A Motion Capture Framework for Table Tennis Using Optimized SVM and AdaBoost Algorithms

Feng Shi
Institute of Physical Education, Hunan University of Science and Engineering Yongzhou 425199, China
E-mail: shifeng198912@126.com

Keywords: support vector machine; adaptive boosting algorithm; table tennis; motion capture; collection

Received: July 26, 2024

Table tennis requires high technical and tactical skills. The application of motion capture technology can improve athletes' training effectiveness and competition strategies. To further improve the collection and capture efficiency of table tennis sports data, a high-precision optical motion capture system, and inertial measurement unit sensors are first used to collect table tennis sports data. Data preprocessing is carried out using action windows and sliding windows. Secondly, a support vector machine classifier optimized with a Gaussian radial basis kernel function is used for training. The sample weights are updated based on the feature classification results in each iteration. Finally, the adaptive boosting algorithm is combined with it to propose a new type of table tennis motion capture model. The experimental results showed that the optimized model achieved the highest classification accuracy of 96% with the best kernel parameters and a normalization factor of 1.0. The model's motion capture errors ranged from 1.5% to 9.7% and had the shortest runtime of 7.66 seconds. In addition, the model achieved the highest capture accuracy of 93%, 92%, 91%, and 90% for the four motions of forehand kill, forehand putt, backhand kill, and backhand putt, respectively. This demonstrates significant advantages in terms of accuracy and computational efficiency. Therefore, the proposed model not only improves the accuracy and efficiency of motion capture but also performs well in terms of resource consumption. This model has high practical application value and can provide a new reference for technological development in this field.

Povzetek: Narejena je analiza gibanja pri namiznem tenisu z optimiziranimi algoritmi SVM in AdaBoos za namene analize igralnih tehnik ter treningov.

1 Introduction

Table tennis is a sport that requires high technical and tactical skills. Athletes need to have quick reactions and precise hitting movements during competitions [1]. As technology advances, motion capture technology is gradually being applied to table tennis. This can be used to analyze athletes' technical movements, enhance training effectiveness, reduce injury risks, and refine competition strategies. By capturing athletes' hitting movements, various details of their movements can be analyzed in detail, including hitting angle, swing speed, body posture, etc. [2]. These data can help coaches and athletes identify deficiencies in their movements and make targeted improvements. Li et al. believed that traditional motion capture methods might be more time-consuming and less efficient in analyzing the rotation trajectory of table tennis balls. To this end, the research team proposed a novel table tennis motion capture method by combining spatial upsampling and reconstruction encoders with long short-term memory networks. The table tennis movement trajectory under this method achieved an accuracy of 96.5% and a prediction speed of 15 seconds [3]. Wu et al. proposed a novel table tennis pose estimation method to improve the processing effect of pose estimation technology in table tennis by combining graphics processor optimization and OpenPose. This algorithm performed well in estimating table tennis players' posture in videos and was more feasible compared to before improvement [4]. Ren et al. explored a unique method of motion capture for table tennis players. After combining graph convolutional neural networks, they proposed a novel method for detecting incorrect postures and movements of table tennis players. This method performed well on a large number of existing datasets, with the highest accuracy of 96.4%. However, this method relied on a large amount of data for training. The initial data acquisition cost was relatively high [5]. Therefore, a large number of table tennis motion capture methods have demonstrated superior performance. However, there are still problems such as complex structure, large training quantity, difficulty in parameter adjustment, and the need to improve capture accuracy. Support Vector Machine (SVM) classifies data points by finding the optimal hyperplane in high-dimensional space, which can handle high-dimensional data and complex relationships. Adaptive Boosting (AdaBoost) generates a series of weak classification algorithms and then combines them according to certain rules to form a strong classification algorithm. In recent years, these two methods have been widely used in many motion capture scenarios. For example, Duan et al. added SVM to the built-in sensors of wearable devices. After data collection, classification and recognition were carried out to achieve the goal of capturing and recognizing human lower limb movements. The average recognition rate of this method after 10 cross-validations was 97.01% [6]. Feradov et al. attempted to develop a baseline detector for automatically detecting incorrect sitting posture, which combined SVM and time-domain parameters for activity movement integrity assessment. The average recognition classification rate of this detector in sitting posture detection was 98.4% [7]. In addition, Ding et al. found that it was increasingly difficult to evaluate human activity status by identifying fine-grained channel state information. To this end, they proposed a human motion perception model by combining AdaBoost and Internet of Things technology. This model achieved an average recognition accuracy of 96% for human activities in corridor environments [8]. Zhu et al. developed a human arm motion model suitable for remote operation, which integrated AdaBoost and inertial measurement units. This method could capture the position and orientation information of the end of the arm for transformation regression modeling. This model demonstrated an average angle error of 1.1 degrees in arm motion capture during on-site experiments, demonstrating good capture performance [9]. The results of the comparison of the indicators for each method are shown in Table 1.

Table 1: Comparison results of indicators for various methods.

Algorithm	Dataset	Accuracy (%)	Runtime (s)	Key Limitation	Reference
Spatial Upsampling + LSTM Encoder	Custom Table Tennis Motion Trajectory Dataset	96.5	15.33	Complex structure, high computational resource demand	[3]
OpenPose + GPU Optimization	Video-based Table Tennis Player Pose Estimation Dataset	92.7	12.18	Mainly applicable to video data, limited generalizability	[4]
Graph Convolutional Network	Large-Scale Table Tennis Player Pose Dataset	96.4	18.24	Requires large amounts of training data, high initial data acquisition cost	[5]
SVM (for lower limb motion capture)	Wearable Device Sensor Dataset	97.1	10.03	Limited to lower limb motion, restricted application scenarios	[6]
SVM + Temporal Parameters	Posture Detection Dataset	98.4	8.89	Mainly for posture detection, limited motion capture types	[7]
AdaBoost + IoT Technology	Human Activity Recognition in Corridor Environment Dataset	96.6	14.17	Depends on IoT-based environment limitations	[8]
AdaBoost + Inertial Measurement Unit	Arm Motion Dataset	94.3	11.56	Limited to arm motion, angular error of 1.1 degrees	[9]

In summary, existing table tennis motion capture methods have demonstrated good performance. However, there are also some shortcomings. The main problems include complex model structure, high computational resource requirements, large training data demands, and difficulty in parameter adjustment. To solve these problems, the study adopts a hybrid approach based on SVM and AdaBoost to improve the accuracy and efficiency of table tennis motion capture. First, the motion data are collected using a high-precision optical capture system and inertial measurement unit sensors. Then, the key feature vectors are generated by data preprocessing. Subsequently, a Gaussian Radial Basis Function (GRBF)-optimized SVM is employed for preliminary classification. Thereafter, the sample weights are adaptively adjusted by the AdaBoost algorithm to enhance the overall performance of the classifier, thereby facilitating efficient table tennis motion capture. The research innovation lies in the combination of SVM and AdaBoost. This significantly improves motion capture models' classification accuracy, reduces computational resource requirements, and enhances model robustness. The research contribution lies in providing more precise and efficient training tools for table tennis players, while promoting the application and development of motion capture technology in sports.

2 Methods and materials

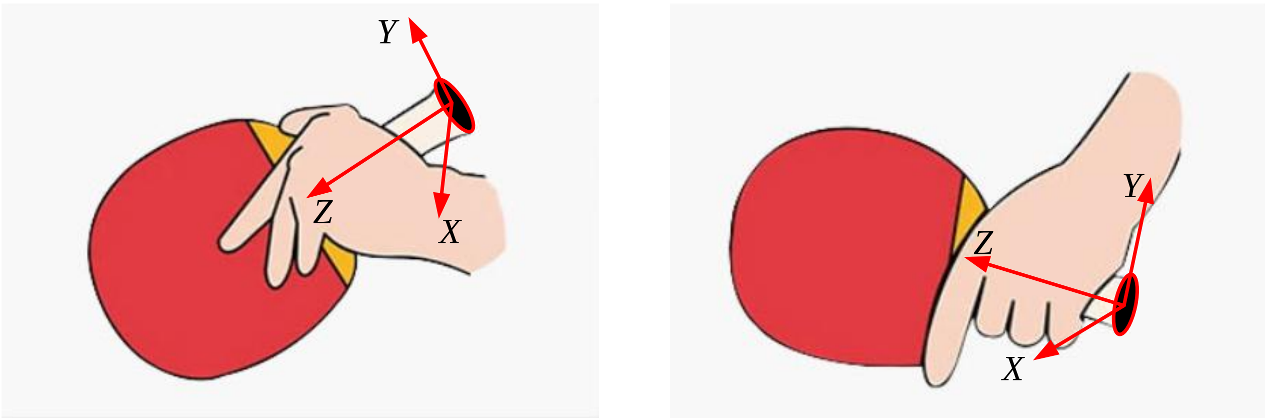
The study first uses high-precision optical motion capture systems and inertial measurement unit sensors to record athletes' swing motion data to improve the accuracy and efficiency of table tennis motion capture. Key feature vectors are generated through denoising, standardization, and feature extraction. Secondly, the sliding window technique is utilized to segment these data and extract segments containing swinging movements. Then, by initializing the sample weights, a GRBF-optimized SVM classifier is utilized to train these data. The sample weights are updated based on the classification results in each iteration. Finally, a table tennis motion capture model that integrates SVM and AdaBoost is proposed.

2.1 Table tennis sports data collection and preprocessing

In table tennis, data collection and preprocessing are key steps in achieving efficient motion capture and accurate analysis. To ensure the accuracy and consistency of the collected data, a high-precision optical motion capture system OptiTrack Prime 13 is studied. This system is equipped with 12 cameras with a resolution of 1280×1024 pixels, a frame rate of 240 frames per second, a capture accuracy of less than 0.5 millimeters, and a working range of 5m×5m×3m [10-11]. In addition, the inertial measurement unit sensor Xsens MVN Awinda is utilized to supplement data acquisition. This sensor consists of 17 units, with a gyroscope range of ±2000 degrees per second, an accelerometer range of ±16G, a magnetometer range of ±8 Gauss, and a data transmission rate of 100 Hz [12-13]. The high-speed camera Phantom VEO 640S has a resolution of 2560×1600 pixels, a frame rate of 1000 frames per second, and an exposure time of 1 microsecond. To standardize the collection of table tennis movements, two types of grip methods have been defined in Figure 1.

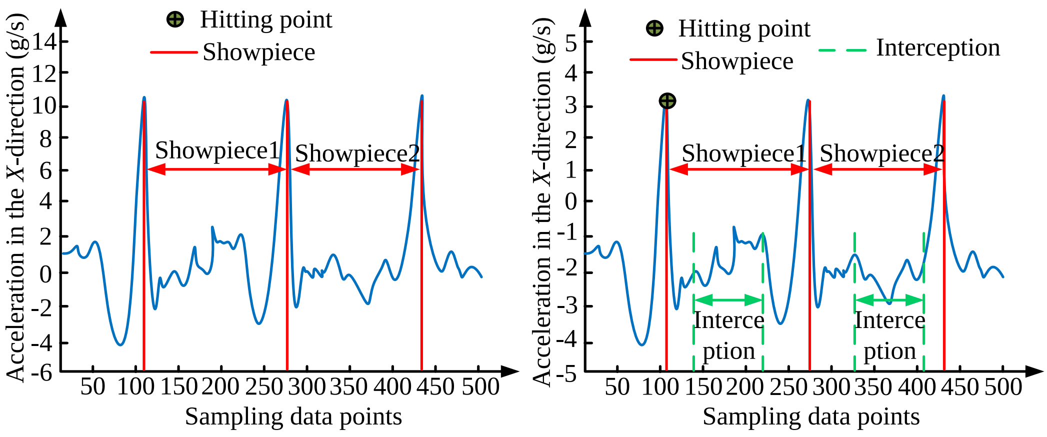
Figure 1 (a) is a schematic diagram of the first type of table tennis racket grip posture. Figure 1 (b) is a schematic diagram of the second type of table tennis racket grip posture. The area below the grip of a table tennis racket is defined as the original point. Space is established along three axes. When the X-axis points towards the inside of the hand, it is defined as the first type of grip. When the X-axis points to the outside of the hand, it is defined as the second type of grip. Repeated data collection is conducted on commonly used basic actions. After completion, to enhance the action analysis accuracy and efficiency, the study introduces action windows and sliding windows for data processing and analysis. Compared to other methods, action windows and sliding windows have simpler operations, richer detail capture, and higher accuracy in action processing [14]. Figure 2 shows the segmentation of the action window and sliding window.

Figure 2 (a) shows the data segmentation of the action window. Figure 2 (b) shows the data segmentation of the sliding window. In Figure 2 (a), the sample point between the two data peaks is labeled as an action. The peak point is the hitting point. According to this logic, the data



(a) Grip mode 1 (b) Grip mode 2

Figure 1: Schematic diagram of the two types of grips.



(a) Action window (b) Sliding window

Figure 2: Segmentation of the action window and sliding window.

sampling frequency is set to 50Hz. After multiple data collections, 30 sets of window data are established with a ping pong swing action of 0.7 seconds. The first 14 groups are one segment, the 15th group is the hitting point, and the last 15 groups are the second segment. In Figure 2 (b), the swinging motion of table tennis has discontinuity. Therefore, real-time data are captured using a sliding window, that is, window data are captured before and after the hitting point. To classify table tennis movements more accurately, this study introduces a MeACC threshold for acceleration to make judgments. The acceleration judgment is represented by equation (1) [15].

$$Accelerations = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{3}\left|x_{ij}\right| \tag{1}$$

In equation (1), x_{ij} refers to the data in row i and column j. $Accelerations$ refers to the judgment criteria. n refers to window specifications. The criterion for determining angular velocity is represented by equation (2).

$$\begin{cases} S_{\omega\,\mathrm{var}} = \dfrac{1}{n-1}\sum\limits_{i=1}^{n}(\omega_i - \dfrac{1}{n}\sum\limits_{i=1}^{n}\omega_i)^2 \\ \omega = \sqrt{\omega_x^2 + \omega_y^2 + \omega_z^2} \end{cases} \tag{2}$$

In equation (2), $S_{\omega\,\mathrm{var}}$ refers to the criterion for determining angular velocity. ω_i refers to the angular velocity value of the i th data. ω_x, ω_y, and ω_y refer to the angular velocities of the x, y, and z, respectively. ω refers to the sum of angles in three directions. If the acceleration of the window data is greater than $MeAcc$, the data in the window are defined as the hitting action. On the contrary, it is a non-hitting action. In addition, if the angular velocity of the window data is greater than $S_{\omega\,\mathrm{var}}$, it is also defined as a hitting action. On the contrary, it is a non-hitting action. Through the above methods, the real-time collected table tennis movement data have been reasonably divided. This not only improves the effectiveness of the original table tennis motion data, but also facilitates subsequent motion capture and recognition. This can improve the training quality and competition performance of table tennis players.

2.2 Table tennis swing action classification based on SVM

After collecting and optimizing table tennis movements data through windowing preprocessing, this study attempts to introduce SVM to recognize and classify different table tennis swing postures. Table tennis is a sport that requires extremely high technical and tactical skills, covering a variety of basic movements. At present, there are many basic movements in table tennis, including serving, receiving, pulling, dunking, drawing, blocking, chopping, rubbing, etc., as well as comprehensive movements such as side pulling, picking, and rotation control [16-17]. To reduce the dimensionality of the data and facilitate subsequent action analysis, the study only considers the frequent use of forehand kills, forehand pushes, backhand kills, and backhand pushes. When faced with such high requirements for sports action classification, SVM performs well in handling high-dimensional data and small sample datasets compared to traditional classification methods, which has strong generalization ability [18-19]. Figure 3 is a schematic diagram of a linearly separable SVM [20].

In Figure 3, a simple dataset is divided into two categories by a dashed line: a red circular dataset and a blue square dataset. This dashed line is the decision boundary of SVM. The data points closer to the decision boundary are called support vectors. To define this decision boundary, generally, two parallel lines with intervals of d_1 and d_2 are chosen. By infinitely differentiating these two lines, the most realistic decision boundary is approximated. This process is represented by equation (3).

$$D = \{(x_i, y_i) | i = 1, 2, \cdots, n\}, x_i \in R, y_i \in \{1, -1\} \tag{3}$$

In equation (3), D refers to a complete dataset. x_i and y_i correspond to two 2D data points on the dataset. The calculation for correctly classifying the dataset into straight lines is represented by equation (4).

$$wx + b = 0 \tag{4}$$

In equation (4), both w and b belong to the weights and biases of the samples in the given dataset. x represents the feature space. At this point, the two parallel lines and the decision boundary's distance is represented by equation (5).

$$\begin{cases} d1 = \dfrac{|wx+b|}{\|w\|} = \dfrac{1}{\|w\|} \\ d2 = \dfrac{2}{\|w\|} \end{cases} \tag{5}$$

The related variables in equation (5) are explained in a consistent manner. However, for problems that cannot be classified by a 2D straight line, they need to be mapped to a high latitude environment before classification. This method generally converts data from low latitudes to high latitudes in the form of vector points. Figure 4 is a schematic diagram of a linearly inseparable SVM [21-22].

In Figure 4, SVM can map 2D data to 3D space. In high-dimensional environments, data are easier to separate, effectively partitioning the original data. However, this process usually requires the assistance of kernel functions. Choosing the appropriate kernel function can improve mapping and classification efficiency as well as reduce computational complexity. Therefore, in calculations, the focus should be on the kernel function rather than the mapping itself. Common kernel functions include linear kernel, polynomial kernel, GRBF, and Sigmoid kernel [23-24]. Considering the diverse types and complex characteristics of table tennis swing movements, combined with the sample size, a new model for classifying table tennis swing movements is proposed by combining SVM with GRBF. Figure 5 shows the model operation process.

In Figure 5, the entire process includes 6 steps. Firstly, initial data are collected. Various sensors and camera devices are utilized to record the raw data of table tennis

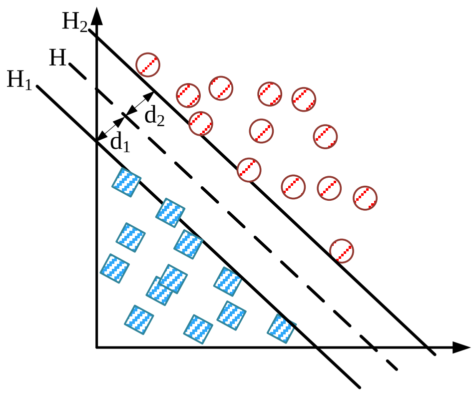


Figure 3: Schematic diagram of linearly separable SVM.

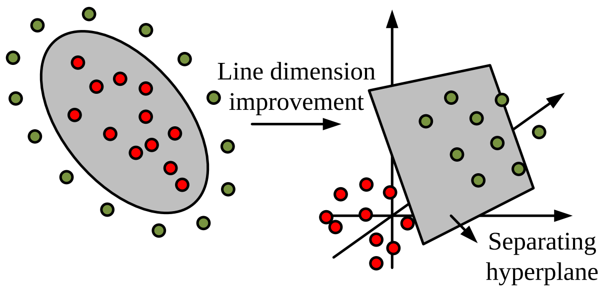


Figure 4: Schematic diagram of linearly inseparable SVM.

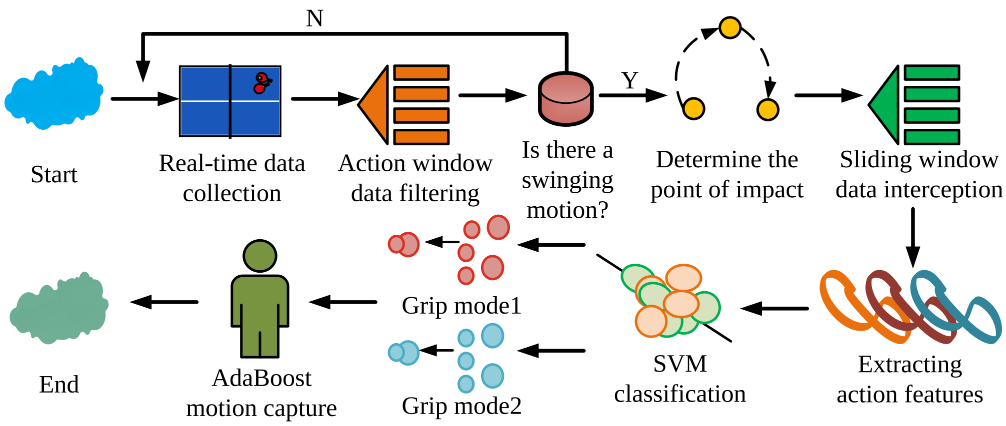


Figure 5: Flowchart of the new table tennis swing movement classification model.

swing movements. Then, sliding window technology is utilized to segment the collected data and extract segments containing swinging movements. Thirdly, the hitting point is determined and the corresponding swing data are captured. Next, these action data are preprocessed, noise is removed, and these data are standardized. Then the features of these preprocessed data are extracted to generate feature vectors for classification. Finally, the feature vectors are input into the trained SVM model for classifying swing actions. The acceleration and angle characteristics of table tennis swing during this period are represented by equation (6).

$$\begin{cases} A(t)=(\frac{d^2x(t)}{dt^2},\frac{d^2y(t)}{dt^2},\frac{d^2z(t)}{dt^2}) \\ \theta(t)=\arctan(\frac{y(t)}{x(t)}) \end{cases} \tag{6}$$

In equation (6), $A(t)$ refers to the acceleration vector at time t. $\theta(t)$ refers to the racket angle at time. In addition, the classification decisions of GRBF and SVM are represented by equation (7).

$$\begin{cases} K(x_u,x_r)=exp(-\gamma\|x_u-x_r\|^2) \\ f(x)=sgn(\sum_{i=1}^{N}\alpha_i y'K(x_u,x_r)+b') \end{cases} \tag{7}$$

In equation (7), both x_u and x_r refer to the input feature vectors. $\|x_u-x_r\|$ refers to the Euclidean distance between two vectors. γ refers to a kernel parameter that can control the width of the kernel function. b' refers to a bias term. α_i stands for Lagrange multiplier. $K(x_u,x_r)$ stands for GRBF. y' refers to category labels that

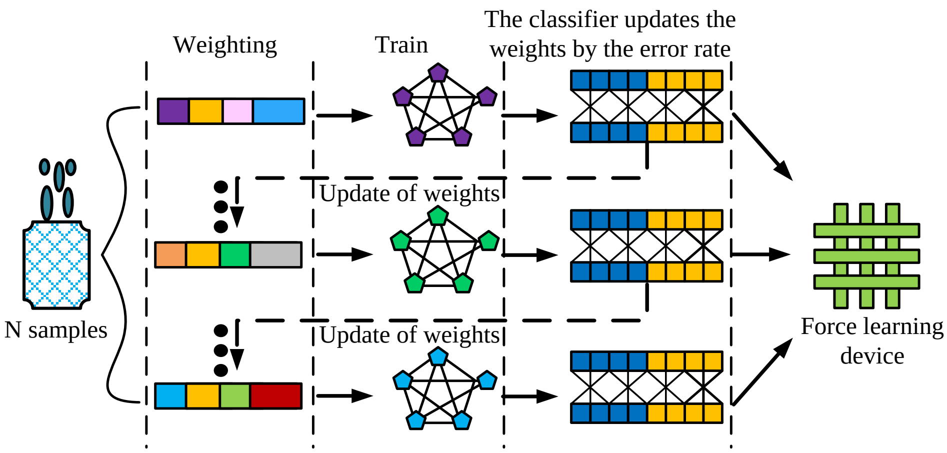


Figure 6: Schematic of how the AdaBoost algorithm works.

correspond to different swing movements, such as forehand pull and backhand pull.

2.3 Construction of table tennis motion capture model integrating SVM and AdaBoost algorithm

After classifying the swing posture of table tennis using SVM with GRBF, the study further introduces AdaBoost to enhance the motion capture effect. Boosting is a machine learning ensemble meta-algorithm that transforms weak learners into strong learners. Compared to other methods, AdaBoost has the advantage of adaptively adjusting sample weights and combining multiple weak classifiers in table tennis motion capture. This significantly improves classification accuracy and model robustness [25-26]. Its ability to automate feature processing and reduce overfitting makes it particularly suitable for handling high-dimensional and complex table tennis motion data. Figure 6 shows the principle of AdaBoost [27-28].

In Figure 6, starting from the initial dataset, each sample is assigned the same weight. Then the separator training, classification testing, and weight updating are repeated. After repeated multiple times, a new weak classifier is generated in each round [29-30]. Ultimately, the AdaBoost combines the weighted voting results of all weak classifiers to form a strong classifier, thereby improving the overall classification accuracy. The classification error is represented by equation (8).

$$\varepsilon_e = \sum_{e=1}^{N} w_e I(y_e \neq h_e(x_i)) \qquad (8)$$

In equation (8), ε_e refers to the weak classifier e's classification error. w_e refers to the i th sample's weight. $I(y_e \neq h_e(x_i))$ refers to the indicator function, which is 1 when y_e is not equal to $h_e(x_i)$. Otherwise, it is 0. y_e refers to e's actual category. $h_e(x_i)$ refers to e's prediction result for the i th sample. At this point, the weights of the weak classifier are represented by equation (9).

$$w_e = \frac{1}{2}\ln\left(\frac{1-\varepsilon_e}{\varepsilon_e}\right) \qquad (9)$$

In equation (9), w_e increases as ε_e decreases, indicating that the classifier's error rate is smaller and its proportion in the final classifier is larger. The weight update is represented by equation (10).

$$w_{e+1} \leftarrow \frac{w_e}{Z_e}\exp(-y_e h_e(x_i)) \qquad (10)$$

In equation (10), Z_e refers to a normalization factor. w_{e+1} refers to a weight value updated after $e+1$ times. After combining SVM with AdaBoost, the weight adjustment needs to be incorporated into optimization, represented by equation (11).

$$L(\alpha) = \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i y_e K(x_u, x_r) \quad (11)$$

In equation (11), $L(\alpha)$ represents a Lagrangian function. The final classifier after adding the sign function is represented by equation (12).

$$G(x) = sign(g(x)) = sign\left(\sum_{e=1}^{E} w_e h_e(x_i)\right) \quad (12)$$

In equation (12), $g(x)$ refers to a classifier composed of multiple weak classifiers combined. The combination of SVM and AdaBoost can adaptively adjust sample weights and enhance attention to difficult to classify samples. SVM is combined with AdaBoost for motion capture in table tennis. The adaptive weighting mechanism of AdaBoost and the powerful classification ability of SVM are utilized. A strong classifier is formed by combining multiple SVM weak classifiers and applied to the motion capture of table tennis. Figure 7 shows the model process.

In Figure 7, the entire table tennis motion capture includes three major modules, namely data acquisition and preprocessing, feature extraction and SVM training, as well as AdaBoost optimization and classification. Firstly, high-precision motion capture systems and inertial sensors are utilized to record table tennis players' swing motion data. Secondly, the collected data are denoised, standardized, and feature extracted to generate feature vectors for classification. These feature vectors include key features such as position, velocity, acceleration, and angle. Then, these sample data weights are initialized. In each iteration, an SVM classifier is trained using the

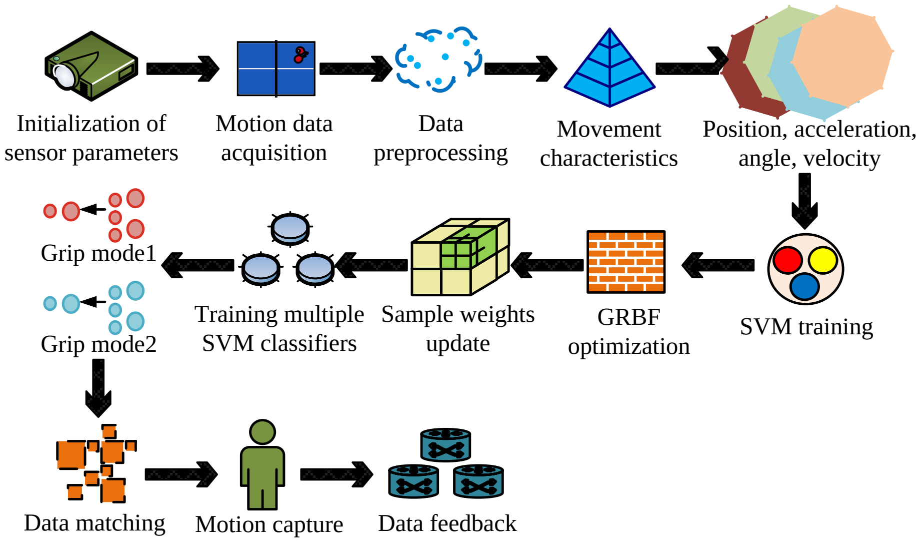


Figure 7: SVM-AdaBoost table tennis ball motion capture modeling process.

current sample weight w_e. GRBF and related parameters are selected for classifier performance optimization. After completion, these sample weights are updated based on the weak classifier's weights and the classification results.

After repeating multiple times, multiple SVM weak classifiers are trained until the predetermined number of weak classifiers or classification error meets the requirements.

3 Results

The study first established a suitable experimental environment to validate the new table tennis motion capture model. Secondly, the optimal parameters, namely kernel parameters and gauge factors, were determined. Ablation testing and comparative testing of similar methods were conducted. In addition, several types of basic table tennis movements were randomly selected for simulation testing. Quantitative and visual methods were utilized for comparative verification. The results show:

- Classification accuracy: the classification accuracy of the proposed model was studied up to 96% on the Table Tennis Stroke Dataset (TTSD) and 95.13% on the Table Tennis Trajectory Dataset (TTTD).

- Running time: The running time of the proposed model was 7.66 seconds on the TTSD dataset and 8.15 seconds on the TTTD dataset, which is significantly faster than the other compared models.

- Memory usage: The memory usage of the proposed model was studied to be 120MB and 125MB on TTSD and TTTD datasets respectively, which shows a high efficiency of resource utilization.

- Robustness: The motion capture errors of the proposed model ranged from 1.5% to 9.7% under different noise conditions, which outperforms other models such as SVM-KPCA and AdaBoost-RF.

- Multi-metric performance: The Precision, Recall, and F1 values of the proposed model were 95.79%, 94.58%, and 95.18% on the TTSD dataset, and 92.69%, 91.35%, and 92.02% on the TTTD dataset, respectively.

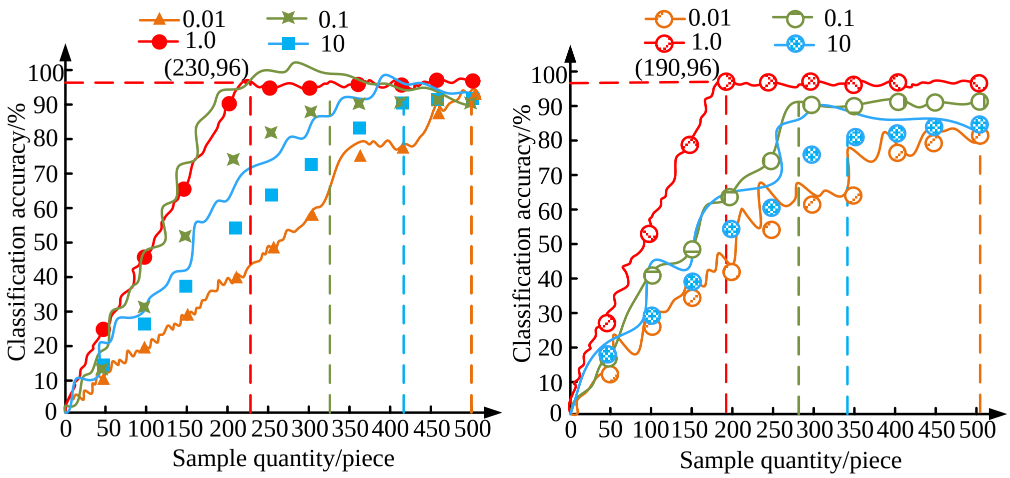
- Specificity: The specificity of the proposed model reached 92.75% in the TTSD dataset and 90.34% in the TTTD dataset, both of which are better than the other compared models.

- Real-time application potential: The performance of the proposed model in real-time sports training and motion capture analysis was outstanding. It is suitable for providing fast and accurate feedback in real-time scenarios.

3.1 Performance testing of table tennis motion capture model

The CPU used in this experiment is Intel Xeon Gold 6230, 2.1 GHz, with 20 cores. The GPU is NVIDIA Tesla V100, 32 GB. The operating system is Ubuntu 20.04 LTS. The software environment is Python 3.8. GRBF's kernel parameter is 0.1. The maximum iteration is 1000 times. The weak classifier's iteration is 50 times. The learning rate is 0.005. The TTSD and TTTD are utilized as experimental test data sources. TTSD is a high-quality dataset containing various table tennis movements, mainly used for action recognition and classification research. This dataset contains 5000 table tennis action samples, covering common actions such as serving, dribbling, and dunking. TTTD is a dataset focused on table tennis trajectory analysis, widely used in sports science and computer vision research. This dataset contains 3000 table tennis trajectory samples, covering various hitting actions and rotation effects. The study first utilizes classification accuracy as an indicator to test the model performance under different kernel parameters γ and normalization factors Z_e to determine the optimal hyperparameter values. Figure 8 shows the test results.

Figure 8 (a) shows the model components under different γ. Figure 8 (b) shows the model components under Z_e. In Figure 8, as the test samples increase, γ is not necessarily bigger. On the contrary, the model performs the best in classification performance when γ is 1.0, with a maximum classification accuracy of 96% and



(a) Sub-kernel parameter testing (b) Sub-normalization factor test

Figure 8: Comparison of model performance tests with different kernel parameters and specification factors.

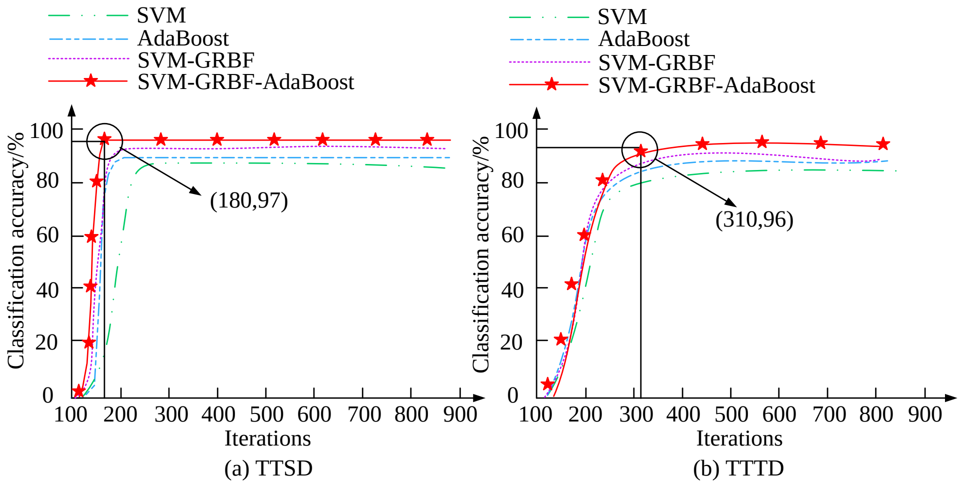


Figure 9: Motion capture models' ablation test results.

only 230 iterations. In addition, Z_e is similar to γ, with the highest classification performance of 96% and 190 iterations at 1.0. Upon investigation, both γ and Z_e control the influence range of a single training sample. The larger these values, the smaller the sample's influence range, making the model more complex and prone to overfitting. The smaller these values, the larger the sample's influence range, and the simpler this model will become, which is prone to underfitting. Therefore, the study selects 1.0 GG and 1.0 HH as fixed parameters for subsequent experimental testing. The study continues to validate the proposed new motion capture model through ablation testing, with training time as the indicator. Figure 9 shows the measurement results.

Figure 9 (a) shows the ablation test results of the motion capture model under TTSD. Figure 9 (b) shows the ablation test results of the motion capture model under TTTD. The influence of each component on the overall performance is determined by comparing the models with varying configurations, such as SVM without AdaBoost and SVM with distinct kernel functions, as illustrated in Figure 9. The SVM-GRBF-AdaBoost model achieves a classification accuracy of 97% after about 180 iterations and remains stable. In contrast, the SVM and AdaBoost models has lower classification accuracy and slower convergence. Meanwhile, when only SVM is used and AdaBoost is not combined, the classification accuracy decreases significantly. This indicates the importance of AdaBoost in dealing with difficult-to-classify samples and improving model robustness. In addition, the optimization improvement of each module in the SVM-GRBF-AdaBoost model is crucial and valuable. For example, the GRBF kernel function is selected for optimization in SVM, which improves the classification accuracy of SVM from a maximum of 85% to nearly 90%. The study continues to introduce methods similar to SVM and AdaBoost to explore the motion capture errors changes of various algorithms under different noise environments. They include Support Vector Machine-Kernel Principal Component Analysis (SVM-KPCA), Support Vector Machine-Genetic Algorithm (SVM-GA), and Adaptive

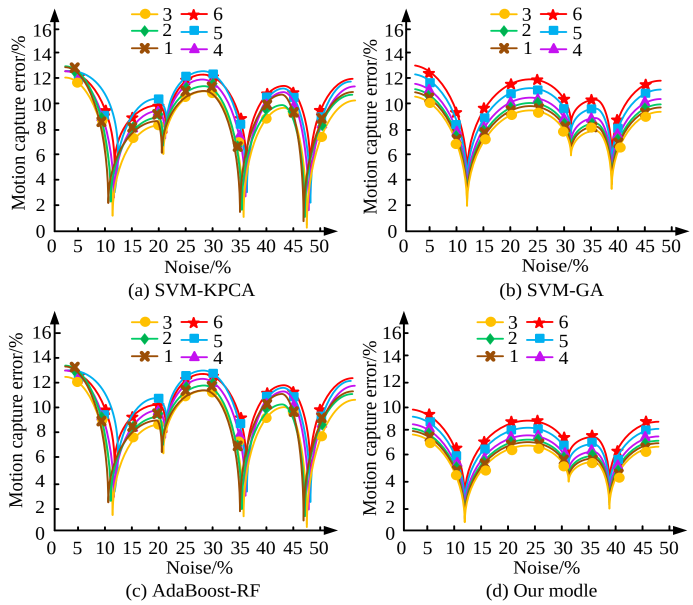


Figure 10: Motion capture error test results under different noise interference.

Table 2: Multi-metric test results for different algorithms.

Data set	Algorithm	P/%	R/%	F1/%	Average accuracy/%	Running time/s
TTSD	SVM-KPCA	88.21	85.83	87.02	87.11	12.34
	SVM-GA	91.27	90.19	90.73	91.28	15.67
	AdaBoost-RF	93.86	92.14	93.03	92.36	10.45
	Our model	95.79	94.58	95.18	96.25	7.66
TTTD	SVM-KPCA	86.44	82.37	84.45	84.77	13.29
	SVM-GA	85.41	87.56	86.48	88.96	16.58
	AdaBoost-RF	91.32	91.28	91.37	91.32	11.34
	Our model	95.07	93.22	94.14	95.13	8.15

Boosting-Random Forest (AdaBoost-RF). Figure 10 shows the test results.

Figure 10 (a) shows the motion capture error under SVM-KPCA. Figure 10 (b) shows the motion capture error under SVM-GA. Figure 10 (c) shows the motion capture error under AdaBoost-RF. Figure 10 (d) shows the motion capture error under the proposed method. As the proportion of noisy data continues to increase, there is some interference in the actual motion capture performance of various methods, especially SVM-KPCA and AdaBoost-RF. These two methods exhibit significant fluctuations in motion capture errors under different noise interference, with the highest capture error being 13% and the lowest being only 2%. These results indicate that these methods' performance is easily affected and their robustness is poor in strong interference environments. Although the SVM-GA's performance has improved and the capture errors have significantly reduced, the overall motion capture error is still relatively large. Relatively speaking, the proposed method has a smaller range of motion capture errors, with a maximum error of only 9.7% and a minimum error of 1.5%. Therefore, the proposed method has demonstrated certain functional superiority and stability among these existing similar methods. The study conducts tests using Precision (P), Recall (R), F1, average accuracy, and running time as indicators for action classification. Table 2 shows the test results.

From Table 2, the proposed model performs well on a number of performance metrics, especially on the TTSD and TTTD. Its F1 values reach 95.18% and 92.02%, respectively, indicating that the model maintains a good balance between precision and recall. Meanwhile, the precision and recall of the proposed model reach 95.79% and 94.58% on TTSD, and 92.69% and 91.35% on TTTD. This further proves that the new model not only accurately classifies most of the samples but also efficiently identifies all the positive class samples. In addition, the running time of the research model in the TTSD is 7.66 seconds, while

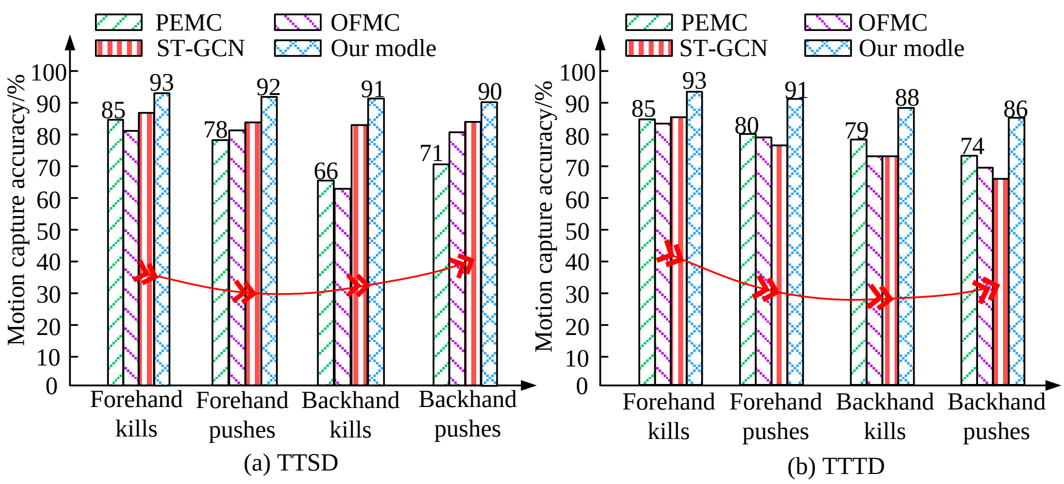


Figure 11: Different motion capture accuracy tests for different models.

the running time of other models such as AdaBoost-RF and SVM-GA is 10.45 seconds and 15.67 seconds, respectively, which is significantly longer. This indicates that the research model has a significant advantage in processing speed. This high efficiency makes the research model particularly suitable for live sports training and real-time motion capture analysis, providing athletes and coaches with fast and accurate feedback to enhance training effects and game strategy development. In contrast, other models perform slightly less well in these metrics, especially in specificity and memory usage. The research model shows higher efficiency and stability. These results suggest that the proposed model has a significant advantage in classification performance, especially when dealing with unbalanced datasets.

3.2 Simulation testing of table tennis motion capture model

Multiple simulation tests are conducted on TTSD and TTTD to validate the performance of the new table tennis motion capture model in real-world scenarios. These tests include capturing and analyzing different table tennis movements, covering common actions such as serving, receiving, pulling, smashing, drawing, are utilized for recording. The average duration, resolution, frame rate, acceleration range, and angular velocity range are optimized according to the actual needs. To reduce the dimensionality of the test data, the study randomly selects four actions: forehand kill, forehand push, backhand kill, and backhand push. More advanced models are introduced for comparison, such as Pose Estimation-based Motion Capture (PEMC), Optical Flow-based Motion Capture (OFMC), and Spatio-Temporal Graph Convolutional Network (ST-GCN). Figure 11 shows the test blocking, chopping, and rubbing. Table 3 shows the data collection and related parameters for each action.

In Table 3, to ensure the accuracy and consistency of the data, the data samples for each action are set to 500. High precision capture devices such as OptiTrack Prime 13, Xsens MVN Awinda, and Phantom VEO 640S results.

Figure 11 (a) shows the capture accuracy of four table tennis movements by different models under TTSD. Figure 11 (b) shows the capture accuracy of four table tennis movements by different models under TTTD. In the forehand and backhand kill movements, the proposed model has the highest capture accuracy, reaching 93% and 91% respectively. PEMC has the lowest capture accuracy among these actions, at 83% and 71% respectively. The new model's capture accuracy for forehand and backhand push movements reaches 78% and 90%, respectively. In contrast, other models have relatively lower capture accuracy in these actions. Especially in the forehand push and backhand kill movements, PEMC's capture accuracy is only 66% and 71%, respectively. Overall, this new model has high capture accuracy in various movements, demonstrating its potential for application in table tennis motion capture, especially in forehand and backhand kills. However, PEMC performs relatively poorly and has lower capture accuracy. Although OFMC performs relatively stable in different movements, its overall capture accuracy is slightly lower than the proposed model's. The study visually verifies different methods' motion capture results through chaos testing in Figure 12.

Figure 12 (a) shows the visualization results of four types of motion capture for PEMC. Figure 12 (b) shows the visualization results of four types of motion capture for OFMC. Figure 12 (c) shows the visualization results of four types of motion capture for ST-GCN. Figure 12 (d) shows the visualization results of four types of motion capture for this proposed model. There is a significant overlap in the distribution of PEMC's capture results, especially in the blurred boundary between forehand and backhand kills, resulting in low accuracy. OFMC's capture results have improved compared to PEMC. However, there is still a certain degree of motion confusion, especially in the low distinction between backhand push and backhand kill. ST-GCN's capture results are significantly better than the first two methods, and the distribution of each action is clearer. However, there is still some overlap between the forehand and backhand push movements. In contrast, the proposed model has a

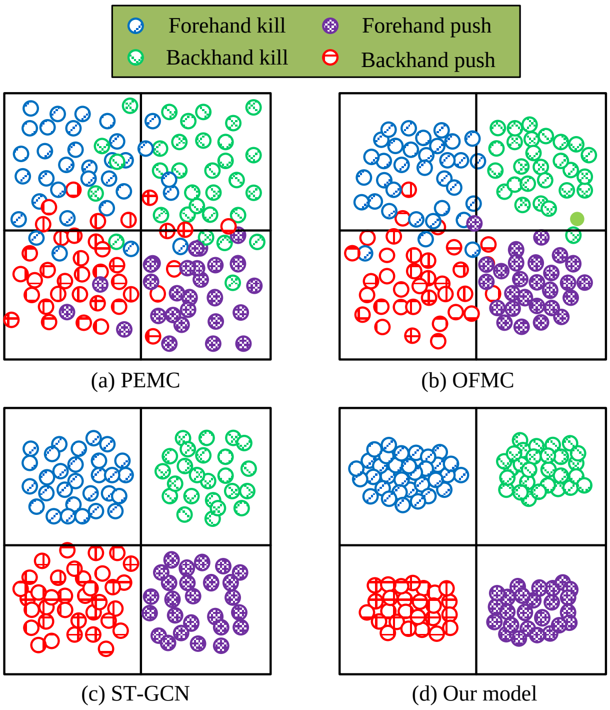


Figure 12: Motion capture visualization results for different methods.

Table 4: Multi-metric test results for different models.

Data set	Algorithm	MAE	MSE	RMSE	Memory usage/MB	Specificity/%
TTSD	PEMC	0.15	0.03	0.17	150	85.32
	OFMC	0.14	0.02	0.14	140	87.45
	ST-GCN	0.13	0.02	0.14	135	89.23
	Our model	0.10	0.01	0.10	120	92.75
TTTD	PEMC	0.17	0.03	0.18	160	82.47
	OFMC	0.16	0.03	0.17	150	85.64
	ST-GCN	0.14	0.02	0.15	145	87.98
	Our model	0.11	0.01	0.11	125	90.34

significant advantage in capturing results with high accuracy. These four actions' distribution is very clear, with almost no overlap. The boundaries between each type of action are distinct. This indicates that the proposed model has high recognition and accuracy in capturing table tennis movements, which can better distinguish different movements. Finally, the study conducts tests using Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), memory usage, and specificity as indicators. Table 4 shows the test results.

In Table 4, the proposed model outperforms other algorithms in all metrics on both TTSD and TTTD. Specifically, this new model has MAE, MSE, and RMSE of 0.10, 0.01, and 0.10 on TTSD, with a memory usage of 120 MB and a specificity of 92.75%. The MAE, MSE, and RMSE on TTTD are 0.11, 0.01, and 0.11, respectively, with a memory usage of 125 MB and a specificity of 90.34%. In contrast, other algorithms have shown certain shortcomings in these indicators, especially in terms of specificity and memory usage, which are significantly different from the proposed model. This indicates that the proposed model not only performs well in accuracy and robustness, but also is more efficient in resource consumption, making it suitable for widespread promotion in practical applications.

4 Discussion

The study proposes a novel table tennis motion capture model by combining SVM with AdaBoost, and experimentally verifies its significant advantages in terms of classification accuracy and computational efficiency. Compared with the model proposed by Li et al., which combines spatial sampling reconstruction encoder with LSTM, the new model achieves a classification accuracy of 96%. Although the two are similar in accuracy, the running time of the new model is significantly reduced to 7.66 seconds, while the latter requires 15 seconds. This result is attributed to the effective combination of SVM and AdaBoost, which improves the robustness and adaptability of the classifier by dynamically adjusting the sample weights, especially when dealing with complex ping-pong actions. In addition, the pose estimation algorithm based on OpenPose and GPU optimization proposed by Wu et al. performs well in video data. However, its 92.7% accuracy and limitations mainly applied to video data limit its generalization ability in real table tennis motion capture. In contrast, in this study, TTSD and TTTD are used for training and testing in the selection of datasets to further validate the applicability and robustness of the model for the characteristics of table tennis movements. The two datasets cover a rich variety of table tennis movement types. Especially in noisy environments, the new model significantly reduces the error range when dealing with high noise and complex backgrounds through the powerful classification ability of SVM and the adaptive sample weight adjustment of AdaBoost, while maintaining high accuracy in motion capture. Its motion capture error range is 1.5% to 9.7%, which is significantly lower than other models. Other studies, such as that conducted by Li W in 2023, which employed a CNN-LSTM combined with a PCA optimization algorithm for lower limb motion prediction, encountered overfitting issues in noisy environments. However, the model demonstrated satisfactory performance on clean datasets [31]. In contrast, the test results of the proposed model on multiple datasets show that it is more robust in dealing with different complex scenarios. In summary, the study develops a model that can effectively capture table tennis movements by combining SVM and AdaBoost algorithms, especially performing well under noise interference. Compared with other methods, the new model has significant advantages in terms of accuracy, efficiency, and robustness, which provides new references and ideas for future sports motion capture technology.

5 Conclusion

The rapid development of artificial intelligence and big data technology has gradually made table tennis training and competition more technological and intelligent. However, the existing motion capture technology for table tennis is affected by factors such as the field, environment, and other non-objective factors, whose accuracy and efficiency need to be improved. To this end, a motion capture model that integrates SVM and AdaBoost is developed. When both γ and Z_e were 1.0, the model performed the best in classification performance, with a maximum classification accuracy of 96%. In TTSD and TTTD, the shortest iteration of the proposed model was 180 times, and the highest classification accuracy was 97%, which remained stable. In contrast, SVM and AdaBoost had lower classification accuracy and slower convergence speed. In addition, the proposed method had a smaller range of motion capture errors, with the highest value being only 9.7% and the lowest value being 1.5%, which was about 3.3% lower compared to SVM alone. After multiple repeated tests, its highest P, R, F1, average accuracy, and shortest running time were 95.79, 94.58%, 95.18%, 96.25%, and 7.66 seconds, respectively. Compared with PEMC, OFMC, and ST-GCN, the proposed model had the highest capture accuracy of 93%, 92%, 91%, and 90% for the four movements of forehand kill, forehand push, backhand kill, and backhand push, respectively. In the visualization results, the proposed model had a very clear capture distribution for the four actions, with almost no overlap. The boundaries between various actions were clear. The minimum MAE, MSE, RMSE, and memory usage were 0.10, 0.01, 0.10, and 120MB, respectively, with the highest specificity of 92.75%. In summary, the model that integrates SVM and AdaBoost has significant advantages in the accuracy and efficiency of motion capture. This method can provide more accurate training data for table tennis players, improving their training effectiveness and competition performance. However, this study only tested a few common table tennis movements. Although existing TTSD and TTTD are sufficient to validate the performance and robustness of the model, future work can be further expanded. Considering expanding to more action types, future research can use larger datasets to test the model's generalization and robustness.

6 References

[1] J. Zhang, Y. Ren, L. Lin, Y. Xing, and J. Ren, "Table tennis motion recognition based on the bat trajectory using varying-length-input convolution neural networks," Scientific Reports, vol. 14, no. 1, pp. 3549-3550, 2024. https://doi.org/10.1038/s41598-024-54150-5

[2] R. Zhu, X. Yang, L. C. Chong, S. Shao, B. István, and Y. Gu, "Biomechanics of topspin forehand loop in table tennis: an application of opensim musculoskeletal modelling," Healthcare, vol. 11, no. 9, pp. 1216-1217, 2023. https://doi.org/10.3390/healthcare11091216

[3] W. Li, "Analyzing the rotation trajectory in table tennis using deep learning," Soft computing, vol. 27, no. 17, pp. 12769-12785, 2023. https://doi.org/10.1007/s00500-023-08962-8

[4] C. H. Wu, T. C. Wu, and W. B. Lin. "Exploration of applying pose estimation techniques in table tennis," Applied Sciences, vol. 13, no. 3, pp. 1896-1897, 2023. https://doi.org/10.3390/app13031896

[5] W. Ren, "A novel approach for automatic detection and identification of inappropriate postures and movements of table tennis players," Soft computing, vol. 28, no. 3, pp. 2245-2269, 2024. https://doi.org/10.1007/s00500-023-09587-7

[6] L. T. Duan, M. Lawo, Z. G. Wang, and H. Y. Wang, "Human lower limb motion capture and recognition based on smartphones," Sensors, vol. 22, no. 14, pp. 5273-5274, 2022. https://doi.org/10.3390/s22145273

[7] F. Feradov, V. Markova, and T. Ganchev, "Automated detection of improper sitting postures in computer users based on motion capture sensors," Computers, vol. 11, no. 7, pp. 116-118, 2022. https://doi.org/10.3390/computers11070116

[8] J. Ding, Y. Wang, H. Si, S. Gao, and J. Xing, "Multimodal fusion-adaboost based activity recognition for smart home on wifi platform," IEEE Sensors Journal, vol. 22, no. 5, pp. 4661-4674, 2022. https://doi.org/10.1109/JSEN.2022.3146137

[9] H. Zhu, X. Li, L. Wang, Z. Chen, Y. Shi, S. Zheng, and M. Li, "IMU motion capture method with adaptive tremor attenuation in teleoperation robot system," Sensors, vol. 22, no. 9, pp. 3353, 2022. https://doi.org/10.3390/s22093353

[10] M. Z. Chen, X. Wang, Q. Chen, Y. Ma, L. M. Lanzoni, and W. K. Lam, "An analysis of whole-body kinematics, muscle strength and activity during cross-step topspin among table tennis players," International Journal of Performance Analysis in Sport, vol. 22, no. 1, pp. 16-28, 2022. https://doi.org/10.1080/24748668.2022.2025712

[11] R. J. Mishra, "A systematic literature review of kinematics involved in table tennis at different achievement levels," NeuroQuantology, vol. 21, no. 6, pp. 1708-1729, 2023. https://doi.org/10.48047/nq.2023.21.6.nq23171

[12] G. W. Intyanto, D. A. F. Yuniarti, and A. S. Pawening, "Implement of the transformation for markerless facial motion capture in 3D animation," Journal of Physics: Conference Series, vol. 1836, no. 1, pp. 12-33, 2021. https://doi.org/10.1088/1742-6596/1836/1/012033

[13] R. Mena-Yedra, J. López Redondo, H. Pérez-Sánchez, and P. Martinez Ortigosa, "ALMERIA: Boosting pairwise molecular contrasts with scalable methods," Informatica, vol. 35, no. 3, pp. 617-648, 2024. https://doi.org/10.15388/24-INFOR558

[14] D. Luvizon, M. Habermann, and V. Golyanik, "Scene-Aware 3D multi-human motion capture from a single camera," Computer Graphics Forum, vol. 42, no. 10, pp. 47-68, 2023. https://doi.org/10.1111/cgf.14768

[15] T. Ning, C. Wang, M. Fu, and X. Duan, "A study on table tennis landing point detection algorithm based on spatial domain information," Scientific Reports, vol. 13, no. 1, pp. 20656-20657, 2023. https://doi.org/10.1038/s41598-023-42966-6
[16] X. Yang, Q. Mei, S. Shao, W. Gu, Y. He, R. Zhu, and Y. Gu, "Understanding sex-based kinematic and kinetic differences of chasse-step in elite table tennis athletes," Bioengineering, vol. 9, no. 6, pp. 246-247, 2022. https://doi.org/10.3390/bioengineering9060246
[17] K. Xing, L. Hang, Z. Lu, C. Mao, D. Kang, C. Yang, and Y. Sun, "Biomechanical comparison between down-the-line and cross-court topspin backhand in competitive table tennis," International Journal of Environmental Research and Public Health, vol. 19, no. 9, pp. 5146-5147, 2022. https://doi.org/10.3390/ijerph19095146
[18] J. Jacquier-Bret and P. Gorce, "Kinematics of the tennis serve using an optoelectronic motion capture system: are there correlations between joint angles and racket velocity?," Sensors, vol. 24, no. 11, pp. 3292-3293, 2024. https://doi.org/10.3390/s24113292
[19] Y. He, G. Fekete, D. Sun, J. S. Baker, S. Shao, and Y. Gu, "Lower limb biomechanics during the topspin forehand in table tennis: a systemic review," Bioengineering, vol. 9, no. 8, pp. 336-338, 2022. https://doi.org/10.3390/bioengineering9080336
[20] H. T. Chiang, B. Y. Tseng, J. L. Chen, and H. C. Hsieh, "Trajectory analysis in UKF: Predicting table tennis ball flight parameters," IT Professional, vol. 26, no. 3, pp. 65-72, 2024. https://doi.org/10.1109/MITP.2024.3380837
[21] S. Tamaki, S. Yamagata, and S. Hashizume, "Spin measurement system for table tennis balls based on asynchronous non-high-speed cameras," International Journal of Computer Science in Sport, vol. 23, no. 1, pp. 37-53, 2024. https://doi.org/10.2478/ijcss-2024-0003
[22] S. Qiu, H. Zhao, N. Jiang, D. Wu, G. Song, and H. Zhao, "Sensor network oriented human motion capture via wearable intelligent system," International Journal of Intelligent Systems, vol. 37, no. 2, pp. 1646-1673, 2022. https://doi.org/10.1002/int.22689
[23] J. Xu, "Application of optical motion capture device based on android intelligent platform in sports field auxiliary recognition system," Optical and Quantum Electronics, vol. 56, no. 3, pp. 296-297, 2024. https://doi.org/10.1007/s11082-023-05878-3
[24] W. W. T. Lam, Y. M. Tang, and K. N. K. Fong, "A systematic review of the applications of markerless motion capture (MMC) technology for clinical measurement in rehabilitation," Journal of Neuroengineering and Rehabilitation, vol. 20, no. 1, pp. 57-59, 2023. https://doi.org/10.1186/s12984-023-01186-9
[25] Y. Q. Zhu, Y. M. Cai, and F. Zhang, "Motion capture data denoising based on LSTNet autoencoder," Journal of Internet Technology, vol. 23, no. 1, pp. 11-20, 2022. https://doi.org/10.53106/160792642022012301002
[26] Z. Hu, J. Tang, L. Li, J. Hou, H. Xin, X. Yu, and J. Bu, "MarkerNet: A divide-and-conquer solution to motion capture solving from raw markers," Computer Animation and Virtual Worlds, vol. 35, no. 1, pp. 2228-2229, 2024. https://doi.org/10.1002/cav.2228
[27] Y. Ren, C. Zhao, and Y. He, "Lidar-aid inertial poser: Large-scale human motion capture by sparse inertial and lidar sensors," IEEE Transactions on Visualization and Computer Graphics, vol. 29, no. 5, pp. 2337-2347, 2023. https://doi.org/10.1109/TVCG.2023.3247088
[28] E. Warmerdam, C. Hansen, R. Romijnders, M. A. Hobert, J. Welzel, and W. Maetzler, "Full-body mobility data to validate inertial measurement unit algorithms in healthy and neurological cohorts," Data, vol. 7, no. 10, pp. 136-137, 2022. https://doi.org/10.3390/data7100136
[29] M. Zhu, X. Guan, Z. Li, L. He, Z. Wang, and K. Cai, "sEMG-based lower limb motion prediction using CNN-LSTM with improved PCA optimization algorithm," Journal of Bionic Engineering, vol. 20, no. 2, pp. 612-627, 2023. https://doi.org/10.1007/s42235-022-00280-3
[30] M. Hasanvand, M. Nooshyar, E. Moharamkhani, and A. Selyari, "Machine learning methodology for identifying vehicles using image processing," AIA, vol. 1, no. 3, pp. 170-178, 2023. https://doi.org/10.47852/bonviewAIA3202833
[31] W. Li, "Analyzing the rotation trajectory in table tennis using deep learning," Soft Computing, vol. 27, no. 17, pp. 12769-12785, 2023. https://doi.org/10.1007/s00500-023-08962-8