Wang D et al. (2025) ENHANCING PHYSICAL EDUCATION MOVEMENT PRECISION THROUGH AI-DRIVEN DEEP LEARNING CALIBRATION SYSTEMS. Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte vol. 25 (99) pp. 452-469.
DOI: https://doi.org/10.15366/rimcafd2025.99.029

ORIGINAL

ENHANCING PHYSICAL EDUCATION MOVEMENT PRECISION THROUGH AI-DRIVEN DEEP LEARNING CALIBRATION SYSTEMS

Qiang Yin[1], Dong Wang[2]*, Lin Zhang[3]

[1] Public Course Department, Loudi Vocational and Technical College, Hunan, China.
[2] Physical Education Department, Hunan railway vocational and technical college, Hunan, China.
[3] Teaching-research Office, School of Basic Education, Yiyang Normal College, Yiyang, Hunan, China.
E-mail: 447417267@qq.com

Recibido 15 de Marzo de 2024 **Received** March 15, 2024
Aceptado 03 de Octubre de 2024 **Accepted** October 03, 2024

ABSTRACT

The aim of this study is to enhance the efficacy of sports teaching movements and to promptly correct erroneous forms, by integrating artificial intelligence (AI) and deep learning technologies into the recognition of sports movements. This paper commences by computing the correlation matrix for a set of selected features, subsequently establishing a threshold to eliminate features with high cross-correlation, thereby reducing redundancy and optimizing the feature set. To preprocess the imagery, a Gaussian function is initially applied to perform convolution operations. Subsequently, a Gaussian kernel function is utilized to filter the images, constructing a hierarchical structure known as the Gaussian pyramid, wherein variable Gaussian filter coefficients are employed at each level of image processing. Ultimately, this research develops a precise calibration system for physical education movements and implements it within the context of physical education to enhance teaching outcomes. The experimental results demonstrate that the system developed in this study effectively satisfies the practical requirements of physical education.

KEYWORDS: Artificial Intelligence; Deep Learning; Physical Education; Movement; Precise School Position.

1. INTRODUCTION

The general steps of video behaviour recognition include three main

parts: video preprocessing, feature extraction and feature classification. In the video preprocessing part, it is usually necessary to separate the actors making movements in each frame from the background environment (Ferguson et al., 2019). The foreground of each frame is extracted through the Gaussian mixture model for background modelling. After extracting the foreground of each frame, we must consider how to express the features of the video samples effectively. If the features abstracted from the video samples can distinguish the characteristics of different behaviors well, then the classification will have a multiplier effect. A video itself is a high-dimensional data related to time, space, and frequency (color). It is a complicated task to directly obtain useful features from it. Using MHI to express the movements in the sample video in a two-dimensional form is an effective method (Kimasi et al., 2019).Although acquiring MHI has high requirements on the video background, considering that most of the surveillance camera equipment is in a static state, the extracted video background is relatively single. Therefore, MHI still has stricter application conditions. After converting the video samples into MHI, image recognition and classification techniques can be used for behavior recognition. In the classifier stage, this paper will mainly test k-nearest neighbors, SVM and nuclear non-linear classification methods (Reinhart & Wichmann, 2020). Upon securing a video sample, the initial analytical step involves the extraction of dynamic subjects from the static background, a process referred to as foreground detection, which is essential for subsequent action feature extraction and classification. In scenarios where surveillance is conducted using a stationary camera, the backdrop is relatively invariant and can be predefined. The moving targets within the frame, such as pedestrians or vehicles, constitute the primary area of interest. To segment these foreground objects from the background, background modelling is conducted, followed by a comparison between the current frame and the established background model to identify moving objects. Foreground extraction is a fundamental component of intelligent surveillance systems. Various techniques are employed for foreground detection, including the frame difference method, median filter method, optical flow method, background statistical model method, and the codebook model. Foreground detection is applicable in two primary environments: with a stationary camera and with a moving camera. When the camera is in motion, the optical flow method is often the preferred approach for foreground detection. This method calculates the optical flow field of the image sequence to infer the camera's motion state. However, due to the computational demands of the optical flow method, real-time processing is challenging. In the majority of surveillance settings, the camera remains stationary, eliminating the need for extensive optical flow calculations. The Gaussian Mixture Model (GMM), one of the background statistical model methods employed in this study, is particularly suited for distinguishing between the background and foreground in video sequences captured by a stationary camera. This method provides an effective approach for background subtraction in static camera surveillance scenarios.

This paper applies artificial intelligence technology and deep learning technology to sports movement recognition, and combines intelligent technology to perform sports teaching auxiliary work, improve the accuracy of sports movement recognition, and lay the foundation for subsequent smart sports teaching.

2. Related Work

In the past ten years, great progress has been made in the field of movement recognition. Various movement recognition systems have emerged and many of them have been put into practical use. These systems have penetrated into all areas of people's lives. It is generally divided into single-sensor-based research (Abanazir, 2019) and multi-sensor-based data fusion (Gerke et al., 2018).Literature (Pogrebnoy & Komlev, 2018) proposes a method to recognize simple daily human activities in an environment that is not affected by location by using the built-in sensor of a smartphone. This algorithm can solve the problem that users can freely switch the position of the smartphone, and improve the recognition rate by using a combination of different sensors and a deep neural network to classify and remove transitional data. The final recognition rate can reach 93%. The literature (Ilies et al., 2018) developed a smart home system based on multi-sensor data fusion technology and its smart monitoring interface. The system uses a number of different types of sensors (inertial sensors, temperature sensors, light sensors) to realize human motion recognition, fire detection, and temperature monitoring in smart homes. The literature (Kondrukh, 2017) used accelerometers, gyroscopes and magnetic sensors for data complementary fusion, greatly improving human body gestures recognition. The literature (Sá et al., 2015) proposed real-time processing of data to achieve real-time dynamic detection of daily behaviours for the battery life and memory limitation of smart phoneature (Komisar et al., 2017), solving two major problems that affect the recognition rate: zero drift and sensor orientation change. In the recognition system based on inertial sensors, the change of the wearing direction of the sensor will directly affect the recognition accuracy of the system. It proposes a method that combines the constant gravity in the vertical direction and the static state value, which greatly reduces the error caused by the change of the sensor wearing direction and improves the recognition rate. The literature (Sá et al., 2015) proposed real-time processing of data to achieve real-time dynamic detection of daily behaviours for the battery life and memory limitation of smartphone terminals. Moreover, it used sensors such as the accelerometer, gyroscope and magnetic field meter of the smartphone to perform data fusion at the feature level, and adopted a low-complexity classification algorithm to achieve a high recognition rate. In addition, it verified that the combination of accelerometer and gyroscope shows the best fusion effect through a combination of multiple sensors and showed that the increase in frequency domain features does not greatly improve the recognition rate but will occupy a large amount of memory. The

study referenced as (Ángel-López & de la Peńa, 2015) integrates information fusion techniques into context-aware systems, proposing a novel multi-sensor fusion approach that is predicated on the time evidence theory for the purpose of recognizing indoor activities. This approach extends the traditional Dempster-Shafer (D-S) evidence theory by incorporating a temporal dimension, which is instrumental in enhancing the accuracy of activity recognition within a multi-sensory setup. The method is designed to incrementally resolve conflicts and is particularly effective in applications aimed at aiding daily life, such as life assistance systems. In literature (Kim et al., 2015), accelerometers and gyroscopes in wearable devices are used to recognize daily behaviors such as standing, walking, running, jumping, up and down stairs, etc. The recognition rate can reach about 85%. In addition, the multi-sensor fusion system occupies an important position in the field of action recognition due to its robustness and stability. Literature (Zhi-Min & Zhong-Wen, 2015) improves the system's robustness by fusing multi-source heterogeneous sensor data, and applies fuzzy logic information fusion algorithms to In the fusion architecture, human behavior is recognized, and feature layer fusion is performed to improve the recognition rate. However, the disadvantage of these algorithms is that the higher feature dimensionality leads to the increase of algorithm complexity. Therefore, the use of principal component analysis, linear discriminant analysis and other algorithms to reduce the feature dimensionality is a common method in practical applications (Khan, 2017). There have been some achievements in domestic research on action recognition, but there is still room for innovation in feature extraction and the establishment of classification models. Literature (Puupponen et al., 2015) explored the contribution of wrist-worn multi-sensor for elderly activity recognition, and found that the combination of accelerometer sensor and heart rate data can be used to improve the classification and recognition of elderly activity. Here, a genetic algorithm-based approach is proposed. The method of fusion weight selection uses genetic algorithm to find the fusion weight, and the fusion result can reach about 98% of the recognition rate of the simple daily actions of the elderly. Literature (Lee & Yoo, 2017) recognizes the actions when playing table tennis, such as rubbing the ball, attacking the ball, and pushing the gear. The fixed threshold is used to determine the start and end points of the acceleration signal and extract the signal's characteristics through the processing of the three-axis acceleration signal. The classification algorithm of K nearest neighbors and decision tree is used to recognize it. The result shows that the action when hitting a table tennis can be recognized quickly and effectively, but the disadvantage is that there may be redundant features, so the feature selection algorithm is also the direction of improvement in the future. All in all, in the past ten years, the research on action recognition at home and abroad has made great progress, but it is difficult to achieve accurate recognition of complex human behaviors, and most systems cannot perform real-time and effective active monitoring and the algorithm complexity is relatively high. Therefore, the balance between

recognition accuracy and algorithm complexity remains to be studied.

3. The Application of Artificial Intelligence Deep Learning Technology in Sports Teaching Movement Recognition

Generally, the radial basis function is defined as $\Phi(x,y) = \phi(\|x-y\|)$, where $\|x\|$ represents the Euclidean norm. A necessary condition of the radial basis function is: when $\|x_1\| = \|x_2\|$, there must be $\Phi(x_1) = \Phi(x_2)$. Among the common radial basis functions, the Gaussian type is the most commonly used(Zhang et al., 2020).

The Gaussian distribution function of the Kriging method is $\phi(r) = e^{-\frac{r^2}{\sigma^2}}$.

Hardy's M-Q inverse function is $\phi(r) = (c^2 + r^3)^{\beta}$ or $\phi(r) = (c^2 + r^3)^{-\beta}$.

Duchon's thin plate spline function is $\phi(r) = r^{2k}\, ln(r)$ or $\phi(r) = r^{2k+1}$.

The above-mentioned radial basis functions all have the characteristic of radial symmetry, and when the function deviates from the center of the function, the value of the function decreases very quickly, and the rate of decrease is proportional to the strength of the selectivity(Liu & Hodgins, 2018). In practical applications, RBF networks have two structures: regularized networks and generalized networks. The generalized network can relax this condition to as few samples as possible. Regularized network is a universal ability (Li et al., 2021). As long as the training sample is large enough, in theory, any multivariate continuous function can be approximated with arbitrary precision, and for any unknown nonlinear function, there is always a set of coefficients to make the network approach the most Optimization, although this makes its approximation performance theoretically guaranteed. The network structure is shown in Figure 1.

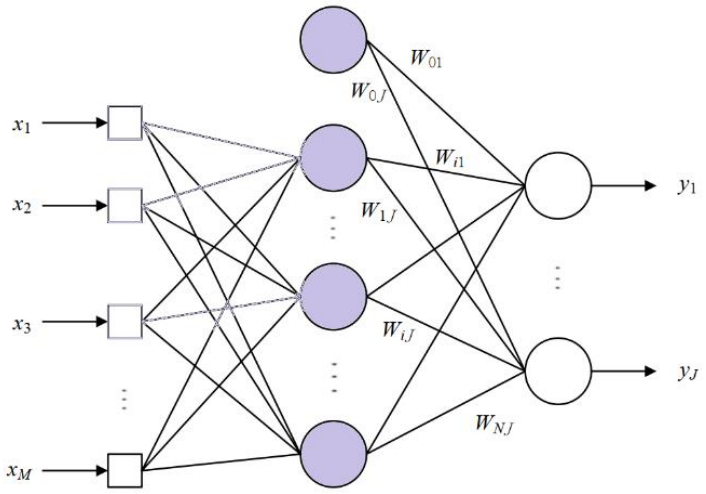

Figure 1: Generalized RBF network structure

Here, we assume that the generalized network has M input nodes, and the number of hidden layer nodes is I. The characteristic that is different from the regularized network is that I<M, the basis function of the i-th node in the hidden layer is $\Phi(\|X - X_i\|)$, the center of the basis function is $X_i = [x_{i1}, x_{i2}, \ldots, x_{im}]$, the number of neurons in the output layer is J, and the threshold Φ_0 is added here.Its output is 1, and the connection weight is ω_{oj}. If the actual output is set as $Y_k = [y_{k1}, y_{22}, \ldots, y_{kj}, \ldots, y_{kJ}]$ and J is the number of output units, when the training sample X_k is input, the output of the j-th neuron in the output layer is as shown in formula 1(Liu et al., 2020).

$$y_{kj} = \omega_{oj} + \sum_{i=1}^{I} \omega_{ij} \phi(X_k, X_i) . j = 1,2, \ldots J \quad (1)$$

If the Gaussian radial basis function is used, then

$$\phi(X_k, X_i) = G(X_k, X_i) = G(\|X_k - X_i\|) \quad (2)$$

$$G(\|X_k - X_i\|) = exp\left(-\frac{1}{2\sigma^2}\|X_k - X_i\|\right) \quad (3)$$

RBF learning is different from a general neural network. It is by continuously adding neurons to the hidden layer until the setting error is reached. This saves time and converges faster than a general neural network using all samples for training.

(1) Randomly select a fixed center: The rudimentary learning strategy entails the selection of a fixed center, with the primary objective being the establishment of connections between the hidden and output layers through weighted associations. This strategy employs the standard deviation and center of a predefined basis function. The center of the hidden layer's basis function is randomly selected from the input training dataset and is thereafter considered constant. Once the center is ascertained, the standard deviation of the basis function is derived as delineated in the following equation (Wang et al., 2017):

$$\sigma = \frac{d_{max}}{\sqrt{2n}} \phi(r) = e^{-\frac{r^2}{\sigma^2}} \quad (4)$$

When the center and standard deviation of the basis function are determined, the basis function can be expressed as:

$$\phi\left(\left\|X_k - X_i\right\|\right) = \exp\left(-\frac{1}{d_{\max}^{\ 2}}\left\|X_k - X_i\right\|^2\right), i = 1,2,\ldots,n \quad (5)$$

When the basis function is obtained, the output weight ω is calculated by the pseudo-inverse method. Here, we assume that d_{kj} is the expected

output of the j-th output node for the k-th input vector, and ω_{ij} is the connection weight from the i-th node in the hidden layer to the j-th node in the output layer. Then, the expected output weight matrix can be expressed as (Bao et al., 2020):

$$\omega = G^{+}d \tag{6}$$

Among them, $G = \{g_{ki}\}$, and $\omega = \omega_{ij}$, and G^{+} represents the pseudo-inverse matrix of G, which can be solved by singular value decomposition. The specific calculation of g_{ki} is as follows:

$$g_{ki} = \phi(\|X_k - X_i\|^2), k = 1,2,\ldots,I \tag{7}$$

(2) Orthogonal least squares method: Orthogonal Least Square (OLS) is an important method of learning weights for RBF networks. First, when it assumes that the output layer has only one node, it compares the basis function network to a special case of linear regression:

$$d(n) = \sum_{i=1}^{I} p_i\omega_i + e(n), n = 1,2,\ldots,N \tag{8}$$

Among them, the number of hidden layer nodes is I, the number of input training samples is N, and ω_i is the weight from the i-th node of the hidden layer to the output node (here it is assumed that there is only one output node). d(n) is the expected output value of the network, e(n) is the network error, and $p_i(n)$ is the model regression factor. When the Gaussian function is selected as the basis function, $p_i(n)$ is expressed as follows:

$$p_i(n) = exp\left(-\frac{1}{2\sigma^2}\|X_n - t_i\|^2\right) \tag{9}$$

It is expressed in matrix form as:

$$\left.\begin{array}{l} d = Pw + e \\ d = [d_1, d_2, \ldots, d_K]^T \\ w = [w_1, w_2, \ldots, w_I]^T \\ P = [p_1, p_2, \ldots, p_I]^T \\ p_i = [p_{i1}, p_{i2}, \ldots, p_{iK}]^T \\ e = [e_1, e_2, \ldots, e_K]^T \end{array}\right\} \tag{10}$$

Here, $d = Pw + e$ can be solved by orthogonal triangular decomposition matrix P.

$$P = UA \tag{11}$$

Then, for the matrix U, the diagonal matrix H can be obtained by formula (12).

$$U^T U = H \tag{12}$$

Through the above derivation, we can get:

$$d = Pw + e = UAw + e = Ug \tag{13}$$

When both sides are multiplied by matrix U^T at the same time, we can get:

$$U^T d = U^T U g = Hg \tag{14}$$

Then, by continuing to solve, we can get:

$$g = H^{-1} U^T d \tag{15}$$

So far, according to the principle of least squares method, Aw = g, A and g have already been obtained, and then the connection weight w can be calculated.

(3) Supervised selection center: This learning strategy determines cluster centers and other parameters through supervised learning. In the supervised selection center method, the cost function is defined as(Bhat et al., 2021):

$$E = \frac{1}{2}\sum_{K=1}^{N} e_K^2 \tag{16}$$

Among them, E is the output error of a certain node, N is the number of training samples, and e_j is the error between the output result of the j-th input training data through the network and the actual expected result.

$$e_k = d_k - \sum_{i=1}^{I} \omega_i G\left(\|X_k - t_i\|_{C_i}\right) \tag{17}$$

I is the number of hidden layer nodes in the network structure. In the training and learning process, the gradient descent method is used to find the free parameters t_i, ω_i and Σ_i^{-1} (related to C_i, denoted by S_i in the following derivation) to minimize the cost function, then the network parameter optimization formula can be derived as follows:

a) The output weight is ω_i,

$$\frac{\partial E(n)}{\partial \omega_i(n)} = \sum_{k=1}^{N} e_k(n) G\left(\|X_k - t_i\|_{C_i}\right) \tag{18}$$

$$\omega(n+1) = \omega_i(n) - \eta_1 \frac{\partial E(n)}{\partial \omega_i(n)}, i = 1,2,\ldots,I \tag{19}$$

b) The center of the hidden layer is t_i,

$$\frac{\partial E(n)}{\partial \omega_i(n)} = 2\omega_i(n)\sum_{j=1}^{N} e_k(n)G'\left(\|X_k - t_i\|_{C_i} S_i\left(X_k - t_i(n)\right)\right) \quad (20)$$

$$t_i(n+1) = t_i(n) - \eta_2 \frac{\partial E(n)}{\partial t_i(n)}, i = 1,2,\ldots,M \quad (21)$$

c) The center extension of the hidden layer is S,

$$\frac{\partial E(n)}{\partial S_i(n)} = -\omega_i(n)\sum_{j=1}^{N} e_k(n)G'\left(\|X_k - t_i\|_{C_i}\right) Q_{ki}(n) \quad (22)$$

$$Q_{ki}(n) = \left(X_k - t_i(n)\right)\left(X_k - t_i(n)\right)^T \quad (23)$$

$$S_i(n+1) = S_i(n) - \eta_3 \frac{\partial E(n)}{\partial S_i(n)}, i = 1,2,..,M \quad (24)$$

In addition, the learning rate η_1, η_2, η_3 should take different values. The Relief algorithm belongs to the Filter-style feature selection algorithm. The extracted features are selected without building a classifier. The basic principle of feature selection is to retain or remove features according to the feature's ability to distinguish samples. Relief algorithm has low complexity and high efficiency, so it is widely used in feature selection and data classification, but it is limited to the feature selection and classification problems of two types of data (Sun et al., 2018). Relief-F solves the limitation of Relief on the number of action categories, changing the number of neighbor selections for sample R to K neighboring samples, and then repeatedly updating the feature weights, thus playing a huge role in the feature selection of multi-type label data. Aiming at this limitation, this paper improves the Relief-F feature selection algorithm. This paper calculates the correlation matrix of the selected features, and sets a threshold for the features with higher cross-correlation to remove them, so as to remove the redundancy between the selected features of the algorithm, and further achieve the optimization of the features. Its realization block diagram is shown as in Figure 2.

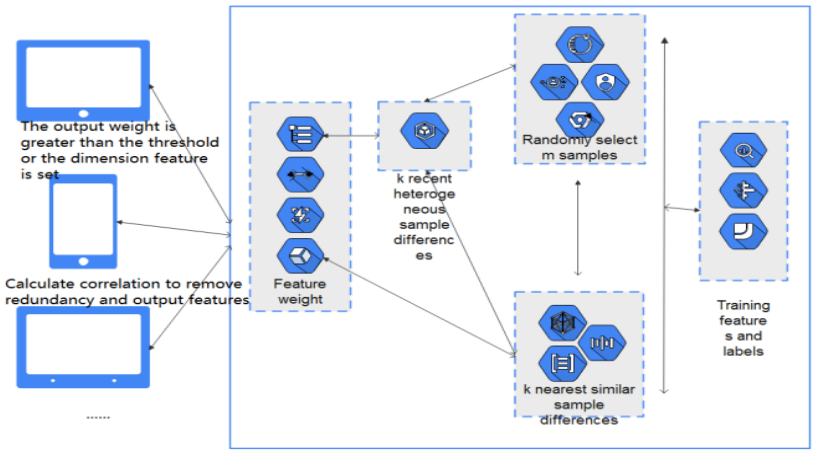

Figure 2: Relief-F improved feature selection

According to Figure 2, the Relief-F algorithm determines the weight of the feature by calculating the strength of the feature's ability to distinguish movements, and sets the threshold for output. The specific algorithm is described as follows:

(1) We assume that the training data set is D, the number of sample sampling rates is m, the number of neighboring samples is k, and the threshold is δ.We randomly select a sample R and look for K neighboring samples, and calculate the feature difference on a certain feature. The calculation formula is as follows.

$$diff(F, R_1, R_2) = \frac{|value(F,R_1) - value(F,R_2)|}{max(F) - min(F)} \tag{25}$$

$$diff(F, R_1, R_2) = \begin{cases} 0; value(F, R_1) = value(F, R_2) \\ 1; other \end{cases} \tag{26}$$

When the sample set is a numerical variable, formula (25) is used to calculate the characteristic difference, and when it is a nominal variable, formula (26) is used to calculate the characteristic difference.

(2) The probability of the class distribution is shown below, and P(c) represents the probability of the c-th target.

$$p = \frac{p(c)}{1 - p(class(R))} \tag{27}$$

(3) The weights of all features are updated cyclically, and the specific calculation formula is as follows.

$$W[F] = -\frac{\left(\sum_{i=1}^{k} diff(F,R,H_i)\right)}{km} + \frac{\sum_{c \neq class(R)} \left[p \times \left(\sum_{i=1}^{k} diff\left(F,R,M_j(c)\right)\right)\right]}{km} \tag{28}$$

In the above formula, $class(R)$ represents the category of R, $diff(F, R, H_i)$ represents the difference value of samples of the same type, and $diff(F, R, M_i(C))$ refers to the difference value of samples of different types.

(4) For the number of sampling m, the feature weight $W[F]$ is finally obtained through continuous loop calculation and update, and then the weight is sorted, and the threshold or specific dimension is set to output the selected feature subset.

(5) The features after the decision are calculated to remove the redundancy between similar features. If it is assumed that there are n output features of the judge, the correlation coefficient matrix between the n groups of features is calculated as A, and the Peel correlation coefficient between features X and Y is calculated as follows:

$$r = \frac{1}{N-1}\sum_{i=1}^{N}\left(\frac{X_i-\overline{X}}{\sigma_X}\right)\left(\frac{Y_i-\overline{Y}}{\sigma_Y}\right) \tag{29}$$

(6) According to the correlation coefficient calculation formula between features, the correlation coefficient matrix between all features is A.

$$A = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \vdots & r_{nn} \end{bmatrix} \tag{30}$$

The specific algorithm flow chart is shown in Figure 3.

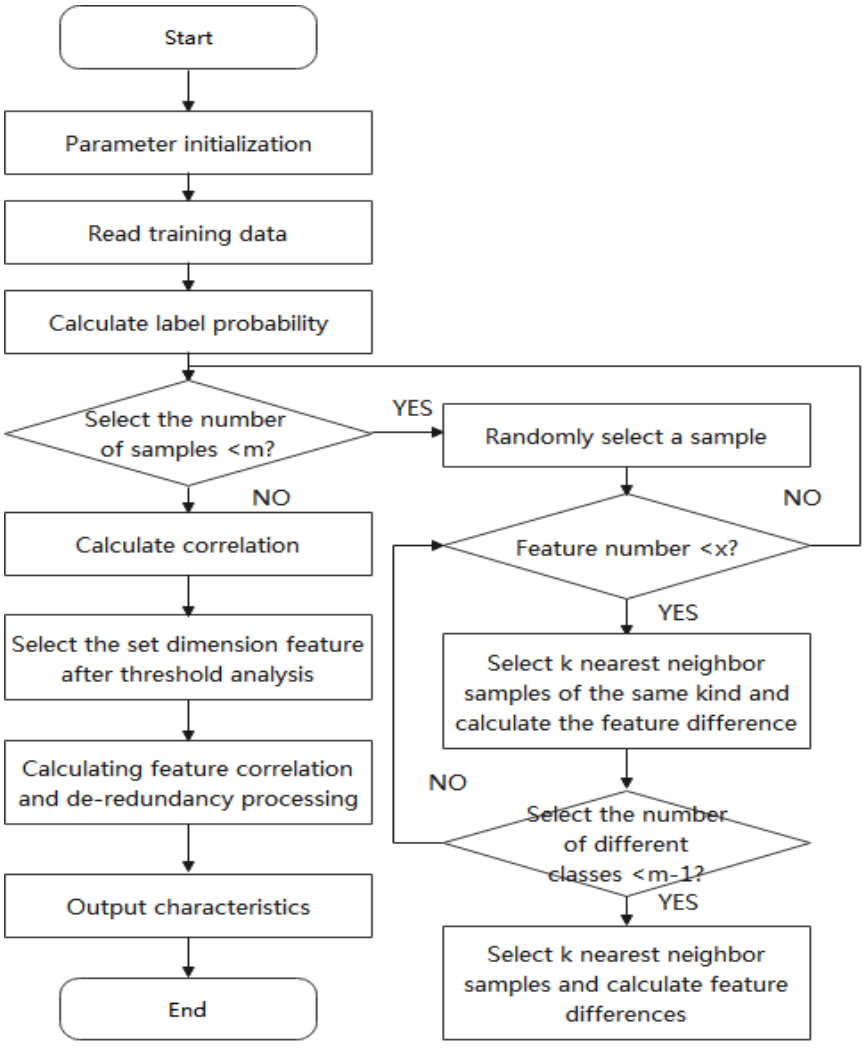

Figure 3: Flow chart of Relief-F improved algorithm

4. Refinement of Sports Teaching Movements through Deep Learning and Artificial Intelligence

In this section, we detail the process of refining sports teaching movements using advanced deep learning techniques and artificial intelligence.

The initial step involves image preprocessing to mitigate blurriness, which is crucial for the subsequent extraction of robust feature points. To achieve this, a Gaussian function is applied to the image through a convolution operation, serving as the initial preprocessing step. Subsequently, we employ a pyramid image hierarchical structure, as depicted in Figure 4. The process begins with the first sampling at an interval of 0.5 pixels, effectively doubling the resolution of the original image to produce the first layer of the pyramid. This initial sampled image serves as the base for subsequent layers. Progressively, the sampling intervals are doubled—1, 2, and 4 pixels—to generate the subsequent layers of the pyramid structure. These sampled images collectively form the pyramid hierarchy, which is then subjected to Gaussian filtering. This results in a Gaussian pyramid image structure, where each layer is filtered using a Gaussian kernel function with varying filter factors. The filter factors for the Gaussian filter at each level of the pyramid are indicated on the left side of the figure. Specifically, the standard deviation ratio between two consecutive layers is set to 1.26, ensuring a consistent and smooth transition across the pyramid layers.

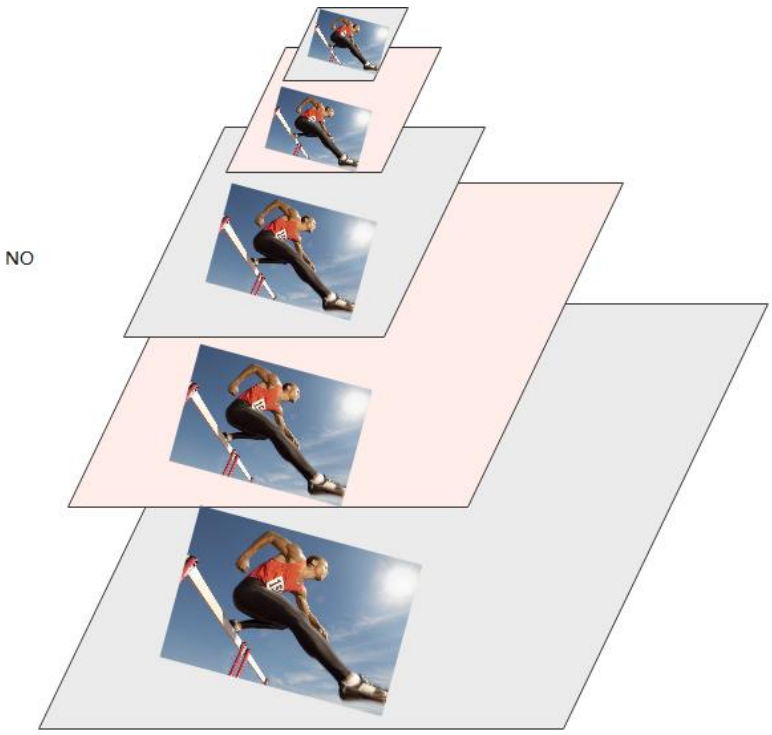

Figure 4: Pyramid image

In Figure 5, the first group of five-layer purple logo images on the right are Gaussian filtered images, and the second group of four-layer green logo images on the right are DOG images after Gaussian difference. The original image is convolved with a Gaussian function to generate a group of images. The right column in the figure shows the result of subtracting adjacent images in each group to generate a Gaussian difference image.

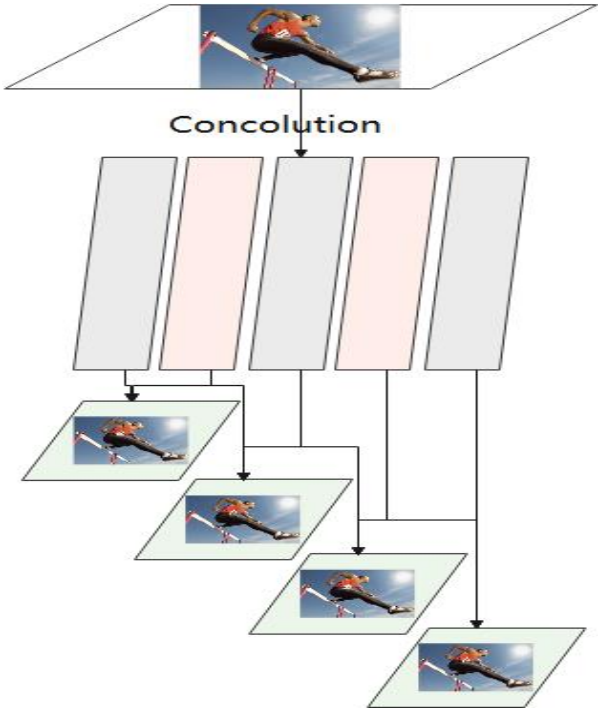

Figure 5: Schematic diagram of Gaussian difference image

The process of obtaining selected feature points commences with the extraction of extrema from the Difference of Gaussians (DOG) image. To identify these extreme points within the Gaussian difference image, each sample pixel is subjected to a comparison with its immediate neighbors, which consist of 8 pixels in the same image layer. Additionally, the pixel under consideration is compared with its 9 counterparts in the directly adjacent layers of the pyramidal image structure. This comprehensive comparison involves a total of 26 pixels, as illustrated in Figure 6.

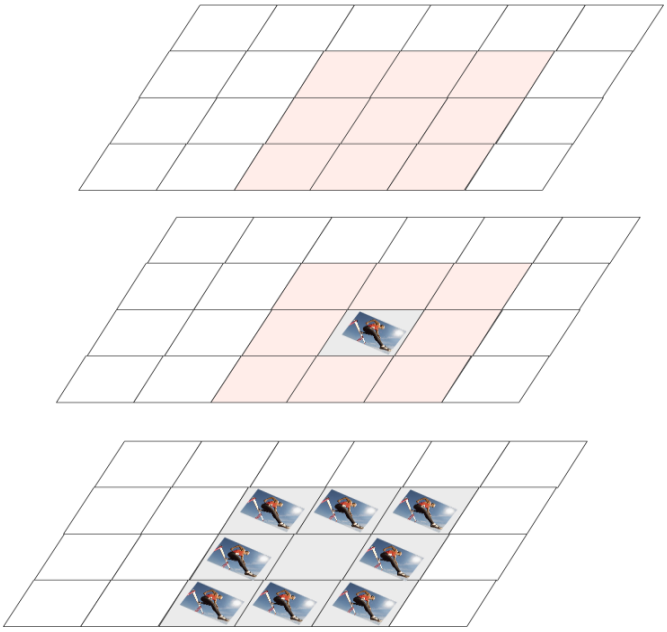

Figure 6: Schematic diagram of comparison of 26 neighborhoods

For the green image on the right side of Figure 7, three images are extracted in order to extract feature points according to the above method. Among the 4 images, the red star image is the intermediate image of 1, 2, 3 and 2, 3, 4 images compared with 26 pixels around the top and bottom. Therefore, we extract feature points from these two images and record grayscale information and coordinate information.

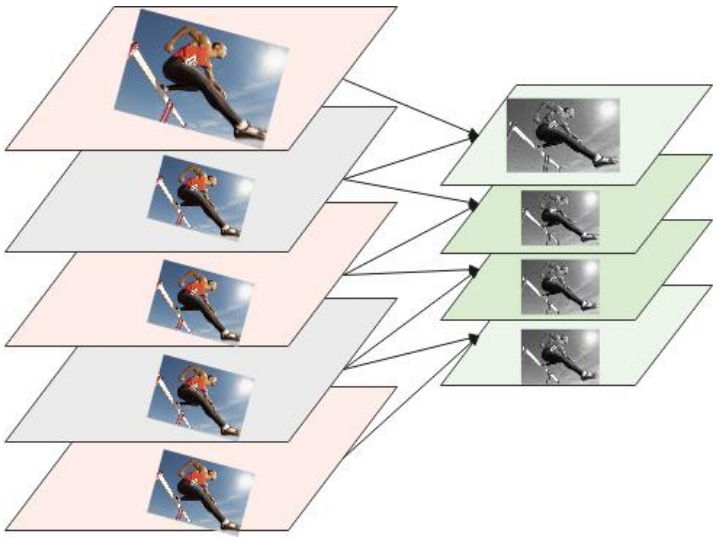

Figure 7: Schematic diagram of Gaussian difference

In order to realize the scalability and versatility of the system, this system is based on the idea of modularization and encapsulates important functional modules into different classes. Based on the above research, the structure diagram of the entire physical education movement recognition system is as follows:

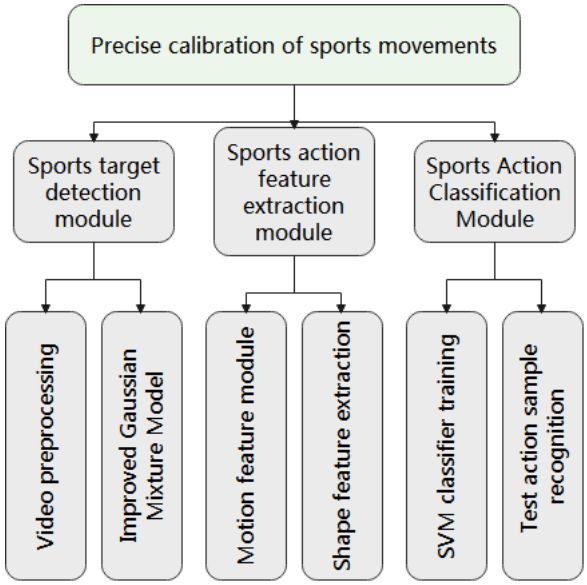

Figure 8: The structure diagram of the precise calibration system of physical teaching movements

5. Performance Verification of the Precise Calibration System for Physical Teaching Movements

This article combines artificial intelligence deep learning technology to construct a precise calibration system for sports teaching movements, and evaluates the effect of systematic teaching in combination with experimental teaching. Moreover, this paper verifies the precise calibration effect of physical teaching movements through experimental teaching methods, and obtains the results shown in Table 1 and Figure 9.

Table 1: Statistical table of the precise calibration effect of physical teaching movements

NO	SPORTS MOVEMENT CALIBRATION	NO	SPORTS MOVEMENT CALIBRATION	NO	SPORTS MOVEMENT CALIBRATION
1	92.48	11	90.38	21	94.78
2	93.99	12	88.45	22	91.79
3	90.86	13	90.35	23	89.81
4	95.15	14	90.73	24	87.93
5	90.54	15	93.33	25	93.30
6	87.73	16	91.30	26	90.86
7	95.72	17	93.26	27	93.09
8	91.97	18	89.45	28	87.48
9	90.56	19	90.15	29	92.54
10	93.50	20	90.03	30	90.51

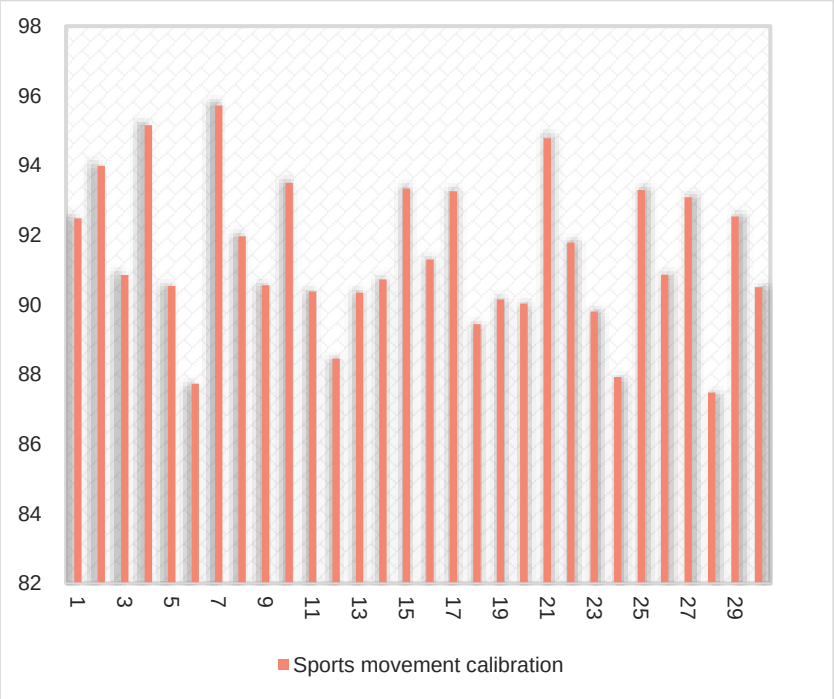

Figure 9: Statistical table of the precise calibration effect of physical teaching movements

From the above experimental research, it can be seen that the precise calibration system of sports teaching movements based on artificial intelligence deep learning technology proposed in this paper has a good recognition effect of sports teaching movements. After that, the effect of this system on the improvement of physical teaching is studied, and the results obtained are shown in Table 2.

Table 2: Evaluation of the improvement of teaching effect

NO	TEACHING EFFECT	NO	TEACHING EFFECT	NO	TEACHING EFFECT
1	84.61	11	91.05	21	92.03
2	88.80	12	83.49	22	89.69
3	93.60	13	78.90	23	75.93
4	96.53	14	78.12	24	88.62
5	86.39	15	76.82	25	90.90
6	85.83	16	79.17	26	79.65
7	77.13	17	95.81	27	81.68
8	83.61	18	77.46	28	91.47
9	86.89	19	95.50	29	90.58
10	86.57	20	89.03	30	79.30

From the above research, it can be seen that the precise calibration system of physical education movements based on artificial intelligence deep learning technology constructed in this paper can effectively improve the effect of physical education.

6. Conclusion

The present study examines the prevalent methodologies for detecting sports targets, which include the optical flow method, inter-frame difference method, and background difference method. Detection methodologies employing optical flow leverage the temporal characteristics of sports objects within imagery, extracting targets by analyzing the displacement vector's optical flow field. While this approach yields targets enriched with sports-specific and structural information, its reliability and precision are often compromised by factors such as noise, multi-light sources, shadows, and occlusions. The computational intensity of this method is substantial, necessitating specialized hardware for real-time detection in practical scenarios, which can be a limiting factor. In this work, we have integrated artificial intelligence and deep learning technologies to develop a precision sports target detection system tailored for physical education applications. The system has been implemented in physical education practice, with the aim of enhancing teaching efficacy. The evaluation of the system through testing indicates that it substantially satisfies the practical requirements of physical education instruction.

REFERENCES

Abanazir, C. (2019). E-sport and the EU: the view from the English Bridge Union. *The International Sports Law Journal*, *18*(3), 102-113.

Ángel-López, J. P., & de la Peña, N. A. (2015). Kinematic hand analysis using motion capture technology. In *VI Latin American Congress on Biomedical Engineering CLAIB 2014, Paraná, Argentina 29, 30 & 31 October 2014* (pp. 257-260). Springer.

Bao, H., Lu, Y., & Wang, Q. (2020). Single target tracking via correlation filter and context adaptively. *Multimedia Tools and Applications*, *79*, 27465-27482.

Bhat, P. G., Subudhi, B. N., Veerakumar, T., Di Caterina, G., & Soraghan, J. J. (2021). Target tracking using a mean-shift occlusion aware particle filter. *IEEE Sensors Journal*, *21*(8), 10112-10121.

Ferguson, L. J., Carlson, K. T., & Rogers, D. (2019). Moving towards reconciliation through sport: sharing our process of exploring team saskatchewan experiences at the North American Indigenous Games. *Journal of Exercise, Movement, and Sport (SCAPPS refereed abstracts repository)*, *51*(1).

Gerke, A., Babiak, K., Dickson, G., & Desbordes, M. (2018). Developmental processes and motivations for linkages in cross-sectoral sport clusters. *Sport Management Review*, *21*(2), 133-146.

Ilies, D., Buhas, R., Ilies, M., Ilies, A., Gaceu, O., Pop, A., Marcu, F., Buhas, S., Gozner, M., & Baias, S. (2018). Sport activities and leisure in Nature 2000 protected area–Red Valley, Romania. *Journal of Environmental Protection and Ecology*, *19*(1), 367-372.

Khan, M. A. (2017). Multiresolution coding of motion capture data for real-time multimedia applications. *Multimedia Tools and Applications*, *76*, 16683-16698.

Kim, M.-K., Kim, T. Y., & Lyou, J. (2015). Performance improvement of an AHRS for motion capture. *Journal of Institute of Control, Robotics and Systems*, *21*(12), 1167-1172.

Kimasi, K., Shojaei, V., & Boroumand, M. R. (2019). Investigation of safety conditions at gymnasia in different organizations. *Journal of Humanities Insights*, *3*(02), 70-74.

Komisar, V., Novak, A., & Haycock, B. (2017). A novel method for synchronizing motion capture with other data sources for millisecond-level precision. *Gait & Posture*, *51*, 125-131.

Kondrukh, A. (2017). Practical shooting sport in Russian sport system: essential specifications and features. *Theory and Practice of Physical Culture*(5), 27-27.

Lee, Y., & Yoo, H. (2017). Low-cost 3D motion capture system using passive optical markers and monocular vision. *Optik*, *130*, 1397-1407.

Li, T., Sun, J., & Wang, L. (2021). An intelligent optimization method of motion management system based on BP neural network. *Neural Computing*

and Applications, *33*, 707-722.

Liu, L., & Hodgins, J. (2018). Learning basketball dribbling skills using trajectory optimization and deep reinforcement learning. *Acm transactions on graphics (tog)*, *37*(4), 1-14.

Liu, S., Wang, S., Liu, X., Lin, C.-T., & Lv, Z. (2020). Fuzzy detection aided real-time and robust visual tracking under complex environments. *IEEE Transactions on Fuzzy Systems*, *29*(1), 90-102.

Pogrebnoy, A., & Komlev, I. (2018). Sport institutions reporting to Ministry of Sport of Russian Federation: intellectual property, invention activity, patenting and legal consulting service analysis. *Theory and Practice of Physical Culture*(2), 2-2.

Puupponen, A., Wainio, T., Burger, B., & Jantunen, T. (2015). Head movements in Finnish Sign Language on the basis of Motion Capture data: A study of the form and function of nods, nodding, head thrusts, and head pulls. *Sign Language & Linguistics*, *18*(1), 41-89.

Reinhart, K., & Wichmann, B. (2020). The TuS Fortschritt Magdeburg-Neustadt (soccer section) in the GDR–an example of amateur socialist sport. *Soccer & Society*, *21*(4), 408-420.

Sá, F., Marques, A., Rocha, N. B., Trigueiro, M. J., Campos, C., & Schröder, J. (2015). Kinematic parameters of throwing performance in patients with schizophrenia using a markerless motion capture system. *Somatosensory & Motor Research*, *32*(2), 77-86.

Sun, K., Li, X., & Shi, W. (2018). The Fusion of Adaptive Color Attributes for Robust Compressive Tracking. *Wireless Personal Communications*, *102*, 879-894.

Wang, L., Lu, H., & Yang, M.-H. (2017). Constrained superpixel tracking. *IEEE transactions on cybernetics*, *48*(3), 1030-1041.

Zhang, J., Jin, X., Sun, J., Wang, J., & Sangaiah, A. K. (2020). Spatial and semantic convolutional features for robust visual object tracking. *Multimedia Tools and Applications*, *79*, 15095-15115.

Zhi-Min, Z., & Zhong-Wen, C. (2015). A survey of motion capture data earning as high dimensional time series. *International Journal of Multimedia and Ubiquitous Engineering*, *10*(9), 17-30.