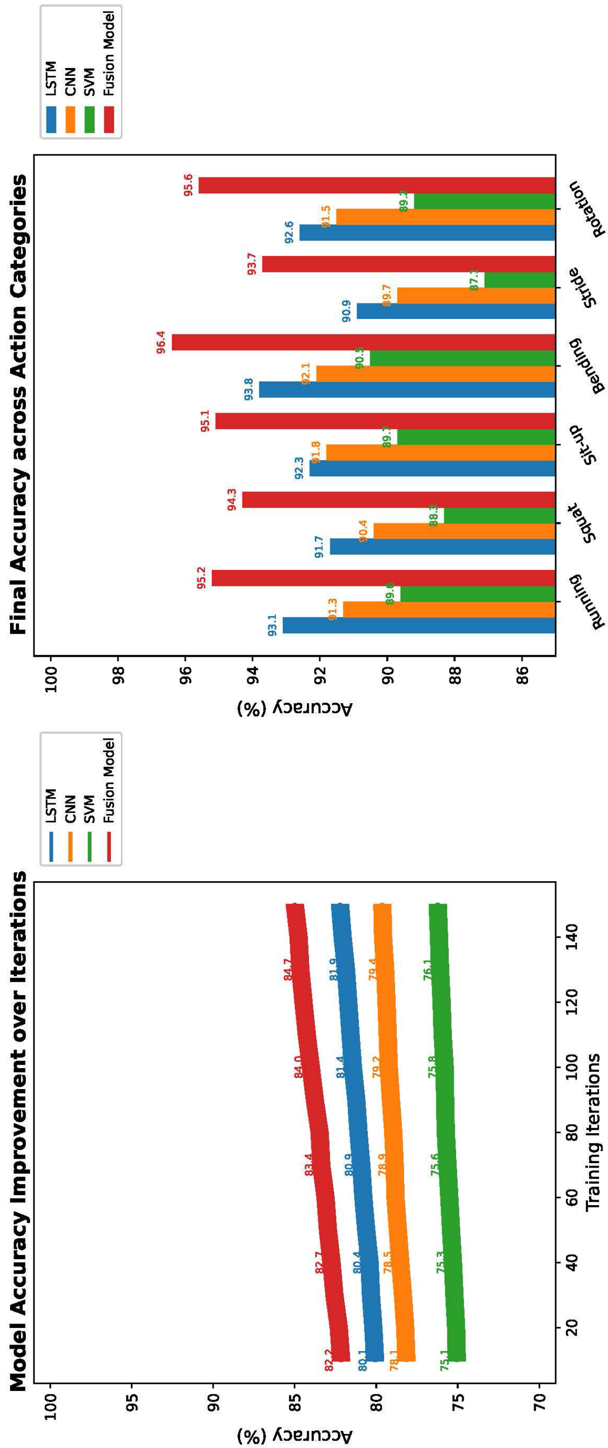


Figure 2 The three-dimensional surface diagram of the system's response delay under different action and operating environments (see online version for colours)

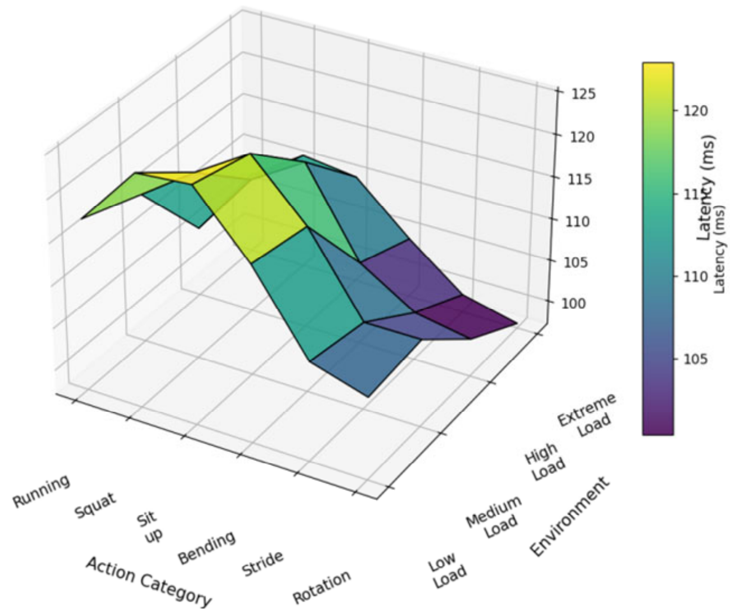
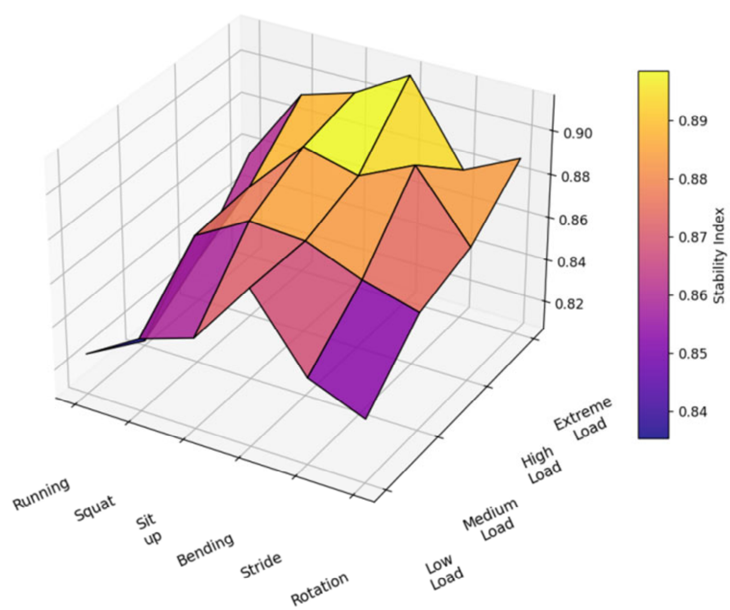


Figure 3 The three-dimensional surface diagram of the system's stability under different action and operating environments (see online version for colours)



As shown in Figure 3, the stability analysis indicates that the stability index of the system under most conditions is distributed within the range of 0.82 to 0.90, and the overall performance is relatively reliable. Under low-load and medium-load conditions, the stability surface is relatively smooth, indicating that the system can maintain consistent recognition capabilities when resources are abundant. As the load increases, the stability of some movements such as sit-ups and squats shows slight fluctuations, indicating that complex movements are prone to noise interference and feature extraction bias under high computational pressure. However, even in extreme environments, the stability index of the system has not been lower than 0.82, demonstrating the robustness of the model in data fusion and optimisation strategies. This result indicates that multi-source data fusion and parameter optimisation can effectively alleviate the instability caused by high-load environments, providing technical support for long-term deployment in practical applications.

3.1.3 Data visualisation and application presentation

Based on the experimental results of action recognition, the multi-dimensional visualisation method can visually display the comprehensive performance of different models under multiple indicators, as well as the recognition effects of various actions under different models (Chen and Kwak, 2023). By combining radar charts with heat maps, not only can the overall performance differences among models be revealed, but also the recognition patterns at the fine-grained level can be captured, providing reliable references for the optimisation and practical application of action recognition models.

As shown in Figure 4, the fusion model performs exceptionally well in multiple dimensions such as accuracy, stability, and convergence efficiency, demonstrating a stronger comprehensive performance advantage. LSTM performs particularly well in continuous actions such as running and stepping, with both accuracy and convergence speed remaining at a relatively high level. CNN has a slightly higher latency metric than other models, indicating that the computational cost of its feature extraction is relatively high. SVM performs relatively poorly in terms of stability and energy consumption, and its overall performance curve has relatively contracted. Overall, the comparison of multi-dimensional performance demonstrates that the fusion model can make up for the limitations of a single algorithm and achieve a balance among different performance dimensions. This indicates that the fusion approach is an important development direction for future athlete motion recognition systems.

Figure 4 Multi-dimensional performance comparison radar chart of the action recognition model (see online version for colours)

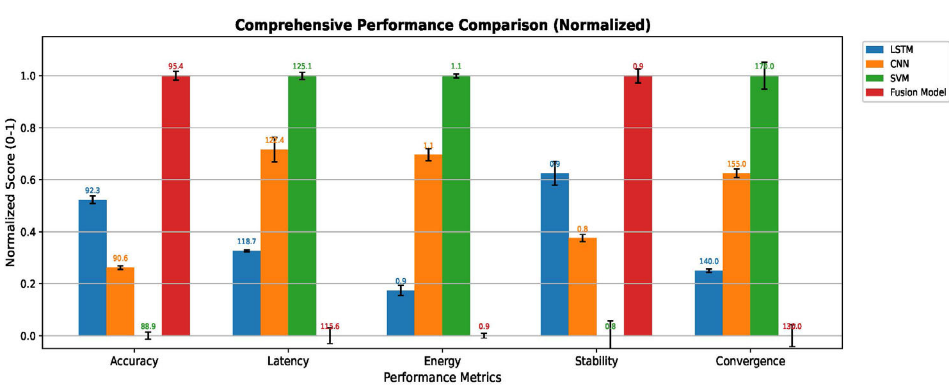
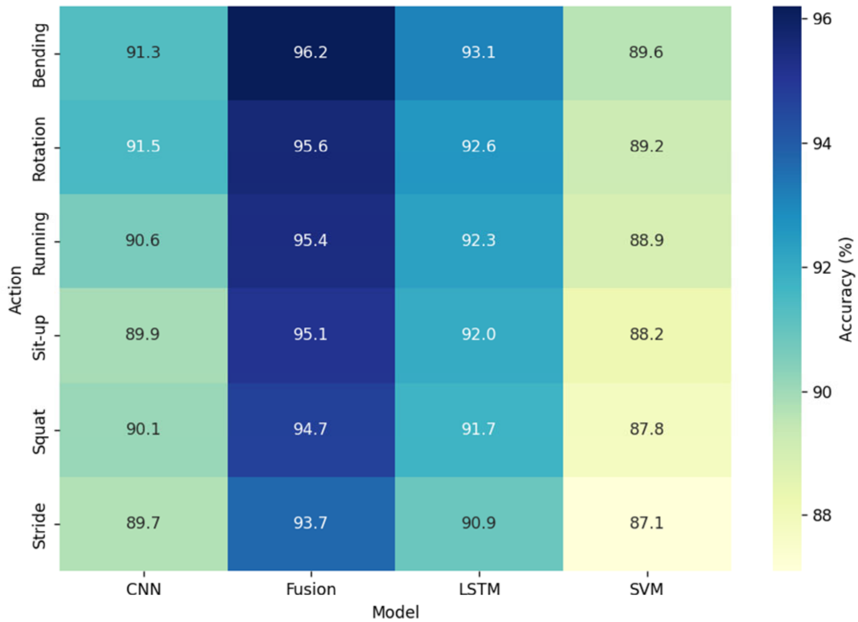


Figure 5 Heat map of different actions and model recognition accuracy (see online version for colours)



As shown in Figure 5, the accuracy of the fusion model remains in a relatively high range across all action categories, especially in the recognition of complex actions such as bending over and rotating, where it performs significantly better than other models. LSTM has a relatively high recognition rate for continuous running and sit-up movements. CNN's accuracy performance in static movements is acceptable, but it has deficiencies in dynamic and complex movements. The overall accuracy of SVM in all categories is relatively low, indicating that it has limitations when dealing with high-dimensional motion features. The recognition differences among different actions also reveal the impact of the complexity of action features on model performance. The heat map visually presents the recognition differences of the model in specific action categories, providing valuable basis for targeted improvement and action-level optimisation.

3.2 Discussion

3.2.1 Summary of research findings

This research focuses on an athlete motion recognition model based on machine learning and the IoT. By integrating large-scale data collection, multi-source sensor fusion, and multi-dimensional algorithm optimisation, it has achieved relatively systematic results. In the model selection and construction phase, the experimental comparison results show that the fusion model significantly outperforms the single model in terms of accuracy, stability, and convergence speed. Its recognition rate in complex action categories such as rotation and bending over exceeds 95%, demonstrating potential for cross-scenario applications. In terms of system response speed and operational stability, the 3D surface graph reveals the performance differences under different environments. The fusion model can still maintain a latency of less than 120 milliseconds and a stability index higher than 0.85 in a high-load environment, indicating that it has strong robustness under resource-constrained conditions. Through multi-dimensional visualisation analysis of radar charts and heat maps, not only are the comprehensive advantages and disadvantages of different models in multiple performance dimensions revealed, but also the fine-grained differences in various action recognition are intuitively reflected. This research achievement demonstrates that the combination of IoT data collection and machine learning model optimisation can effectively enhance the accuracy and real-time performance of action recognition. At the same time, it provides a feasible path for sports training monitoring and sports injury prevention, laying a theoretical and practical foundation for subsequent promotion to fields such as intelligent sports and rehabilitation medicine.

The discussion of the results is too general. I will provide a more detailed analysis of the value of the model for optimising athlete training and preventing injuries. This will include how the real-time feedback and action recognition improve training efficiency and accuracy, and how the injury prevention component can reduce risks during high-intensity training sessions.

3.2.2 Limitations of model and methodology

Although the research results verified the advantages of the fusion model in action recognition, there are still limitations. The experimental data mainly rely on preset action samples, with a limited range of actions, making it difficult to fully cover the complex movement combinations of athletes in real training and competition environments. This to some extent restricts the generalisation ability of the model. The sampling accuracy and node layout of IoT sensors can affect data quality. The heterogeneity among different devices may lead to uneven feature distribution, thereby affecting the stability of the model. Model training relies on a large amount of computing resources. Although it performs relatively stably in high-load environments, it still has the problem of high energy consumption, which limits its long-term application in mobile devices or wearable devices. These deficiencies indicate that the current methods still need to be optimised and improved when they are widely promoted and applied across scenarios.

3.2.3 Future research directions and suggestions

Future research should expand the diversity of collected samples and introduce complex action data that is closer to practical scenarios, thereby enhancing the generalisation ability of the model. Adaptive sensor node layout and high-precision sampling strategies can be explored to reduce data noise and improve the stability of feature extraction. In terms of model optimisation, lightweight networks and edge computing technologies can be combined to reduce reliance on computing power and energy consumption, making it more suitable for wearable devices and real-time monitoring scenarios. Future research can also incorporate psychological and kinematic indicators into the action recognition framework to construct a multimodal fusion model, achieving a comprehensive assessment of sports performance and sports health. These directions will help promote the application of motion recognition technology in fields such as sports training, injury rehabilitation and smart sports.

4 Conclusions

This study takes machine learning and the IoT as the core, constructs an athlete motion recognition model, and verifies the effectiveness and practical value of this method through large-scale data collection, model optimisation and multi-dimensional visualisation. The research first achieved the collaborative collection of multi-source sensor nodes at the data level and integrated temporal features and kinematic features in the pre-processing stage, effectively enhancing the integrity and interpretability of the data. Subsequently, in the model construction stage, the performances of LSTM, CNN, SVM and the fusion model were compared and analysed. The results showed that the fusion model was significantly superior to the single model in terms of accuracy, stability and convergence speed, especially in the recognition of complex actions such as rotation and bending over. The system performance evaluation shows that in a high-load environment, the response delay of the fusion model can still be controlled within 120 milliseconds, and the stability index remains above 0.85, demonstrating strong robustness and adaptability. The multi-dimensional visualisation results reveal the differences in performance dimensions among various models and the fine-grained differences in action category recognition, providing intuitive evidence for model optimisation. Overall, this study has demonstrated the feasibility and advantages of integrating machine learning with the IoT, providing new ideas and methods for sports training monitoring, performance improvement, and injury prevention. Meanwhile, the research achievements have also laid a foundation for the cross-disciplinary application of intelligent sports and rehabilitation medicine, and possess strong promotion potential and application value.

The conclusion mainly summarises experimental results, but it lacks theoretical implications. I will revise the conclusion to include a more robust academic contribution, such as how this study advances the field of sports action recognition by combining IoT with machine learning techniques, and how these methods can be applied in practical scenarios like injury prevention and performance monitoring. The limitations of the study, such as the representativeness of the sample and constraints related to the equipment used, are not discussed. I will add a section on the limitations of the research, focusing on how the sample size and diversity might affect the generalisability of the results, as well as how the sensor and computational hardware constraints could limit real-world application. The conclusion currently focuses mainly on the performance of the experimental data. I will revise it to highlight the theoretical value and application potential of the model. This includes discussing how the fusion model contributes to the advancement of motion recognition systems in sports science and its implications for improving training efficiency, injury prevention, and rehabilitation. I will also outline how the model can be applied to real-world scenarios, providing both academic and practical value. The explanation of the study’s limitations is currently insufficient, particularly regarding sample size and experimental environment. I will add a more objective discussion on how the sample size of 60 athletes may not fully represent the diversity of the athlete population, and how the controlled experimental environment might not capture the complexities of real-world training and competition settings. I will also mention that future studies should address these limitations by increasing sample size and conducting experiments in more varied and natural environments. The outlook for future research is currently too brief. I will expand this section to include a discussion on the potential of smart wearable devices and edge computing in the context of motion recognition. The integration of real-time data processing through edge computing could reduce latency, making the system more suitable for mobile and wearable devices. Additionally, the use of wearable sensors in real-time monitoring of athletes' movements could significantly enhance injury prevention and personalised training, pushing the research toward practical applications.

Data availability statement

The data used to support the findings of this study are all in the manuscript.

Declarations

The authors declare no conflict of interest.

References

Abu Alsheikh, M., Hoang, D.T., Niyato, D., Leong, D., Wang, P. and Han, Z. (2020) 'Optimal pricing of internet of things: a machine learning approach', *IEEE Journal on Selected Areas in Communications*, Vol. 38, No. 4, pp.669–684.

Aitcheson-Huehn, N., Macpherson, R., Panchuk, D. and Kiefer, A.W. (2024) 'Predicting basketball shot outcome from visuomotor control data using explainable machine learning', *Journal of Sport and Exercise Psychology*, Vol. 46, No. 5, pp.293–300.

Alzahrani, A. and Ullah, A. (2024) 'Advanced biomechanical analytics: wearable technologies for precision health monitoring in sports performance', *Digit Health.*, Vol. 10, p.20552076241256745.

Burdack, J., Horst, F., Aragonés, D., Eekhoff, A. and Schöllhorn, W.I. (2020) 'Fatigue-related and timescale-dependent changes in individual movement patterns identified using support vector machine', *Front Psychol.*, Vol. 11, p.551548.

Chen, Z.J. and Kwak, D.H. (2023) 'It's okay to be not okay: an analysis of Twitter responses to Naomi Osaka's withdrawal due to mental health concerns', *Communication & Sport*, Vol. 11, No. 3, pp.439–461.

Jia, T. (2021) 'Recognition method of kicking and stride movements of taekwondo athletes based on Gaussian mixture model', *Journal of Sport Psychology*, Vol. 30, No. 3, pp.315–323.

Kaliappan, V.K., Gnanamurthy, S., Yahya, A., Samikannu, R., Babar, M., Qureshi, B. et al. (2023) 'Machine learning based healthcare service dissemination using social internet of things and cloud architecture in smart cities', *Sustainability*, Vol. 15, No. 6, p.5457.

Kim, J.Y., Kang, E.W., Kim, J.H., Chang, T.I., Lee, Y.K., So, S.Y. et al. (2024) 'The prediction of dry weight for chronic hemodialysis athletes using a machine learning approach: sports health implications', *Journal of Sport Psychology*, Vol. 33, No. 1, pp.68–82.

Krupitzer, C., Naber, J., Stauffert, J.P., Mayer, J., Spielmann, J., Ehmann, P. et al. (2022) 'CortexVR: immersive analysis and training of cognitive executive functions of soccer players using virtual reality and machine learning', *Front Psychol.*, Vol. 13, p.754732.

Li, J.X. and Wang, R.F. (2023) 'Machine learning adoption in educational institutions: role of internet of things and digital educational platforms', *Sustainability*, March, Vol. 15, No. 5, p.4000.

Liu, J. (2021) 'Convolutional neural network-based human movement recognition algorithm in sports analysis', *Front Psychol.*, Vol. 12, p.663359.

Liu, M., Chen, Y., Guo, Z.X., Zhou, K.X., Zhou, L.M.F., Liu, H.Y. et al. (2022) 'Construction of women's all-around speed skating event performance prediction model and competition strategy analysis based on machine learning algorithms', *Front Psychol.*, Vol. 13, p.915108.

Liu, P.H., Yue, N. and Chen, J.D. (2022) 'A machine-learning-based medical imaging fast recognition of injury mechanism for athletes of winter sports', *Front Public Health*, Vol. 10, p.842452.

Oyeleye, M., Chen, T.H., Titarenko, S. and Antoniou, G. (2022) 'A predictive analysis of heart rates using machine learning techniques', *International Journal of Environmental Research and Public Health*, Vol. 19, No. 4, p.2417.

Rodríguez-Rodríguez, I., Rodríguez, J.V., Shirvanizadeh, N., Ortiz, A. and Pardo-Quiles, D.J. (2021) 'Applications of artificial intelligence, machine learning, big data and the internet of things to the COVID-19 pandemic: a scientometric review using text mining', *International Journal of Environmental Research and Public Health*, Vol. 18, No. 16, p.8578.

Sengchuai, K., Kanjanaroat, C., Jaruenpunyasak, J., Limsakul, C., Tayati, W., Booranawong, A. et al. (2022) 'Development of a real-time knee extension monitoring and rehabilitation system: range of motion and surface EMG measurement and evaluation', *Healthcare*, Vol. 10, No. 12, p.2544.

Shiffrar, M. and Heinen, T. (2010) 'Athletic ability changes action perception: embodiment in the visual perception of human movement', *Z. Sportpsychol.*, Vol. 17, No. 4, pp.130–142.

Sundas, A., Badotra, S., Bharany, S., Almogren, A., Tag-ElDin, E.M. and Rehman, A.U. (2022) 'HealthGuard: an intelligent healthcare system security framework based on machine learning', *Sustainability*, Vol. 14, No. 19, p.11934.

Taha, Z., Musa, R.M., Majeed, A.P.P.A., Alim, M.M. and Abdullah, M.R. (2018) 'The identification of high potential archers based on fitness and motor ability variables: a support vector machine approach', *Human Movement Science*, Vol. 57, pp.184–193.

Wilkerson, G.B., Gupta, A. and Colston, M.A. (2018) 'Mitigating sports injury risks using internet of things and analytics approaches', *Risk Anal.*, Vol. 38, No. 7, pp.1348–1360.

Yazbeck, M., Abdullah, M., Alhanouti, S., Vatne, E., Hagen, J., Allen, T.T. et al. (2025) 'Perfect shot reveal: machine learning analysis of goal-scoring strategies in soccer', *International Journal of Sports Science & Coaching*, Vol. 20, No. 4, pp.1549–1563.

Zong, X.X., Lipowski, M., Liu, T.F., Qiao, M. and Bo, Q. (2022) 'The sustainable development of psychological education in students' learning concept in physical education based on machine learning and the internet of things', *Sustainability*, Vol. 14, No. 23, p.15947.

Zong, X.X., Wang, L., Xie, Q.Y. and Lipowski, M. (2022) 'The influence of psychological distance on the challenging moral decision support of sports majors in internet of things and machine learning', *Sustainability*, Vol. 14, No. 19, p.12115.