Enhanced OpenPose Framework for Athlete Motion Analysis via Multi-Classifier Fusion and D-S Decision Rules

Chunzhou Zhang
Department of Physical Education, Zibo Vocational Institute, Zibo, 255000, China
zhangchunzhou1989@163.com

Keywords: OpenPose, athletes, classifiers, skeletal keypoint detection, motion analysis

Received: April 28, 2025

This study aims to address the issue that traditional analysis methods do not fully consider the differences in the number, types, and equipment occlusion of athletes in actual sports scenes. It constructs an athlete skeleton keypoint detection and action analysis model based on the OpenPose algorithm. The original OpenPose pipeline was optimized by introducing a multi classifier module combining SVM, KNN, and Naive Bayes, as well as a decision model based on conflict judgment and D-S rule fusion. The experiment used a dataset containing 2000 athlete training images (divided into training, testing, and validation sets in a 7:2:1 ratio), and compared it with the original OpenPose and DeepPose baseline models. The results showed that the optimized model had a loss rate of 4.1% on the training set and an accuracy rate of 98%. The error rate of keypoint detection was 7.4%, the speed was 14.2fps, the detection accuracy was 0.92 mAP, and the time consumption was 0.9s. In the case analysis, the highest discrimination rates of the model for football players' running and defensive actions were 0.94 and 0.89, respectively. Moreover, it could further distinguish between subdivided actions such as sprinting and backward running. In addition, the model performed well in different types of athletes (including athletics, weightlifting, gymnastics, etc.) and complex scenes (occlusion, multiple people), with significantly better accuracy and recall than the baseline model (mAP=0.84 for the original OpenPose and mAP=0.78 for DeepPose). The research reveal that the proposed model has higher accuracy and robustness in athlete bone keypoint detection and action analysis, and can provide effective support for scientific training.

Povzetek: Okvir OpenPose je izboljšan z večklasifikatorsko fuzijo (SVM, KNN, NB) in D-S odločitvenimi pravili. Omogoča bolj kvalitetno zaznavo skeletnih točk ter analizo gibanja športnikov kot osnovni OpenPose in DeepPose.

1 Introduction

The identification and evaluation of athletes' motions has long been a crucial component in raising the level of competition and training effectiveness in the realm of sports competition. The subtle differences in athletes' movement patterns, power application, coordination, and trajectories can be the key factors affecting victory or defeat [1]. Therefore, accurate capture and in-depth analysis of athletes' movements play a crucial role in optimizing training programs and improving athletic performance. However, most of the traditional movement analysis methods rely on manual observation or simple video analysis. This is not only time-consuming and laborious, but also highly susceptible to the influence of subjective factors, making the analysis results not objective and accurate. To address this problem, in recent years, many scholars have begun to try to use computer technology to build intelligent models, with a view to realizing scientific sports training.

Cronin et al. compared the performance of OpenPose with manually examined data from the SIMI campaign using film shot at 200 Hz from actual competitive settings using two cameras. It was discovered that the joint angle waveforms' multiple coefficients of determination showed significant variations among athletes, with knee angle values being more constant than those for hip and ankle angles [2]. Based on multi-view cross-information and confidence metrics, Xu M et al. presented a reliable multi-view anomalous human pose recognition technique that assessed pose attributes from various angles. Specifically, human skeletal data were extracted using OpenPose. These data were then used as inputs to the YOLOv5s system, which detected abnormal poses, such as falls and bumps. Experimental results found that the model had high accuracy [3]. To address the subjectivity and inefficiencies of conventional manual counting, Shi Z et al. suggested a sit-up counting method based on bone keypoint identification. By enhancing the network topology and adding jump connections, the system used a deep learning algorithm to identify and locate the main areas of the human skeleton [4]. A discontinuous frame filtering module was added to the front-end of the gait feature extraction (FE) network by Han K et al. in an attempt to actualize the limitation of the input image information to the network. This allowed them to propose a study approach based on human body

keypoint extraction. According to the experimental findings, the model demonstrated great robustness and an average recognition accuracy of 79.5% on the gait dataset [5].

Table 1: Literature comparison table

Reference	Method	Key contribution	Performance	Application scenario	Limitations
Cronin et al. [2]	Original OpenPose	Compared with SIMI manual analysis	High knee joint consistency (R^2>0.85)	Track & field (e.g., long jump)	Unsuitable for real-time competition analysis
Xu et al. [3]	OpenPose+YOLOv5s	Multi-view abnormal posture detection	Fall detection accuracy: 92.3%	Security surveillance	High computational complexity
Shi et al. [4]	Enhanced keypoint network	Skip-connection architecture	Sit-up counting: 98.57%, 63.6 FPS	Fitness training monitoring	Limited to specific movements
Han et al. [5]	Keypoint extraction+frame filtering	Discontinuous frame screening module	Gait recognition: 79.5% accuracy	Gait analysis	Poor adaptability in dynamic scenes

Even while earlier research has shown some promising findings, there is still opportunity for advancement. Due to the differences between different sports, athletes have different focuses on their own movements and postures. Therefore, the previous models do not further consider the differences in the number of athletes and types of athletes in the actual sports scene, as well as the existence of the occlusion of equipment. Based on this, the study proposes to use the OpenPose algorithm to construct a skeletal keypoint detection and movement analysis model for athletes, and to optimize the model on the basis of this model. The innovation of the optimized model is that OpenPose is used as the backbone of the model, and three classifiers, namely support vector machine (SVM), K-nearest neighbor (KNN) and naive Bayes (NB), are combined to construct the multi-classifier module. In addition, the multi-classifier module is followed by the decision module, which together enhances the model's overall performance and detection accuracy (DA).

This study aims to address the following key research questions: (1) Can the multi classifier fusion strategy significantly improve the robustness of pose classification under occlusion conditions? (2) In what ways can decision modules based on D-S rules surpass traditional majority voting methods? How can they improve the reliability of final classifications through conflict judgment mechanisms and confidence fusion? (3) What is the generalization ability of the proposed optimization model in different sports projects? The exploration of these issues will provide new technical ideas and theoretical basis for posture analysis in complex motion scenes.

2 Methods and materials

2.1 Construction of an OpenPose-based model for skeletal keypoint detection and movement analysis of athletes

Athletes are professionals who specialize in the training and competition of sports. They are often required to improve their fitness, skills and athleticism through long-term, systematic and scientific training in order to achieve results in sports competitions. As a result, sports athletes are often required to perform highly precise and coordinated movements during training and competition. This is not only to improve the athletes' professional level, but also to avoid sports injuries, as incorrect posture can easily lead to injuries [6-7]. Based on this, the detection and analysis of athletes' movements has become the focus of the field today. OpenPose, as a posture estimation algorithm, recognizes and detects the key points of human skeleton in images or videos to analyze human posture and movements. Therefore, with the help of OpenPose algorithm, the detection of key points of the athlete's skeleton can be realized so as to analyze the movement. This can provide scientific data support and feedback for athletes' training [8]. Among

them, the specific processing flow of the OpenPose algorithm is shown in Figure 1.

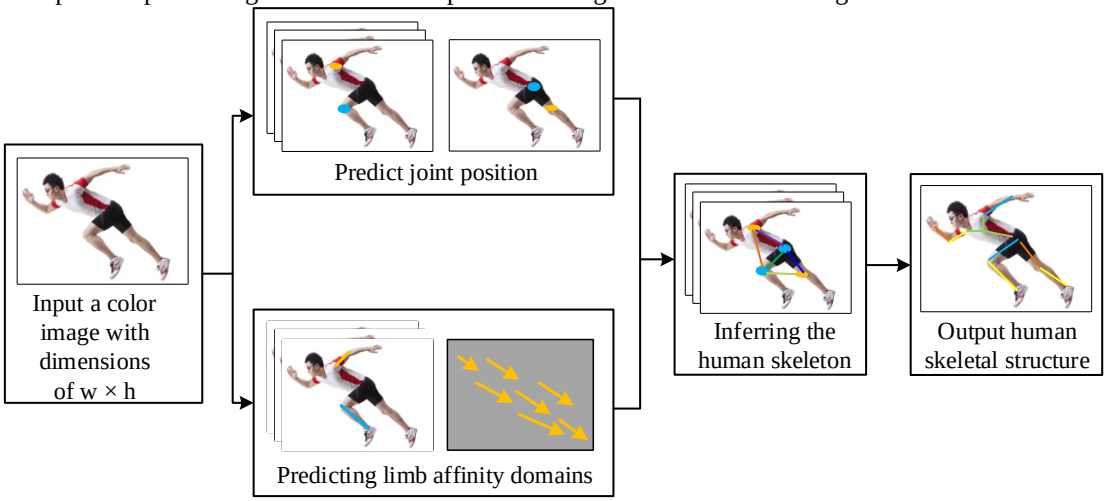


Figure 1: Flow of OpenPose algorithm

Figure 1 shows the specific flow of OpenPose algorithm. The input data is an image of a certain size, and OpenPose processes this input image in stages [9-10]. One stage is mainly used to predict the position of joint points and generate 2D confidence maps (CMs), while the other stage is mainly used to predict the affinity domains of limbs and output 2D vector fields. When the two phases are finished, the skeleton of the human body is jointly reasoned, and the 2D positions of the human skeletal structure are finally output. The specific calculation of this process is shown in Equation (1).

$$\begin{cases} S^t = \rho^t(F, S^{t-1}, L^{t-1}), \forall t \geq 2 \\ L^t = \phi^t(F, S^{t-1}, L^{t-1}), \forall t \geq 2 \end{cases} \quad (1)$$

In Equation (1), L^t is the weighting coefficient, which mainly serves to quantify the correlation between the joints and provide a numerical basis for the subsequent analysis. S^t is the CM of joints output in t stage. F is the feature map (FM) after the convolutional network operation. Within the structure of the OpenPose algorithm, there are two branches, the upper and lower branches, and each branch is an iterative prediction architecture, as shown in Figure 2.

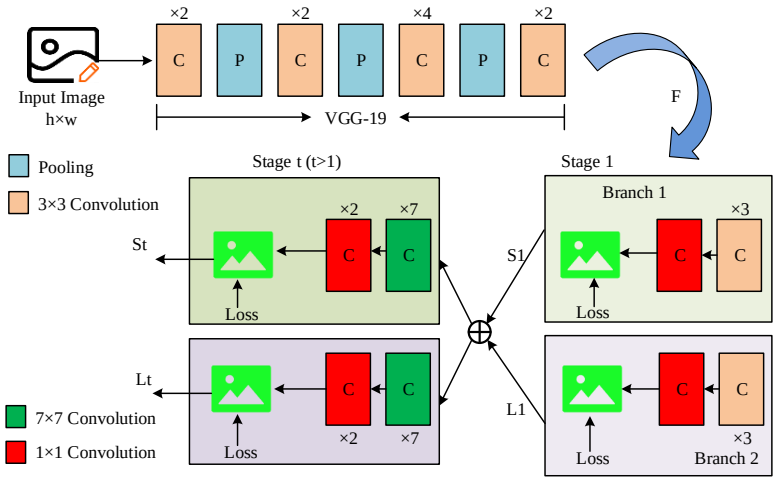


Figure 2: Schematic diagram of the internal structure of OpenPose

Figure 2 shows the internal structure of OpenPose. There are not only two branches up and down inside the standard OpenPose, but there is also a backbone network consisting of the first 10 layers of VGGNet at the front of its branches [11-12]. To extract valuable feature information from the input image, this backbone network is mostly utilized as a feature extractor. To make the detection speed and DA increase, the research has improved on the original OpenPose. The backbone network of the original OpenPose uses VGG-19. To make the model more lightweight, the study replaces VGG-19 with a lightweight neural network MobileNetV1, as shown in Equation (2)

$$\begin{cases} MACC = H \times W \times 1 \times 1 \times \frac{c_1}{g} \frac{c_2}{g} \times g = \frac{h\omega c_1 c_2}{g} \\ FLOP_s = 2 \times MACC \\ Params = g \times \frac{c_2}{g} \times \frac{c_1}{g} \times 1 \times 1 + c_2 = \frac{c_1 c_2}{g} \\ MAC = HW(c_1 + c_2) + \frac{c_1 c_2}{g} \end{cases} \tag{2}$$

In Equation (2), $MACC$ denotes the computation of point-by-point convolution in MobileNetV1. $FLOP_s$ denotes the floating-point operation in the convolutional layer. $Params$ is the number of parameters. MAC is the computational complexity. H and W denote the height and width of the FM, respectively. c_1 and MAC both denote the channel's quantity. g denotes the number of groups of convolutional kernels. At this point, the activation function used is ReLU, as shown in Equation (3).

$$\begin{cases} \mathrm{Re}\,LU = \max(x, 0) = \begin{cases} 0, x < 0 \\ x, x > 0 \end{cases} \\ \delta^l = \frac{\partial L}{\partial z^{l+1}} \frac{\partial z^{l+1}}{\partial z^l} \end{cases} \tag{3}$$

In Equation (3), z^{l+1} is the output of layer $l+1$. δ^l is the bias derivative of the output z^l of the loss function (LF)'s l th layer. It is worth noting that the upper and lower branches of OpenPose use a LF at the end of each stage, as shown in Equation (4).

$$\begin{cases} f_S^t = \sum_{j=1}^{J} \sum_p W(p) \cdot \left\| S_j^t(P) - S_j^*(p) \right\|_2^2 \\ f_L^t = \sum_{c=1}^{C} \sum_p W(p) \cdot \left\| L_c^t(P) - L_c^*(p) \right\|_2^2 \end{cases} \tag{4}$$

Equation (4) is the formula for the loss value in the OpenPose stage. p is the pixel point. $W(p)$ is the missing marker of the pixel point, which is replaced by 1 or 0 only. When the value of the marker is 0, this means that the missing value is not included in the calculation. S_j^* and L_c^* represent the truth value of the joint point CM or partial affinity vector field of the corresponding branching stage, respectively [13-14]. During training, losses are incurred at each stage. Therefore, to avoid gradient vanishing, the study periodically replenishes the gradient at each stage with intermediate supervision. Equation (5) illustrates that just the output of the final layer is utilized for prediction.

$$f = \sum_{t=1}^{T} (f_S^t + f_L^t) \tag{5}$$

Equation (5) is the overall objective of training, i.e., the overall loss value is obtained by relying on the summation of the loss values of each stage. In addition, in OpenPose, the prediction about the CM of joints is mainly realized by convolutional pose machines (CPM). The specific structure is shown in Figure 3.

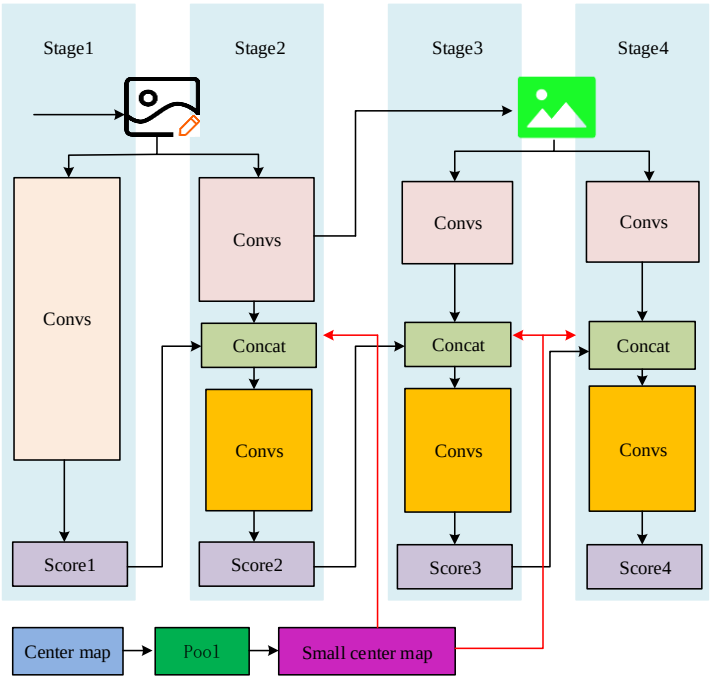


Figure 3: Schematic diagram of CPM structure

Figure 3 shows a schematic of the internal structure of the CPM. The CPM is internally composed of multiple convolutional network stages. These stages process the input image and the CM of the previous stage sequentially to generate increasingly accurate joint position estimates. The system also implicitly learns the

spatial relationships between images and joint points [15-16]. Therefore, as the network training progresses and the CPM stages increases, the accuracy of the output CMs improves, ultimately realizing accurate prediction of athletes' skeletal joint point positions. The specific calculation is displayed in Equation (6).

$$\begin{cases} S_{j,k}^{*}(p)=\exp(-\dfrac{\left\|p-x_{j,k}\right\|_2^2}{\sigma^2}) \\ S_{j}^{*}(p)=\max\limits_{k} S_{j,k}^{*}(p) \end{cases} \tag{6}$$

Equation (6) is the training process of the joint point CM. S^{*} is the true CM generated based on the keypoints. $x_{j,k}$ is the true location of the joints j of individual k. $x_{j,k}\in R^2$ $S_{j,k}^{*}$ denotes the individual CM. CPM effectively infers the occluded joints by obtaining a large receptive field through the use of a large convolution kernel. However, increasing the convolution kernel also increases the number of parameters. Based on this, the study proposes to increase the sensory field of CPM by increasing the step size to optimize CPM. In addition to CPM, another branch of OpenPose is part affinity fields (PAF). Its specific calculation is shown in Equation (7).

$$\begin{cases} L_{c,k}^{*}(p)=\begin{cases} v, if\ \ p \quad on \quad c\lim bc,k \\ 0 \quad otherwise \end{cases} \\ L_{c}^{*}(p)=\dfrac{1}{n_c(p)}\sum\limits_{k} L_{c,k}^{*}(p) \end{cases} \tag{7}$$

In Equation (7), $n_c(p)$ is the quantity of all non-zero vectors at point p. $L_{c,k}^{*}(p)$ denotes the affinity vector, whose main role is to assess the likelihood of joint point connections. v denotes the unit vector of $L_{c,k}^{*}(p)$ from $j1$ to $j2$. After obtaining the CMs of the joints, it is also necessary to calculate the line integrals of the corresponding PAFs. The line integral is calculated to obtain the strength of association score between the two joints. A higher score obtained indicates the possibility of a connection between the joint points. However, when there is more than one person in the input image, there may be many candidate locations for each part detected at this point. To cope with this problem, the study uses the Hungarian algorithm to evaluate each candidate limb. This is shown in Equation (8).

$$\max_{Z} E=\max_{Zc}\sum_{c=1}^{C}\max_{Zc} Ec \tag{8}$$

In Equation (8), Ec is denoted as the overall matching weight value for limb C. Zc then denotes the subset of the Z set with respect to limb C. E is the connectivity strength between the parts. The evaluation criteria are determined by the connectivity strength, and the one with the highest strength is the candidate site [17]. In this way, an athlete skeletal keypoint detection and movement analysis model based on the OpenPose algorithm can be finally constructed, as shown in Figure 4.

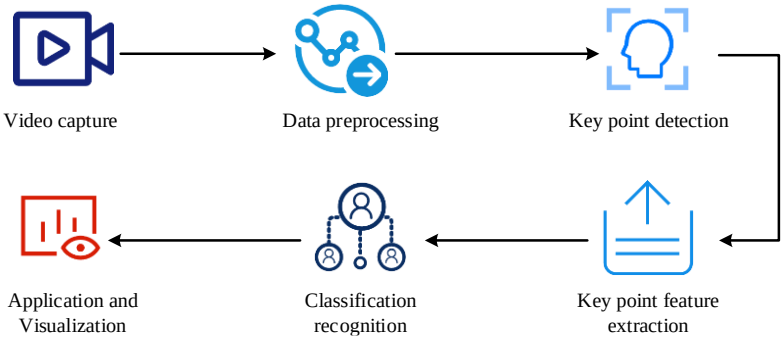


Figure 4: Athlete skeletal keypoint detection and movement analysis model

The OpenPose algorithm-based skeletal keypoint identification and movement analysis model designed for athletes is displayed in Figure 4. In Figure 4, the specific process of the model is that the video of the athlete practicing or competing is captured as the input data, and then preprocessed to form an image. Moreover, the preprocessed image is subjected to a series of operations by OpenPose algorithm, including keypoint detection, extraction of keypoint features, and classification and recognition. Finally, the visualized heatmap is output to realize the analysis of the athlete's movement.

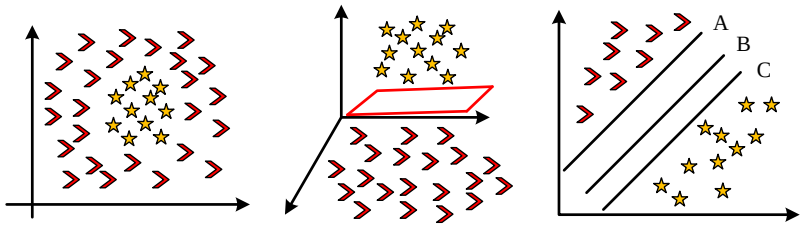
2.2 Optimization model for OpenPose incorporating multiple classifiers

Although the model constructed using the OpenPose algorithm can detect and analyze the movements of athletes, however, in actual sports training or competition, there is often the same occasion. There are several trained athletes, and there are differences in the training programs and movements between different athletes. This difference may cause the original OpenPose algorithm to make judgment errors in detection, which in turn affects the accuracy of movement analysis. To deal with a problem, the study proposes to optimize the detection of skeletal keypoints on the basis of OpenPose algorithm. An optimization model for skeletal keypoint detection and movement analysis of athletes based on OpenPose is constructed, referred to as the optimization model.

During the model optimization process, the researchers

first replaced the VGG19 backbone network with a MobileNetV1 network. They uses a depth multiplier of 0.75 to balance accuracy and efficiency. This adjustment can reduce the number of network parameters from the original OpenPose's 125.6M to 38.3M, and the computational complexity (FLOPs) from 52.4G to 15.8G. Specifically, the depthwise separable convolution structure of MobileNetV1 decomposes standard convolutions into depthwise and pointwise convolutions. In FM computation, the number of convolution kernel parameters is reduced to the original $\frac{1}{N}+\frac{1}{D_k^2}$. It significantly reduces model complexity while maintaining the original FE capability.

The key problem highlighted in the original model is that there are omissions and errors in the detection process of athletes' skeletal keypoints. Therefore, this study introduces multiple classifiers behind the FE module of the original model to improve the model's ability of classification and recognition [18]. To increase the accuracy of the output findings, a decision module is also added after the classifier module to merge the conflict judgment with the detection speed synthesis rules. For the introduction of multiple classifiers, which are NB, SVM, and KNN classification respectively. All the three classifiers each have different characteristics. Classification is achieved by finding the hyperplane that maximizes the spacing between the sample points of different categories, as shown in Figure 5.



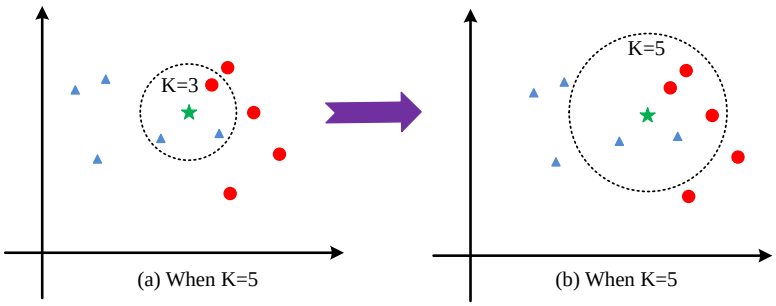
(a) 2D Space (b) 3D Space (c) Cross sectional diagram

Figure 5: Schematic diagram of SVM hyperplane

Figure 5 shows the 2D and 3D views of the SVM hyperplane. Figure 5(c) shows the cross-sectional view. Among them, the line C is the support vector, which is formed by the connection between the nearest two different categories of sample points and is parallel to the hyperplane [19]. The distance between A and B in the figure is then the interval of the SVM. The larger the interval, the more correct the classification is indicated. BN is a classifier based on Bayes' theorem, which obtains the optimal classification by learning the prior probability and conditional probability. Equation (9), which illustrates the fundamental notion of this classifier, reduces the conditional probability to the probability of a single feature because BN assumes that all features are independent of one another.

$$\begin{cases} \arg\max\limits_{c_k} p(y=c|x) \\ \arg\max\limits_{c_k} p(x|y=c_k)\,p(y=c_k) \end{cases} \tag{9}$$

In Equation (9), $n_c(p)$ and $p(y=c_k)$ are the conditional and prior probability. x is the input data to be measured. c_k is the category corresponding to the k th position in the category set. Compared with SVM, NB is more sensitive to the assumption of feature independence. However, its advantages include: (1) Strong robustness to missing data. (2) Higher computational efficiency with high-dimensional data. (3) Stable performance when the feature independence approximation holds, even with small samples. Moreover, in addition to the introduction of SVM with BN, the improved optimization model also introduces KNN as a classifier [20-21]. KNN is a classifier based on distance metric. Therefore, the classification criteria of this classifier and its results mainly depend on the number of neighbors, i.e., the K-value, and the distance metric. The details are shown in Figure 6.



(a) When K=5 (b) When K=5

Figure 6: Classification discrimination of KNN

Figure 6 shows the classification interpretation process when KNN is used as a classifier. The green dot at the center in Figure 6(a) is the point to be predicted. When K is 3, the to-be-predicted point is considered to be blue

because the circle around the center of the point has 2 blue dots and 1 red dot. Moreover, when K is equal to 5, then there are 2 blue points and 4 red points within the new circle. At this point, it is then considered that the electrodynamic point belongs to the red side. Obviously, all three classifiers have their own characteristics and advantages. Since most of the previous methods for the detection and recognition of gesture behavior use a single classifier, multiple classifiers are fused. This enhances DA in addition to recognizing the complementarity of the classifiers' strengths and limitations [22-23]. The three classifiers have complementary strengths. SVM excels in small-sample, high-dimensional scenarios by maximizing the margin. KNN uses local data density to analyze dynamic actions. NB remains robust in sparse or incomplete data when features are independent. Together, they enable a comprehensive analysis of both common (data-rich) and rare (data-scarce) football movements.

The study employs a two-stage fusion strategy: First, initial weights are assigned to SVM, NB, and KNN based on their validation set accuracy. Second, the D-S evidence theory dynamically adjusted the weights when the outputs of the classifiers diverged (below the similarity threshold). This boosted the influence of consistent classifiers while reducing that of conflicting ones. This balances overall performance with sample-specific adaptation. In addition to the incorporation of a multi-classifier module, a decision-making module is also incorporated in the optimization model for skeletal keypoint detection and movement analysis of athletes based on improved OpenPose, as shown in Equation (10).

$$\begin{cases} m_1 = \{m_1(H_1), m_1(H_2), ..., m_1(H_p)\} \\ m_2 = \{m_2(H_1), m_2(H_2), ..., m_2(H_p)\} \\ m_1 = \{m_3(H_1), m_3(H_2), ..., m_3(H_p)\} \end{cases} \tag{10}$$

In Equation (10), m_1, m_2, and m_3 are the detection results of the three classifiers, respectively. p is the type of behavior. $m_1(H_p)$, $m_2(H_p)$, and $m_3(H_p)$ are the detection rates of the three classifiers in the pth behavior, respectively [24-25]. Since three classifiers exist, three results theoretically exist. This may lead to contradictions and conflicts; thus the first step of the decision module is to make a conflict judgment as shown in Equation (11).

$$\rho_{m_i,m_j} = \frac{\mathrm{cov}(m_i, m_j)}{\sigma_{m_i}\sigma_{m_j}} = \frac{E[(m_i - \mu_{m_i})(m_j - \mu_{m_j})]}{\sigma_{m_i}\sigma_{m_j}}, (i, j = 1,2,3,...,n) \tag{11}$$

In Equation (11), ρ_{m_i,m_j} is the similarity between two two classifiers. ρ_{m_i,m_j} and μ_{m_j} display the mean of the i th and j th classifiers, respectively. σ_{m_i} and σ_{m_j} display the variance of the i th and j th classifiers, respectively. When the similarity is 1, it means there is no conflict. When the similarity is -1, it indicates complete conflict [26-29]. After the conflict determination, then the D-S rule fusion operation can be performed as shown in Equation (12).

$$\begin{cases} m(A) = \dfrac{\sum_{B\cap C=A} m_1(B)\cdot m_2(C)\cdot m_3(D)}{1-\sum_{B\cap C=\varnothing} m_1(B)\cdot m_2(C)\cdot m_3(D)} \\ k = \sum_{\bigcap_{i=1}^{n} A_i \neq \Phi} m_1^{'}(H_1)\cdot m_2^{'}(H_2)\cdot .....\cdot m_n^{'}(H_p) \end{cases} \tag{12}$$

Equation (12) uses the classical Dempster combination rule. In this equation, B, C, and D represent the mass function assignment sets of the SVM, KNN, and NB classifiers, respectively. A represents the hypothesis to be verified, such as a specific motion behavior. The sum of the mass products of all empty intersections are calculated using the conflict coefficient $k \cdot m_n^{'}(H_p)$ stands for mass function. The specific implementation process of the D-S decision module is shown in Figure 7.

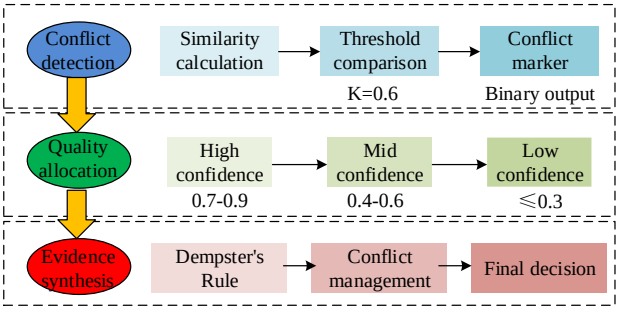


Figure 7: Schematic diagram of D-S fusion process

Figure 7 outlines the three-stage D-S fusion process. First, the system detects classifier conflicts using a similarity matrix (Equation 11). Then, it triggers reweighting when the number of conflicts exceeds the 0.6 threshold. It then dynamically allocates mass function values before finalizing through Dempster's combination rule: 0.7–0.9 for high-confidence/low-conflict predictions, 0.4–0.6 for medium confidence, and ≤0.1 for low-confidence/high-conflict cases. The optimized model structure can be shown in Figure 8.

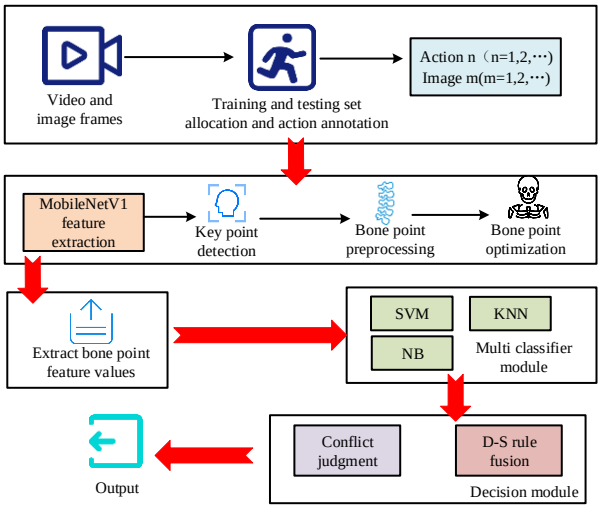


Figure 8: Optimized model of skeletal keypoint detection

and motion analysis for athletes with improved OpenPose

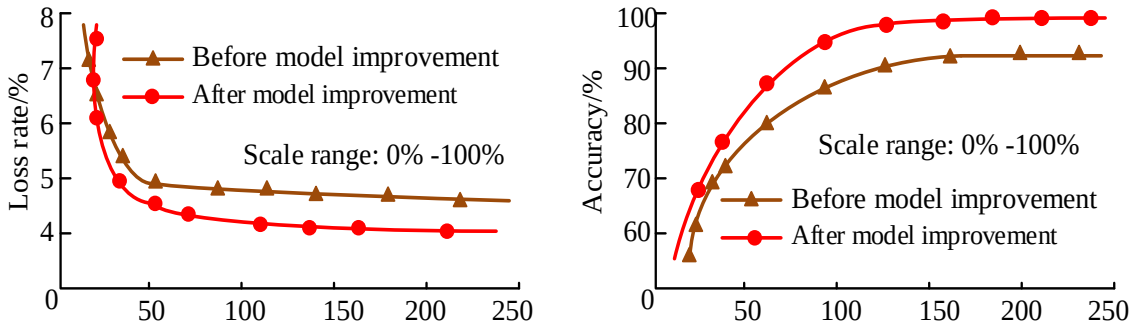
Figure 8 shows the optimized model for athlete skeletal keypoint detection and movement analysis after the introduction of the multi-classifier module and decision module. The optimized model is still based on the OpenPose algorithm, which is mainly used for the detection and FE of athletes' skeletal keypoints. The model's overall DA is increased with the addition of a multi-classifier and decision module. This can avoid the situation of missed detection and false detection due to the complex situation of multiple people and dynamics. In summary, the improved model has great potential in applying to athletes in sports.

3 Results

3.1 Validation analysis of skeletal keypoint detection and motion analysis optimization models for athletes

To examine the actual performance of the proposed model of the study, the study selects the training images of a soccer club's athletes and their training videos since nearly one year as the experimental dataset. The video is intercepted frame by frame and preliminarily screened, and finally a total of 2000 original images that meet the requirements are collected. Additionally, to enhance the effectiveness of picture recognition and classification, the image is normalized. In accordance with 7:2:1, the dataset is separated into training, test, and validation sets. The corresponding number of images are 1400, 400, and 200, respectively. The experiments are run on the integrated development environment PyCharm 2022 and operated on Ubuntu 18.04.6. The dataset is captured by professional sports photographers using a Sony α9 II (4K/60fps), with 2000 frames (1920×1080) extracted from the footage. Three nationally certified football referees independently labeled behaviors (running, shooting, defending, etc.) through a semi-automatic process: OpenPose-generated keypoints are manually verified and corrected, with disputed samples (5.3%) resolved via group arbitration. CVAT ensured precise keypoint-label alignment. The complete processing flow includes: (1) video capture (fixed camera position, 10-15 meters distance, 800-1200 lux lighting), (2) frame extraction and normalization (FFmpeg tool, resolution 640×360), (3) three level annotation system (automatic generation manual verification arbitration), and (4) Model training (Adam optimizer, 150 epochs). For subsequent elaboration, the study names the models as pre-improvement model and post-improvement model, respectively, and the comparison method is used. The two models before and after optimization are trained separately using the training set data. The changes in the loss rate and accuracy of the models are shown in Figure 9.

Figure 9 displays the model's iterative performance on the training set before and after enhancement. The model's loss change curve is displayed in Figure 9(a). The two models, both before and after improvement, show the same overall trend of gradually decreasing loss rates with increasing iterations. In terms of individual models, however, the improved model reaches the inflection point of convergence faster than the pre-improved model. The improved model finally converges completely at a loss rate of 4.1%. The loss rate at the point of complete convergence of the pre-improved model is 4.8%, which is higher than that of the improved model. Figure 9(b) shows the change in accuracy of the two models. The accuracy of both the before and after models gradually increases with each iteration until they reach a stable state. When the enhanced model converges, the comparison findings indicate that its accuracy is roughly 98%. The pre-improved model, on the other hand, reaches full convergence with an accuracy of about 90%, which is lower than that of the improved model. Further, the experiments also test the error rate and detection speed of the before and after models for skeletal keypoint detection. The results are shown in Figure 10.



(a)Loss rate performance (b)Accuracy performance

Figure 9: Loss change curve and accuracy change curve of the model before and after optimization

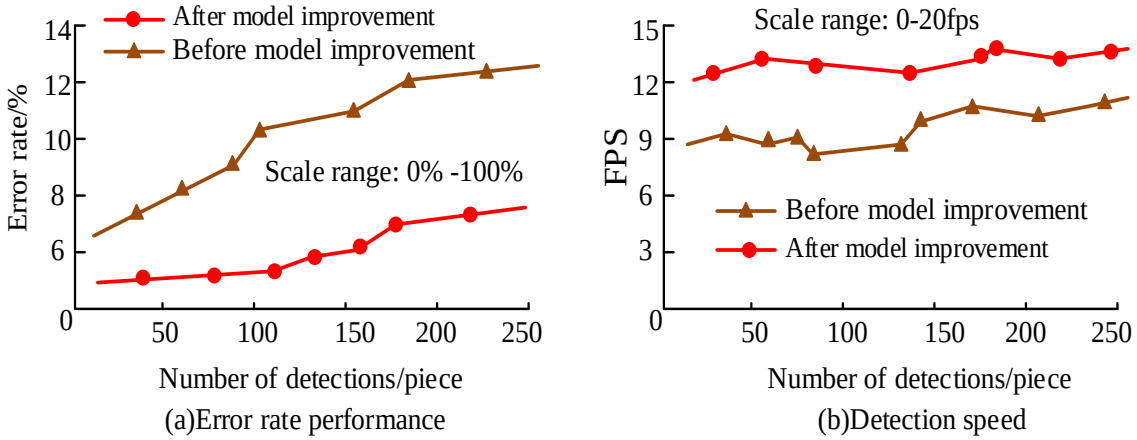


Figure 10: Error rate and detection speed of model keypoint detection

Figure 10 compares the two models' performance in terms of error rate and detection speed for skeletal keypoint identification of athletes before and after improvement. Figure 10(a) displays the error rate of the two models for keypoint detection. As the detections increases, the overall error rate of both models before and after improvement shows an increasing trend. In terms of individual models, the average error rate of the pre-improved model is about 12.3%. The improved model, on the other hand, has an error rate of about 7.4%. In comparison, the improved model has fewer error cases on skeletal keypoint detection. Figure 10(b) displays the detection speed of the two models. TBoth models' overall detection speed is comparatively stable. However, the average detection speed of the improved model is 14.2 fps, while the average detection speed of the pre-improved model is only 9.7 fps, which is lower than that of the improved model. In addition, the mean average precision (mAP) and detection time before and after the model improvement are shown in Figure 11.

Figure 11 shows the comparison of the DA and time-consumption of the two models before and after the improvement. Figure 11(a) shows the keypoint DA of the models. The mAP values of both models increase as the number of keypoints increases. The improved model first reaches a state of convergence, with an mAP value of about 0.92 at full convergence. Whereas the pre-improved model shows the inflection point later. Its mAP value at full convergence is about 0.84, which is lower than that of the improved model. Figure 11(b) displays the time-consumption comparison of the two models. The time-consuming performance of both models is relatively stable. On average, the pre-enhanced model take around 2.9 seconds to complete, whereas the improved model take approximately 0.9 seconds. Obviously, the combined performance of the improved model is better. In addition, the improved model incorporates a multi-classifier module, whereas previous methods for behavior detection mostly use one classifier. Therefore, the experiments are tested to compare the differences in the type and number of classifiers. The mean precision, accuracy, recall, and F1 score are selected as the effect assessment metrics for the classifier module. Table 2 displays the findings.

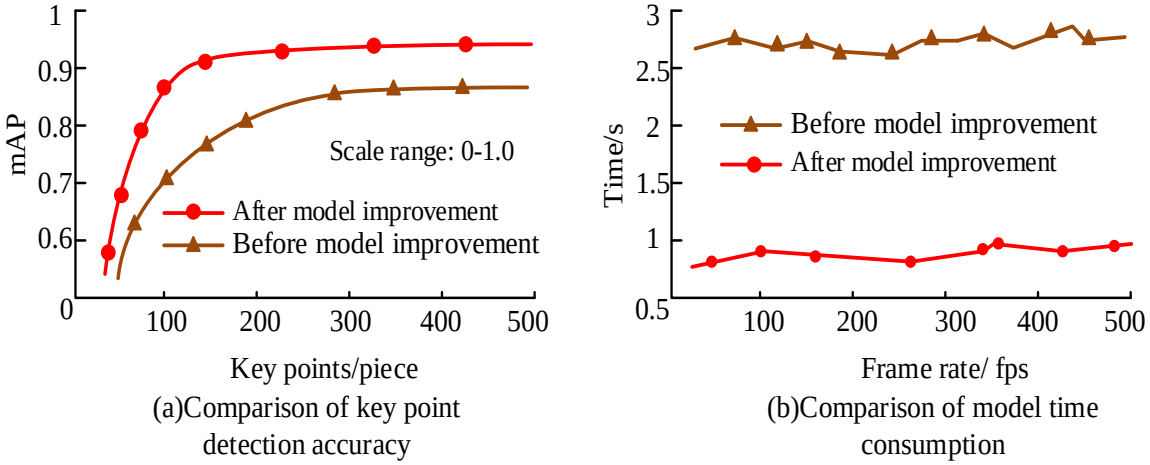


Figure 11: Model key point DA and time-consumption

Table 2: Classification effectiveness evaluation of the classifier module

Classifier	Precision	accuracy	Recall	F1 score
SVM	89.2%	87.3%	86.4%	87.5%
NB	86.4%	85.1%	85.7%	86.6%
KNN	87.8%	84.7%	88.3%	87.2%
SVM-NB-KNN	95.2%	94.6%	91.7%	93.6%

Table 2 shows the performance of single classifiers and multi-classifier modules on the classification effect of athlete behavioral gesture recognition and detection. The performance of the first three single classifiers in terms of precision, accuracy, recall, and F1 score have their own slight advantages and disadvantages, but the overall difference is not significant. Compared with the multi-classifier module containing three classifiers, its precision is 95.2%, accuracy is 94.6%, recall is 91.7%, and F1 score is 93.6%, all of which are higher than any single classifier.

Table 2 shows that while SVM and NB perform well in low-dimensional spaces, KNN is still crucial for (1) capturing spatio-temporal continuity in dynamic actions (as opposed to hyperplane-based approaches), (2) error correction via neighborhood voting when keypoints are inaccurate, and (3) detecting localized density variations in multiplayer interactions to complement the global patterns of SVM/NB.

The multi-classifier module indirectly enhances keypoint detection through a two-level feedback mechanism: (1) Behavior classification identifies kinematically implausible postures, triggering: (a) low-quality frame marking, (b) temporal re-estimation (5 adjacent frames), and (c) subsequent frame weight adjustment. (2) OpenPose's body-part detection weighting is guided by D-S fused action probabilities. This dual-function system improves behavior recognition accuracy directly and refines keypoint detection quality indirectly.

To comprehensively evaluate the performance of the model, statistical significance analysis was supplemented in the study. The results are shown in Table 3.

Table 3: Statistical significance comparison of model performance

Metric	Our Model (95% CI)	Original OpenPose (95% CI)	Lightweight OpenPose	HRNet	OpenPose+MobileNetV3	Ino T et al. [30]	Lee P et al. [31]	*p*-value
Accuracy (%)	97.8±0.3	89.5±0.7	93.2±0.5	95.2±0.4	92.8±0.6	89.2±0.8	99.0±0.2	<0.01
mAP	0.92±0.02	0.84±0.03	0.87±0.03	0.89±0.02	0.86±0.03	0.86±0.04	0.85±0.03	<0.05
Occlusion mAP	0.88±0.03	0.68±0.05	0.80±0.04	0.82±0.04	0.78±0.05	0.76±0.06	0.79±0.05	<0.01
Processing Speed (FPS)	14.2±0.5	9.7±0.8	16.5±0.7	10.3±0.6	18.2±0.9	9.3±0.7	11.7±0.6	<0.05

Table 3 presents the performance comparison with statistical significance. The proposed model significantly outperforms the original OpenPose model and three optimized models (p<0.05). It achieves 97.8±0.3% accuracy, 0.92±0.02 mAP (mAP), and 0.88±0.03 occlusion mAP. Notably, it surpasses HRNet by 6 percentage points in occlusion mAP (0.88 vs 0.82) while being 38% faster (14.2 and 10.3 FPS). Compared to lightweight models, it improves accuracy by 4.6 points (97.8% and 93.2%) with comparable speed. All results are statistically significant (95% CI, t-tests).

Furthermore, in order to verify its effectiveness in computational efficiency, the performance of the model is systematically tested on different hardware platforms. The results are shown in Table 4.

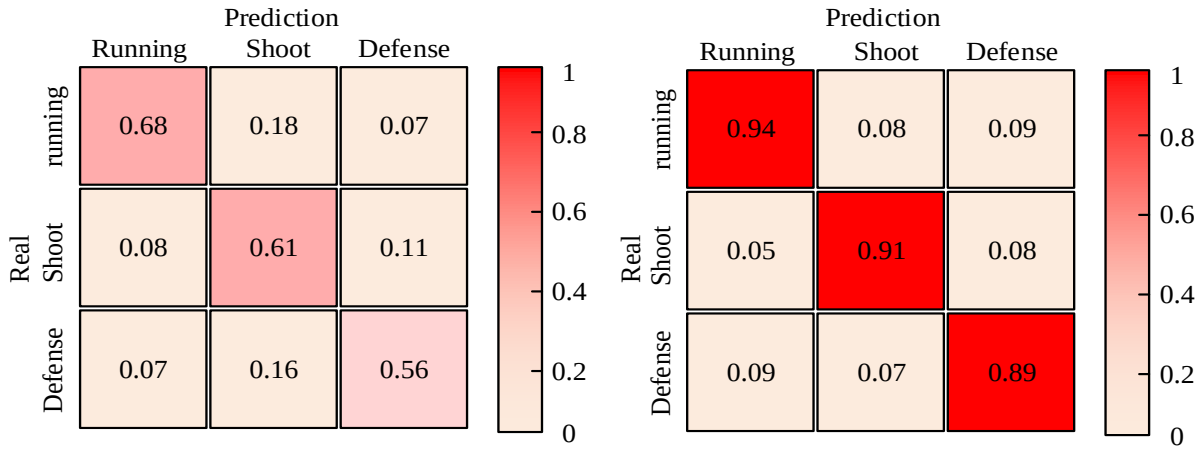
Table 4: Hardware performance benchmarking

Hardware Configuration	Speed (fps)	Latency per Frame (ms)	Memory Usage	Peak Power (W)	Compute Platform
NVIDIA RTX 3090 (GPU)	14.2±0.5	70.4±2.5	2.3 GB GDDR6X	285	CUDA 11.1
Intel Xeon E5-2680 v4 (CPU)	3.1±0.3	322.6±31.2	1.8 GB DDR4	120	OpenMP 4.5
NVIDIA Jetson Xavier NX (Edge)	8.7±0.4	114.9±5.8	1.9 GB LPDDR4	45	JetPack 4.6

Table 4 compares hardware performance. On an NVIDIA RTX 3090 GPU, the model achieves 14.2±0.5 fps (70.4ms/frame) with 2.3GB VRAM usage. On an Intel Xeon E5-2680 v4 CPU, performance drops to 3.1±0.3 fps (322.6ms/frame) with 1.8GB RAM usage. Notably, the multi-classifier module accounts for 23% of the total computation, while the D-S decision module adds only a 7% overhead. This enables deployment on various edge devices.

3.2 Case study of skeletal keypoint detection and motion analysis optimization model for athletes

All of the above are comparisons of the performance of the model before and after this improvement. The investigation will further examine players' behavioral behaviors in particular sports settings with example applications in an attempt to better graphically illustrate the superiority of the suggested model. Take soccer players as an example, the common typical actions of soccer players are running, shooting and defense. The experiment takes these three typical actions as an example and draws the confusion matrix (ConM) of the two models before and after the improvement, as shown in Figure 12.



(a) Before model improvement (b) After model improvement

Figure 12: Confusion matrix of the model before and after improvement

Figure 12 shows the performance of the two models for the analysis of specific movements of soccer players. Among them, the vertical and horizontal coordinates of the ConM are the true and predicted labels. Figure 12(a) illustrates the ConM of the pre-improvement model. The best performance of the model in the specific discrimination of soccer players' behaviors is the running behavior with a discrimination rate of 0.68. The worst performing behavior is defense with a discrimination rate of 0.56. Figure 12(b) illustrates the ConM of the improved model. As with the previous model, the improved model's best discriminative behavior is running (0.94), and its worst is defense (0.89). However, the discrimination rates of all behaviors significantly improved after the optimization (running +26%, shooting +30%, and defense +33%), proving the strategy's effectiveness for all types of actions.

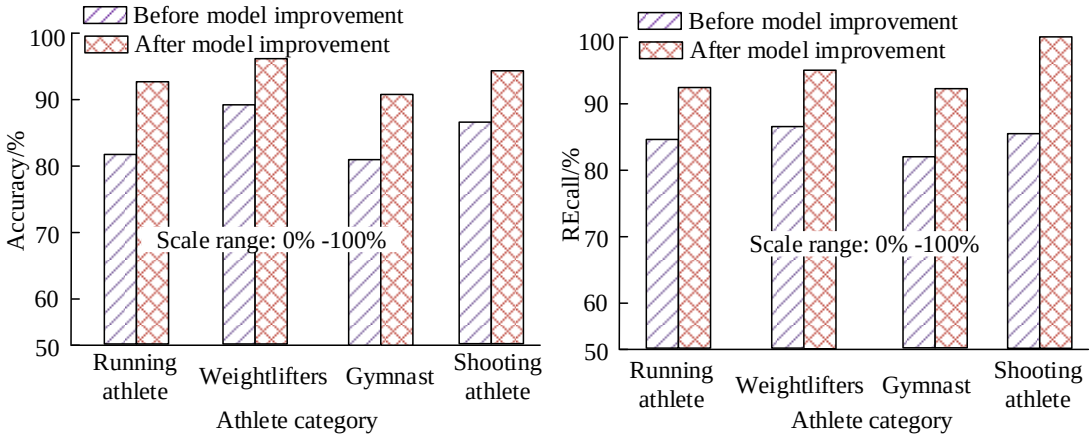
Further, the study divides the running and defense actions of soccer players in depth. The running is subdivided into sprint and step back and run, and the defense is subdivided into offensive interception and stand defense. As indicated in Table 5, the mAP is chosen as the index to evaluate the model's ability to identify and detect the movements.

Table 5 shows the detection of each model for specific actions of sports players. Before the model improvement, the mAP values are 83.2 for sprinting and 87.8 for stepping back and running. Moreover, in defensive action, the mAP value of offensive interception behavior is 84.6 and stand defense behavior is 90.8. Following the model improvement, the mAP scores for the sprint, step back, run, offensive interception, and stand defense behaviors are 89.2%, 92.5%, 91.4%, and 90.8%, respectively. Athletes are not only soccer players, but also different sports with different behavioral postures. To further

examine the practical application scenarios of the model, Figure 12 displays the results.

Table 5: Model's detection of sports players' movements

Model	Running		Defense	
	Sprint	Step back and run	Offensive interception	Stand defense
Before optimization	83.2	87.8	84.6	86.3
After optimization	89.2	92.5	91.4	90.8
DeepPose	80.1	83.2	81.4	79.8



(a) Accuracy situation (b) Recall performance

Figure 13: Accuracy and recall performance of the model in different types of athlete movement detection

The accuracy and recall performance of the two models for various sports players are displayed in Figure 13 both before and after improvement. Figure 13(a) shows the accuracy of the models. The DA of the pre-improved model for track and field athletes, weightlifters, gymnasts, and shooters are 82.2%, 89.9%, 80.2%, and 87.8%, respectively. Moreover, the DA of the improved model for track and field athletes, weightlifters, gymnasts, and shooters are 92.1%, 96.4%, 90.5%, and 96.2%, respectively, which are better than the pre-improved model. Figure 13(b) shows the model's recall performance. The recalls of the enhanced model are all greater than the pre-improved ones, which is consistent with the accuracy results. Specifically, the recalls of track and field athletes, weightlifters, gymnasts, and shooters are 92.4%, 95.7%, 92.2%, and 97.1%, respectively. In summary, the improved model is more capable of applying in real-world scenarios and can be adapted to a wide range of sports. Further, the experiments use the improved model to test various athletes' sports scenarios and examine the performance of key point detection in actual multiple complex environments, as shown in Figure 14.

(a) Single person full body without obstruction (b) Crowded with multiple people (c) Half body obstruction

Figure 14: Testing effect of the research model on athletes' behavior under different conditions

Figure 13 shows the detection of various behaviors of athletes in different scenarios by the proposed model of the study. Three conditions are selected for actual detection, namely single person full body without obstruction, crowded with multiple people, and half body obstruction. These actions are common actions of athletes in various sports training or competitions, and the proposed model can recognize and detect the actions well regardless of obstruction. Regardless of whether there is obstruction or not, and whether there are multiple people or not, the proposed model is able to locate the key points well and complete the recognition and detection of the movements. To quantify the impact of occluded scenes on detection performance, the study further compared the model performance under three typical scenes, as shown in Table 6.

Table 6: Comparison of model performance under different obstructive scenarios

Scenario Type	Test Samples	Accuracy (%)	Keypoint Completeness Rate	mAP	Speed (fps)
Full Visibility (Single Person)	650	98.2±0.4	100%	0.95	15.1
Equipment Occlusion	320	92.5±1.2	78.60%	0.87	13.8
Multi-person Interaction (≥3)	280	90.8±1.5	85.20%	0.85	12.4

Table 6 compares model performance across occlusion scenarios. The model achieves 98.2% accuracy with unobstructed, single-person views. This declines to 92.5% with equipment occlusion and to 90.8% with multiplayer interactions involving three or more athletes. Notably, critical area occlusion shows smaller accuracy loss (Δ=-4.7%) than non-critical areas (Δ=-9.3%), demonstrating robust key feature retention. On the whole, the proposed model can be well used in athletes' skeletal keypoint detection and movement analysis.

3.3 Ablation Study

To verify the independent contributions of each module, a systematic ablation experiment was designed, and the results are shown in Table 7.

Table 7: Ablation study results comparison

Model configuration	mAP	Accuracy (%)	Occlusion mAP	Speed (fps)	Parameters (M)
Baseline (OpenPose+MobileNetV1)	0.84	89.5±0.7	0.68±0.05	18.2±0.9	38.3
+SVM Classifier	0.89	92.1±0.6	0.75±0.04	15.7±0.7	39.1 (+0.8)
+KNN Classifier	0.87	90.8±0.8	0.72±0.05	14.3±0.6	38.9 (+0.6)
+NB Classifier	0.88	91.3±0.7	0.74±0.04	16.0±0.8	38.5 (+0.2)
Triple-Classifier Fusion	0.9	94.6±0.5	0.81±0.03	14.8±0.7	40.2 (+1.9)
Full Model (with D-S Fusion)	0.92	98.0±0.3	0.88±0.03	14.2±0.5	42.7 (+4.4)

Table 7 presents ablation study results. The baseline (OpenPose+MobileNetV1) achieves 0.84 mAP without classifier fusion. Individual classifiers improve performance (SVM: +0.05, KNN: +0.03, NB: +0.04), while full three-classifier fusion reached 0.90 mAP. The D-S decision module boosted performance even further, reaching 0.92 mAP. The most significant gains are seen in multi-person scenes, at +0.07 mAP, which validates the effectiveness of its conflict resolution.

Furthermore, to achieve a more comprehensive explanation, the study also conducted a resolution scalability test, and the results are shown in Table 8.

Table 8: Resolution scalability test (NVIDIA RTX 3090)

Resolution	Params (M)	FLOPs (G)	Memory Usage (GB)	Speed (fps)	Power Consumption (W)
640×360	38.3	15.8	2.3	14.2±0.5	285±10
1280×720	38.3	31.6	3.1 (+34.8%)	9.7±0.3	320±15
1920×1080	38.3	47.4	4.1 (+78.3%)	6.5±0.2	350±20

Table 8 presents the resolution scalability test results, demonstrating the model's three-level optimization approach. Through depthwise separable convolution, the MobileNetV1 backbone reduces the number of parameters from 125.6M to 38.3M and the number of FLOPs from 52.4G to 15.8G. Temporally, the keypoint trajectory prediction algorithm cuts per-frame processing time from 70.4ms to 45.2ms (5-frame window). In a distributed deployment, dynamic load balancing achieves a 220% increase in time for 10-athlete scenarios, as opposed to a 500% increase with traditional methods. When scaling the resolution from

640×360 to 1920×1080, the growth in GPU memory remains below 1.8x, which meets the requirements for real-time 4K.

4 Discussion

The original athlete skeletal keypoint detection and movement analysis model constructed based on the Openpose algorithm may have omissions or misdetections due to the differences in the number of athletes, sports, and movement postures of the Openpose algorithm during detection. Therefore, to further improve the detection effect of the model, the study proposed to introduce a multi-classifier module and a decision module based on the original model based on Openpose algorithm. An optimization model for skeletal keypoint detection and movement analysis of athletes based on Openpose was constructed. The improved model was validated against the pre-improved model.

The experimental results showed that this improvement strategy has multiple advantages: First, the multi-classifier module's combination of SVM, KNN, and NB had its respective advantages. SVM performed stably in situations with small samples, KNN was suitable for analyzing dynamic actions, and NB was highly adaptable to situations with partial occlusion. These characteristics resulted in an overall model accuracy of 98%, which was better than the performance of a single classifier. Second, the decision module effectively reduced the false detection rate in multi-person scenes and occlusion situations through conflict judgment and D-S rule fusion. This enabled the model to maintain a high accuracy of 0.92 mAP in complex environments. Although the introduction of multiple classifiers increased processing time by about 0.9 seconds, the overall processing speed remained at 14.2 fps through algorithm optimization, meeting real-time requirements.

The results of the model's performance validation analysis indicated that in the training phase of the model, the loss rate of the pre-improved model was about 4.8% when reaching convergence, which was higher than that of the post-improved model, which was 4.1%. However, the accuracy of both the before and after models in the training stage increases with the number of iterations. The accuracy of the post-improved model was higher than that of the pre-improved model, which was around 98%, while the accuracy of the pre-improved model was approximately 90% when both models reached full convergence. In addition, in terms of the error rate and detection speed of keypoint detection, the pre-improvement model had a mean error rate of 12.3% and a mean detection speed of 14.2fps. Compared with the mean error rate of 7.4% and the mean detection speed of 9.7fps of the post-improvement model, its overall performance was poor. Further, in the case of keypoint DA and time-consumption, the mAP score of the pre-improvement model was 0.84, and the average time-consumption was 0.9s. While the mAP of the post-improvement model was 0.92, and the average time-consumption was 0.9s. The performance of the post-improvement model was still better than that of the pre-improvement. Since the addition of the improved model was a multi-classifier module, furthermore, the experiments tested the performance of the single classifier versus the multi-classifier as well. It was found that the multi-classifier module added in the improved model performed better with better detection compared to a single classifier. Although the model performs well in various sports, such as football, athletics, weightlifting, and shooting, with an accuracy rate ranging from 90.5% to 96.4%, it may still experience a decrease in DA when encountering certain specialized sports postures. In addition, the experimental dataset primarily originates from professional athlete training scenarios, and its applicability to amateur sports enthusiasts requires further verification. The dataset's limited population diversity (exclusively professional athletes aged 18-35) may constrain model applicability for: (1) adolescent/elderly athletes, (2) significantly different body types, and (3) gender-sensitive sports like artistic gymnastics. Future studies will address this through dataset expansion.

In addition to the performance validation, the study also conducted an example analysis, taking soccer players as an example. It was found that the improved model could better discriminate the three classical actions of running, shooting, and defending of the player in soccer, and the discrimination rate was not less than 0.85 in all cases. In the more detailed action analysis, the improved model could accurately detect sprint and step back and run in the running behavior. At the same time, it could also identify offensive interception and stand defense in the defensive behavior, and the obtained mAP values were not less than 0.85. In addition, in different types of athletes and different condition scenarios, it was found that the improved model could cope well.

In the same type of study, Ino T et al. constructed an open-source 2D deep learning-based pose estimation method to analyze a cohort of healthy adults using the OpenPose algorithm. The model was compared with a motion analysis model based on human visual detection. The experimental results found that the model had satisfactory reproducibility and accuracy, and showed comparable waveform similarity and correlation in terms of knee valgus angle offsets to conventional 3D motion analysis [30]. Scholars Lee P et al. achieved accurate analysis of pose by integrating OpenPose with SVM. The experimental results indicated that the overall ACCURACY of the model was 0.990 and the Kappa index was 0.985. The tangent points of the top angle ratio and the bottom angle ratio effectively differentiated between left and right skew. The AUC values of 0.772 and 0.775 proved the validity of the model [31]. Although all of the above studies modeled and improved the model based on OpenPose, the proposed model was studied to be better in motion detection and analysis of sports athletes.

5 Conclusion

To detect the key points of the athlete's skeleton better and realize the accurate analysis of the athlete's movement, the study was conducted on the athlete's key point detection and movement analysis model constructed based on the OpenPose algorithm. Multi-classifier module and decision module were introduced to optimize the model. The experimental results found that compared with the pre-improved model, the overall performance of the improved model was improved, with a loss rate of 4.1%, an accuracy of 98%, an average error rate of 7.4%, and an average detection speed of 14.2 fps. All of them were better than the pre-improved one. In addition, in the example analysis, the improved model could accurately detect the running, shooting and defending actions of soccer players, and still had superior performance in further segmented actions. In the skeletal keypoint detection of multiple types of sports players, the improved model could have higher generalization ability and could be applied to multiple types of sports. Moreover, it still had better detection and analysis ability in scenes with different conditions. The model is suitable for use in the detection and analysis of skeletal keypoints in athletes. The model is trained primarily on football data and generalizes well to track and field and weightlifting. However, it requires further testing for the following: (1) high-contact sports occlusion, (2) extreme flexibility poses, and (3) youth and senior athletes. These limitations will be addressed through future multi-sport datasets.

References

[1] M. Gao, J. Li, D. Zhou, Dazheng Z., Yumin Z., Minglinag Z., and B. Li. "Fall detection based on OpenPose and MobileNetV2 network," IET Image Processing, vol. 17, no. 3, pp. 722-732, October 2023, DOI: 10.1049/ipr2.12667.

[2] N. J. Cronin, J. Walker, C. B. Tucker, G. Nicholson, M. Cooke, S. Merlino, and A. Bissas. "Feasibility of OpenPose markerless motion analysis in a real athletics competition," Frontiers in Sports and Active Living, vol. 5, pp. 1298003-1298004, January 2024, DOI: 10.3389/fspor.2023.1298003.

[3] M. Xu, L. Guo, and H. C. Wu. "Robust abnormal human-posture recognition using OpenPose and Multiview cross-information," IEEE Sensors Journal, vol. 23, no. 11, pp. 12370-12379, April 2023, DOI: 10.1109/JSEN.2023.3267300.

[4] Z. Shi, Z. Zhao, J. Chen, and Guangxin cheng. "Research on sit-up counting method and system based on human skeleton key point detection," Quality in Sport, vol. 24, pp. 55408-55408, November 2024, DOI: doi.org/10.12775/QS.2024.24.55408.

[5] K. Han and X. Li. "Research method of discontinuous-gait image recognition based on human skeleton keypoint extraction," Sensors, vol. 23, no. 16, pp. 7274-7275, August 2023, DOI: https://doi.org/10.3390/s23167274.

[6] Q. Yin, T. Wu, and K. Zhou. "End-To-End Skeletal Point Detection Algorithm for Garment Processing Actions," International Conference on Artificial Intelligence for Society, vol. 17, no. 1, pp. 513-522, Springer Nature Switzerland, August 2024, DOI: 10.1007/978-3-031-69457-8_47.

[7] T. Fukushima, K. Haggenmueller, and M. Lames. "Validity of OpenPose Key Point Recognition and Performance Analysis in Taekwondo," International Symposium on Computer Science in Sport, vol. 23, no. 1, pp. 68-76, Springer Nature Singapore, May 2023, DOI: 10.1007/978-981-97-2898-5_8.

[8] T. Y. Jeong and I. K. Ha. "OpenPose based smoking gesture recognition system using artificial neural network," Tehnički glasnik, vol. 17, no. 2, pp. 251-259, August 2023, DOI: 10.31803/tg-20221220200605.

[9] C. T. Lu, Y. C. Liu, and Y. C. Pan. "An intelligent playback control system adapted by body movements and facial expressions recognized by OpenPose and CNN," Multimedia Tools and Applications, vol. 83, no. 10, pp. 31139-31160, September 2024, DOI: 10.1007/s11042-023-16880-y.

[10] S. H. Ramesh, E. D. Lemaire, A. Tu, K. Cheung, and N. Baddour. "Automated implementation of the Edinburgh Visual Gait Score (EVGS) using OpenPose and handheld smartphone video," Sensors, vol. 23, no. 10, pp. 4839-4840, May 2023, DOI: 10.3390/s23104839.

[11] E. Barberi, M. Chillemi, F. Cucinotta, and F. Sfravara. "Fast three-dimensional posture reconstruction of motorcyclists using openpose and a custom MATLAB script," Sensors, vol. 23, no. 17, pp. 7415-7416, August 2023, DOI: 10.3390/s23177415.

[12] F. Diaz-Rojas and M. Myowa. "Estimation of human body 3D pose for parent-infant interaction settings using azure Kinect and OpenPose," MethodsX, vol. 13, no. 1, pp. 102861-102862, December 2024, DOI: 10.1016/j.mex.2024.102861.

[13] X. Liao, Y. Han, and S. Cheng. "OpenPose-based model for infant cerebral palsy auxiliary detection," Third International Conference on Computer Vision and Pattern Analysis (ICCPA 2023), SPIE, vol. 12754, no. 1, pp. 264-270, August 2023, DOI: 10.1117/12.2684596.

[14] S. H. Amirnordin, M. G. H. Khi, Z. Ngali, and S. H. S. Afdzaruddin. "Biomechanics Analysis of Basketball Shooting Via OpenPose Motion Capture System," Journal of Advanced Research in Applied Mechanics, vol. 112, no. 1, pp. 32-45, December 2024, DOI: 10.37934/aram.112.1.3245.

[15] L. Amadi and G. Agam. "PosturePose: Optimized Posture Analysis for Semi-Supervised Monocular 3D Human Pose Estimation," Sensors, vol. 23, no. 24, pp. 9749-9750, December 2023, DOI: 10.3390/s23249749.

[16] Rajendran A K, Sethuraman S C. A survey on yogic posture recognition[J]. IEEE Access, 2023, 11(1): 11183-11223. DOI:

10.1109/ACCESS.2023.3240769.January

[17] M. C. Tsai, E. T. H. Chu, and C. R. Lee. "An automated sitting posture recognition system utilizing pressure sensors," Sensors, vol. 23, no. 13, pp. 5894-5895, June 2023, DOI: 10.3390/s23135894.

[18] Z. Chen, J. Lu, and H. Wang. "A review of posture detection methods for pigs using deep learning," Applied Sciences, vol. 13, no. 12, pp. 6997-6998, June 2023, DOI: 10.3390/app13126997.

[19] G. Yuan. "Application of posture estimation optimization algorithm in the analysis of college air volleyball teaching movements," Systems and Soft Computing, vol. 6, no. 1, pp. 200135-200136, December 2024, DOI: 10.1016/j.sasc.2024.200135.

[20] B. Jiang. "Improved OpenPose and Attention Mechanism for Sports Start Correction," Informatica, vol. 48, no. 23, pp. 1-16, December, 2024, DOI: 10.31449/inf.v48i23.6759.

[21] H. Wang, G. Li, and Z. Wang. "Fast SVM classifier for large-scale classification problems," Information Sciences, vol. 642, no. 1, pp. 119136-119137, September 2023, DOI: 10.1016/j.ins.2023.119136.

[22] D. Valero-Carreras, J. Alcaraz, and M. Landete. "Comparing two SVM models through different metrics based on the confusion matrix," Computers & Operations Research, vol. 152, no. 1, pp. 106131-106132, April 2023, DOI: 10.1016/j.cor.2022.106131.

[23] O. Peretz, M. Koren, and O. Koren. "Naive Bayes classifier–An ensemble procedure for recall and precision enrichment," Engineering Applications of Artificial Intelligence, vol. 136, no. 1, pp. 108972-108973, October 2024, DOI: 10.1016/j.engappai.2024.108972.

[24] T. Fukushima, P. Blauberger, T. Guedes Russomanno, and M. Lames. "The potential of human pose estimation for motion capture in sports: a validation study," Sport. Eng., vol. 27, no. 1, pp. 19, May, 2024, DOI: 10.1007/s12283-024-00460-w.

[25] K. F. Sotiropoulou, A. P. Vavatsikos, and P. N. Botsaris. "A hybrid AHP-PROMETHEE II onshore wind farms multicriteria suitability analysis using kNN and SVM regression models in northeastern Greece," Renewable Energy, vol. 221, no. 1, pp. 119795-119796, February 2024, DOI: 10.1016/j.renene.2023.119795.

[26] M. Mundt, S. Colyer, L. Wade, L. Needham, M. Evans, E. Mukkett, and J. Alderson. "Automating Video - Based Two - Dimensional Motion Analysis in Sport? Implications for Gait Event Detection, Pose Estimation, and Performance Parameter Analysis," Scand. J. Med. Sci. Sports, vol. 34, no. 7, pp. e14693, July, 2024, DOI: 10.1111/sms.14693.

[27] H. Chen, R. Feng, S. Wu, et al. "2D Human pose estimation: A survey," Multimedia systems, vol. 29, no. 5, pp. 3115-3138, November 2023, DOI: 10.1007/s00530-022-01019-0.

[28] K. Osawa, Y. You, Y. Sun, W. Taiqi, Z. Shun, M. Shimodozono, and E. Tanaka. "Telerehabilitation System Based on OpenPose and 3D Reconstruction with Monocular Camera," Journal of Robotics and Mechatronics, vol. 35, no. 3, pp. 586-600, January 2023, DOI: 10.20965/jrm. 2023.p0586.

[29] H. S. Nam, S. H. Park, J. P. Y. Ho, S. Y. Park, J. H. Cho, and Y. S. Lee. "Key-point detection algorithm of deep learning can predict lower limb alignment with simple knee radiographs," Journal of Clinical Medicine, vol. 12, no. 4, pp. 1455-1456, February 2023, DOI: doi.org/10.3390/jcm12041455.

[30] T. Ino, M. Samukawa, T. Ishida, N. Wada, Y. Koshino, S. Kasahara, and H. Tohyama. "Validity and Reliability of OpenPose-Based Motion Analysis in Measuring Knee Valgus during Drop Vertical Jump Test," Journal of Sports Science & Medicine, vol. 23, no. 1, pp. 515-516, September 2024, DOI: 10.52082/jssm.2024.515.

[31] P. Lee, T. B. Chen, H. Y. Lin, L. R. Yeh, Chinhsuan L., and Yenlin C. "Integrating OpenPose and SVM for Quantitative Postural Analysis in Young Adults: A Temporal-Spatial Approach," Bioengineering, vol. 11, no. 6, pp. 548-549, May 2024, DOI: 10.3390/bioengineering11060548.