Hindawi
Wireless Communications and Mobile Computing
Volume 2023, Article ID 9805789, 1 page
https://doi.org/10.1155/2023/9805789

Retraction

Retracted: Application of Machine Learning to Badminton Action Decomposition Teaching

Wireless Communications and Mobile Computing

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] J. Long and S. Rong, "Application of Machine Learning to Badminton Action Decomposition Teaching," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 3707407, 10 pages, 2022.

Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 3707407, 10 pages
https://doi.org/10.1155/2022/3707407

Research Article

Application of Machine Learning to Badminton Action Decomposition Teaching

Jiahuai Long and Shuguang Rong

Department of Physical Education, Bozhou University, Bozhou, 236800 Anhui, China

Correspondence should be addressed to Shuguang Rong; 2011050045@bzuu.edu.cn

Received 30 January 2022; Revised 18 February 2022; Accepted 19 February 2022; Published 20 April 2022

Academic Editor: Mohammad Farukh Hashmi

The study was aimed at realizing the identification of athletes' actions in badminton teaching. The teaching process is segmented into many independent actions to help learners standardize their movements in badminton play, improving the national physical quality. First, the principle and advantages of machine vision sensing are introduced. Second, the images and videos about the action decomposition of badminton teaching are collected, and the image data are extracted by Haar-like. Subsequently, badminton players' actions are recognized and preprocessed, and a dataset is constructed. Furthermore, a new algorithm model is implemented and trained by using Haar-like and Adaptive Boosting (AdaBoost). Finally, the badminton players' action recognition algorithm is tested and compared with the traditional hidden Markov model (HMM) and support vector machine (SVM). The results show that action images improved by machine vision can process the captured actions effectively, making the computer better identify different badminton teaching actions. The proposed method has a recognition rate of more than 90% for each action, the average recognition accuracy of actions reaches 95%, the average recognition rate of the same person's actions is 96.5%, and the average recognition rate of different people's actions is 94.8%. The badminton teaching action recognition model based on Haar-like and AdaBoost can recognize and classify badminton actions and improve the quality of badminton teaching. This study shows that the image processing technology can effectively process the players' static images, which gives the direction for physical education (PE) under artificial intelligence (AI).

1. Introduction

Due to the development of Internet technology, artificial intelligence (AI) [1] technology is widely used in all aspects of peoples' lives, such as production, education, and research fields. Nowadays, it is also used for national physical fitness and professional physical training.

Vision sensors can provide information to the machine vision sensing system, and computer vision conduction focuses on the simulation of animal vision by using computer technology. The simulation involves a variety of technologies, including image processing technology, mechanical engineering technology, mechanical and electronic control technology, lighting technology, optical image capture, processing technology and sensor technology, analogy and digital video technology, computer software and hardware, and support technology (for image enhancement and algorithm analysis). Its main task is computer processing and recognition of images [2, 3]. The application system of machine vision includes the image acquisition module, light-source processing module, image digitization module, digital image processing module, decision-making module, and hardware control implementation module. Machine vision is realized by improving intelligent automation and the quality of production and service. It is used in the environments where manual operation cannot meet the requirements. It can monitor the product's quality and greatly improves the production efficiency in large-scale mechanical production [4–6].

Zhou et al. compared the efficiency of support vector machine (SVM), logistic regression (LR), and artificial neural network (ANN) for shoulder motion pattern recognition of surface electromyogram (EMG). They studied the effect of sliding time window epochs on the recognition accuracy and

verified the accuracy of the LR recognition algorithm [7]. Storey et al. proposed an end-to-end framework named 3D-PalsyNet. The framework is used for mouth motion recognition and facial paralysis scoring. 3D-PalsyNet utilizes a 3D-convolutional neural network (CNN) architecture with a ResNet backbone to predict these dynamic tasks. A pretrained 3D-CNN on the kinetic dataset is used for transfer learning for general action recognition. The model is modified to combine center and softmax for supervised learning [8]. Zhang et al. proposed a solution to identify table tennis motion using commercial smart watches and developed a data acquisition system based on IoT architecture. The system is used to obtain data on the acceleration, angular velocity, magnetic induction, etc. of the watch. Based on the features of the extracted data, the main machine learning (ML) classification algorithms such as k-nearest neighbors, SVM, naive Bayes model (NBM), LR, decision tree, and random forest (RF) are used for verification experiments [9]. Most of the previous studies have used ML techniques such as supervised learning. The innovation lies in the model constructed by using Haar-like features combined with Adaptive Boosting (AdaBoost). The AdaBoost algorithm makes good use of weak classifiers for cascading and can use different classification algorithms as weak classifiers with high accuracy. Relative to the RF algorithm, AdaBoost fully considers the weight of each classifier.

In the information age, there appear various emerging computer technologies, and intelligent image recognition technology applied to badminton teaching is included. In the use of intelligent technology, the teaching action images are collected by cameras, and Haar-like is used for feature extraction and image denoising. The action images are preprocessed and the training set and test set are constructed. The model is implemented by AdaBoost, and the performance of the constructed recognition algorithm is tested. According to computer vision technology, the collected images are preprocessed and the implemented model is trained, achieving intelligent image processing and recognition timely, efficiently, and accurately.

2. Methods

2.1. Machine Vision Sensing. Vision is originally a way for organisms to obtain external information. Now, it is one of the core components in promoting biological intelligence. It is known that 80% of information is obtained by vision. Figure 1 shows some ways to obtain information through the vision. Inspired by this, researchers install "eyes" on machines, so that machines can get necessary external information through "seeing" like humans, which is called computer vision. Machine vision is a comprehensive technology, and it is mainly applied to establish an image capture system, a light source system, an image digitization module, and so on. After that, researchers make a machine vision system by analyzing biological vision systems. The key to vision sensor technology is image processing; that is, an image is processed by intercepting the signal on the object surface [10, 11].

The vision sensor has a huge number of pixels, which can capture the light of an image. And the number of pixels determines the clarity and fineness of an image. After an image is captured, the vision sensor compares it with the given image for analysis. Visual sensing technology includes 3D visual sensors, which is applied in many fields, such as multimedia mobile phones, network cameras, digital cameras, robot visual navigation, automobile safety systems, biomedical pixel analysis, man-machine interfaces, virtual reality (VR), monitoring, industrial detection, wireless remote sensing, microscope technology, astronomical observation, marine autonomous navigation, and scientific instrument tests. In particular, it can be used in industrial control and automobile autonomous navigation. A complete machine vision system contains the following five parts: lighting systems, lens, high-speed cameras, image acquisition cards, and vision processors [12, 13]. The requirements of each part are shown in Table 1.

Intelligent visual sensing technology is also a visual sensing technology. It is also an intelligent camera with a small machine vision system, which can collect and process images and transmit the relevant information. It assembles the image sensors, digital processors, communication modules, and other peripherals into a camera, simplifying the system and improving its reliability. This widens the application range of vision technology. It can help to build a reliable detection system because it is easy to learn, use, and install. Its image collection unit comprises a charge-coupled device (CCD)/a complementary metal-oxide-semiconductor (CMOS), an optical system, a lighting system, and an image acquisition card. The optical image is converted into the digital image and transmitted to the image processing unit, as shown in Figure 2.

The machine vision system can improve flexibility and automation in the production process. In some dangerous environments, it can replace artificial vision. In mass repetitive production, it can detect and improve working efficiency.

CCD has the functions of photoelectric conversion, information storage, delay, and sequential transmission of electrical signals and has high integration and low power consumption, so it has developed rapidly. CCD is an indispensable key device for image acquisition and digital processing. It is widely used in scientific, educational, medical, commercial, industrial, military, and consumer fields. CCD system mainly includes optical system (microlens), CCD, and image processing module, and some also include color filter. A CCD image sensor is an array of capacitors arranged according to certain rules. A very thin layer of silicon dioxide (SiO_2 about 120 nm) is formed on the silicon substrate, and then metal or doped polysilicon electrodes (gates) are sequentially deposited on the SiO_2 thin layer to form a regular capacitor array, thus forming a CCD chip. The working process of CCD is as follows: (a) the generation of signal charge. The CCD can convert the incident light signal into a charge output. The principle is the photoelectric effect (photovoltaic effect) in the semiconductor. The metal-oxide-semiconductor capacitor is the most basic unit that constitutes a CCD. (b) Storage of signal charge: it is the

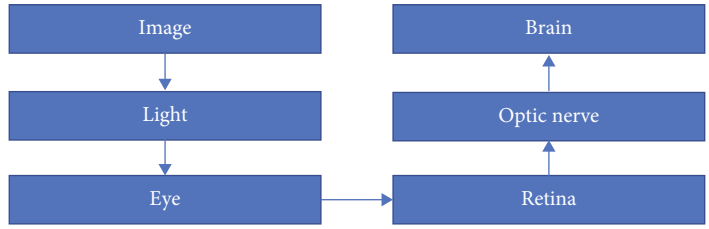

Figure 1: Visual means to obtain information.

Table 1: Requirements for each part of the machine vision system.

Names	Requirements
Lighting systems	Light sources: incandescent lamps, fluorescent lamps, mercury lamps, and sodium lamps Irradiation methods: backlighting, forward lighting, structured lighting, and frequency flash lighting
Camera lens	Field of vision (FOV) = target resolution ∗ subpixel ∗ camera size/part measurement tolerance ratio
High-speed cameras	Lens selection conditions: (a) focal length, (b) target height, (c) projected image height, (d) magnification, (e) distance from the projected image to target, (f) graphic center point/node, and (g) image distortion
Image acquisition cards	The interface requirements of the camera: black and white, color, analog, or digital
Visual processors	The high configuration is the programmable logic controller (PLC)

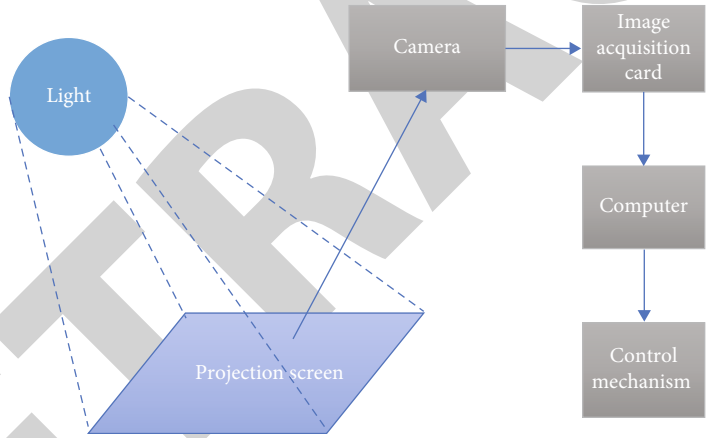

Figure 2: Principle of vision sensors.

process of collecting the charges excited by incident photons and converting them into signal charge packets. (c) Transmission and transfer of signal charges: it is the process of transferring the collected charge packets from one pixel to the next until all the charge packets are output. (d) Detection and output of signal charge: it is the process of converting the charge transferred to the output stage into current or voltage. There are three main output types: current output, floating gate amplifier output, and floating diffusion amplifier output.

2.2. ML

2.2.1. Deep Learning. In 2006, deep learning (DL) appears and it is a branch of ML. It is studied by academia and gradually applied in the industry. In 2012, the Stanford University uses 16000 CPU core parallel computing platforms to expound deep neural networks (DNN), which have an advantage in speech and image recognition. In 2016, "go" is developed from DL and it helps defeat Li Shishi, the world's top master, in the competition. After that, well-known high-tech companies around the world begin to pay more attention to DL and establish research institutes for it, expanding the size of the research team of DL [14].

ML studies how computers simulate or realize the learning behavior of animals and use new knowledge or skills to rewrite the existing data structure, improving the program performance. According to statistics, it can estimate data distribution, learn a data model, and predict new data through this model. ML uses algorithms to analyze data, learn from it, and make decisions. This means that it can teach computers to develop an algorithm and complete assignments instead of compiling programs to perform certain tasks. There are three main types of ML: supervised learning, unsupervised learning, and reinforcement learning. They have specific advantages and disadvantages. Supervised

learning targets labeled data. In the learning process, the computer identifies new samples by the specific patterns, and classification and regression are achieved. In terms of classification, the machine is trained and data are divided into specific classes. This process is like the spam filter on your email account. The filter analyzes the previous email messages and compares them with new ones. If the given ratio is met, the messages are marked and sent to the corresponding folder, and the rest are sent to the destination mailbox. As for regression, the machine can predict future ones according to the labeled data. This is the same as a weather broadcast. With historical data (average temperature, humidity, and precipitation), the APP on mobile phones can predict the weather conditions in the future. Unsupervised learning is for unlabeled data, and it can perform clustering and dimension reduction. Clustering is completed according to the attributes and behaviors. It divides a group into different subgroups (according to age and marital status) and applies them to marketing schemes. Dimensions are reduced under common ground.

Reinforcement learning is implemented on the personal experience of the machine. It is like playing games. It focuses on performance. For example, you play chess on a computer and cannot move the king into the space that the opponent's chess pieces can enter. The experience of playing chess is inferred until the machine beats (and eventually defeat) the top players.

It makes the test data and training data identically distributed. It tries to imitate the transmitting and processing mode of brain neurons. It is mostly applied to computer vision and natural language processing (NLP). DL relies on neural networks in ML. Therefore, DL is called the improved neural network [15, 16]. Figure 3 shows the structure of the convolutional neural network (CNN).

2.2.2. Artificial Neurons. The artificial neuron is a mathematical model created by imitating the basic operation function of biological neurons. The artificial neuron receives the given signal from the front neuron, and each given signal is attached with a weight. Under the joint action of all weights, this neuron shows a corresponding activation state [17], and its principle is shown in Figure 4.

The principle of artificial neurons is expressed by

$$f(x) = \sum_{i=1}^{n} x_i w_i. \tag{1}$$

In equation (1), $f(x)$ represents the final output state, x_i is the input signal, and w_i is its weight. There are n groups in total.

When it receives an input signal, a neuron gives a certain output, and each neuron has a corresponding threshold. If the sum of the inputs received by this neuron is greater than the threshold, its state will change to an active state. When it is less than the threshold, it will show an inhibitory state. The transfer functions of artificial neurons are as follows [18, 19].

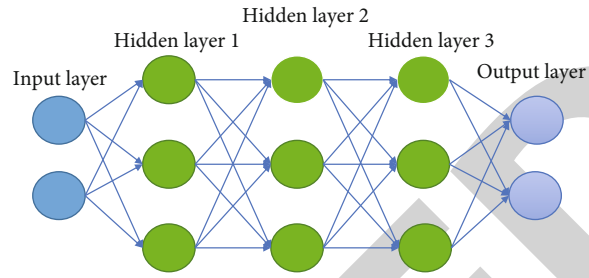

Figure 3: Structure of CNN

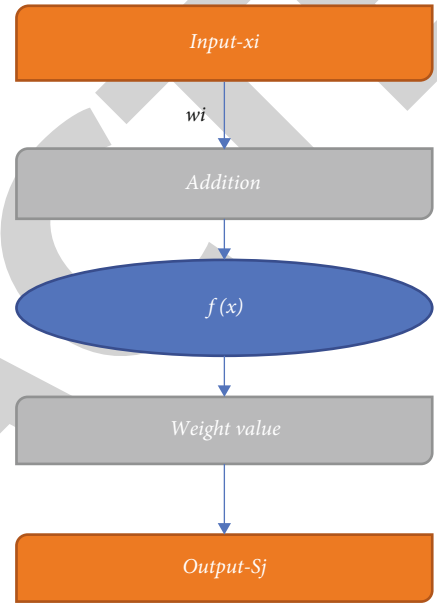

Figure 4: Principle of artificial neurons.

(a) The linear function is calculated by

$$f(x) = kx \tag{2}$$

(b) The slope function is calculated by

$$\begin{aligned} f(x) &= \alpha(x \geq \theta), \\ f(x) &= kx(-\theta < x < \theta), \\ f(x) &= -\alpha(x \leq \theta) \end{aligned} \tag{3}$$

The transition function is calculated by

$$\begin{aligned} f(x) &= \alpha(x \geq \theta), \\ f(x) &= \beta(x \leq \theta). \end{aligned} \tag{4}$$

Sigmoid is calculated by

$$f(x) = a + \frac{b}{1 + \exp(-dx)}. \tag{5}$$

The transfer function needs to be selected according to

the specific range. The linear function amplifies the output signal. The nonlinear slope function prevents the degradation of network performance. The S-type function sets parameters α, β, d, a, b, and θ in the hidden layer. In equation (5), x and $f(x)$ represent the input and output values, respectively.

Neurons can be used for calculating weights and summations. x_i is the input value of the i-th neuron, w_{ji} is the weight between the i-th and j-th neuron, b_j is a threshold, and $f(x)$ is the transfer function.

Net output value S_j of neuron j is calculated in equation (6). When the threshold is 0, b_j is 0.

$$S_j = \sum_{i=1}^{n} w_{ji} \cdot x_i + b_j. \tag{6}$$

2.3. Image Recognition and Processing

2.3.1. Image Recognition. Image recognition is shown in Figure 5.

2.3.2. Feature Extraction. Nowadays, the features of image recognition fall into global and local features. The global features are extracted by the full graph search window, which reflects the general feature. The local feature is the detailed features of the image in the sliding window. Global feature extraction method can make principal component analysis (PCA) [20], gray-level gradient cooccurrence matrix (GLCM) [21], and frequency domain. Local features include scale invariant feature transform (SIFT) [22] and Haar-like features [23].

Haar-like features are extracted for recognition. They have four basic structures and are viewed as windows. This window slides in the image with a step of 1, and a complete image is finally extracted. The length and width of the sliding window are increased. The process is repeated until they are enlarged to the largest.

2.3.3. Classification and Recognition Method. Image recognition needs to use the classifier and a classifier should be designed. And feature extraction is realized by supervised and unsupervised learning. Supervised methods need a huge number of training sample sets and a classifier, including neural networks, SVM [24], and AdaBoost [25]. Unsupervised methods classify data according to their similarity and characteristics. The mainstream method is K-mean [26]. Canadian EC650C camera and 5G are used for data wireless network transmission by AdaBoost.

2.4. Recognition Algorithms. The mainstream recognition algorithm is local vision recognition, but there is less research on image recognition using omnidirectional vision. Therefore, the action recognition model based on DL is used to improve local vision recognition by Haar-like and AdaBoost.

First, the video data are collected and the sample image is preprocessed. Then, the collected images are used as the training set, and the classifier is constructed. After that, the algorithm is used to obtain the frames of images. If tracking fails, badminton players' actions will be retracted.

2.4.1. Sample Set Construction. A sample set should be constructed for feature training and algorithm recognition. It consists of a training set and a test set. Because there will be a lot of training, a large number of images are needed. The training set should include geographical features and light conditions. Finally, 263 players' images and 557 scene images are collected.

2.4.2. Image Preprocessing. The scale and size of the sample set should be adjusted for Haar-like features' calculation. Figure 6 shows the data pre-processing process as follows.

2.4.3. Classifier Algorithm. The AdaBoost algorithm is used for establishing the algorithm of the classifier [27, 28], and the equation is as follows:

(a) Training set (x_i, y_i) is constructed

(b) Initial parameters and their weights are determined by

$$w_i = 1/N,\ i = 1, 2, \cdots, N,\ F(x) = 0 \tag{7}$$

(c) Weighted mean square deviation is estimated by

$$f_m(x) = \arg\min E_w \left[\frac{(y - f(x))^2}{x} \right] \tag{8}$$

(d) The classifier is updated by

$$F(x) = F(x) + f_m(x) \tag{9}$$

(e) The sample weight is updated and normalized by

$$w_i = w_i \exp\left[-y_i f_m(x)\right] \tag{10}$$

(f) Iterate 3-5 steps, and the times are $m = 1, 2, \cdots, M$

(g) The result of the classifier is obtained by

$$\text{Sign}[F(x)] = \text{Sign}\left[\sum_{m=1}^{M} f_m(x) \right] \tag{11}$$

2.4.4. Data Enhancement. D is the data coordinate information of the image, and L is the original labeled sample of the actual value of the human action, denoted as (D, L). The sample after adjustment in the expression frame is denoted as $(D, \bar{L})$.

Data enhancement requires a variety of operations on the original image. In order to describe the rotated coordinates, set the top left corner of the image as point O, and set its coordinates as (0,0). Starting from (0,0), the coordinates from top to bottom are set as the y-axis, and the

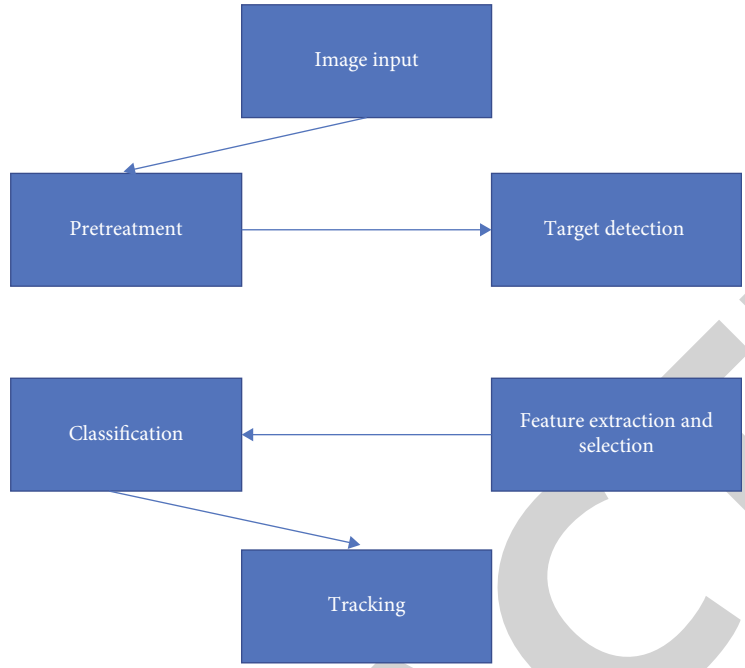

Figure 5: Image recognition

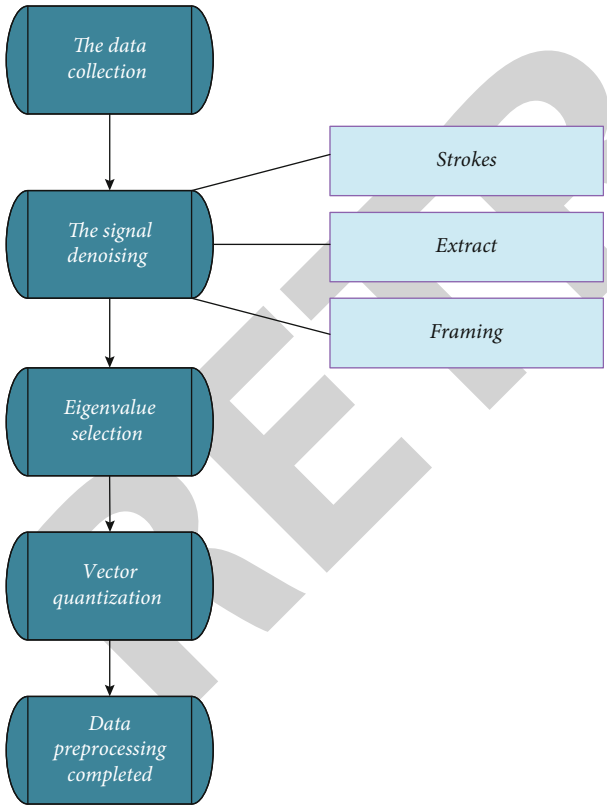

Figure 6: Data preprocessing.

coordinates from left to right are set as the x-axis. The picture size is (D_x, D_y), the rotation angle is φ, and the clockwise direction is the positive direction. $l_i^r = (x_i^r, y_i^r)$ is the coordinate of the rotated human body joint i, and D^r is the rotated image data.

(a) Rotation is shown in

$$\begin{aligned} x_i^r &= \left(x - \frac{D_x}{2}\right)\cos\varphi - \left(y - \frac{D_y}{2}\right)\sin\varphi + \frac{D_x^r}{2}, \\ y_i^r &= \left(y - \frac{Dy}{2}\right)\cos\varphi - \left(x - \frac{Dx}{2}\right)\sin\varphi + \frac{D_y^r}{2}, \end{aligned} \tag{12}$$

$$\begin{aligned} D_x^r &= \left(D_y - D_x\tan\varphi\right)\sin\varphi + D_x\cos\varphi, \\ D_y^r &= \left(D_x - D_y\tan\varphi\right)\sin\varphi + D_y\cos\varphi \end{aligned} \tag{13}$$

The rotation angle is set to be random to obtain many training samples.

(b) Translation: different points are selected on the image to translate the human body. According to equation (14), the minimum clipping region is calculated.

$$\begin{aligned} C_{\text{top-left}} &= \left(\min\left(x_1^r,\cdots,x_K^r\right), \min\left(y_1^r - y_K^r\right)\right), \\ C_{\text{bottom-right}} &= \left(\left(\max\left(x_1^r,\cdots,x_K^r\right), \max\left(y_1^r - y_K^r\right)\right)\right) \end{aligned} \tag{14}$$

$C_{\text{top-left}}$ is the coordinate of the pixel at the top left of the cropped area. $C_{\text{bottom-right}}$ is the coordinate of the bottom right pixel.

(c) Zoom

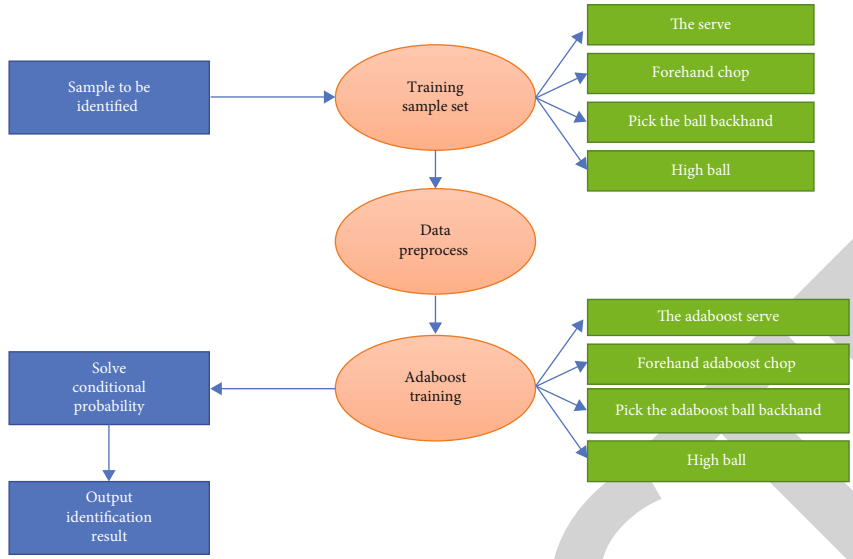

Figure 7: Training and recognition process.

Table 2: Simulation experiment.

Experiment names	Experiment codes	Experimental classification	Description
Comparison of three different segmentation methods	A	Sliding window segmentation in the hitting time	50 sample points before and after hitting the ball as action characteristics
	B	Sliding window segmentation based on the maximum	Sliding windows with a width of 100, the maximum, and the reset
	C	Window segmentation based on events	The starting and ending time of the player's action as those of the experiment
Training and recognition models of different algorithms	D	SVM	Identifying ten hitting actions
	E	Hidden Markov model (HMM) [29]	Comparison of the posterior probability of the model after training
	F	Combination of Haar-like, AdaBoost, and HMM	
The system recognition rate of images	G	One player	Training and test data of one player
	H	Different players	Training and test data of many different players

(d) Horizontal flip: equation (15) is the coordinates after horizontal flipping.

$$x_j^f = D_x^S - x_i^S, \quad y_j^f = y_i^S \tag{15}$$

x_j^f represents the abscissa after flipping. y_j^f represents the ordinate after flipping. D_x^S represents image data after flipping. (x_j^f, y_j^f) represents the pixel coordinate value after flipping. j is the result after the original i joint is flipped.

Figure 7 shows the training and recognition process.

2.5. Testing Environments. The recognition rate of Haar-like is tested. 300 teaching and 500 nonplayer images are extracted, including 600 training images and 200 test images. They are used to identify the athletes' serve, positive rubbing, reverse rubbing, positive flutter, reverse flutter, positive push, reverse push, positive pick, reverse pick, and high and far actions. Among them, each action has a training set of 40 pictures and a test set of 10 pictures. The contents of the three experiments are shown in Table 2.

3. Result

3.1. Influence of Different Sliding Windows on Segmentation. The comparison of recognition rates of three different segmentation methods is shown in Figure 8.

Figure 8 shows that method A based on the hitting time window has a higher recognition rate than B and C. Its recognition rate in each event is greater than 90%, which is much higher than the effect of the other two. Therefore,

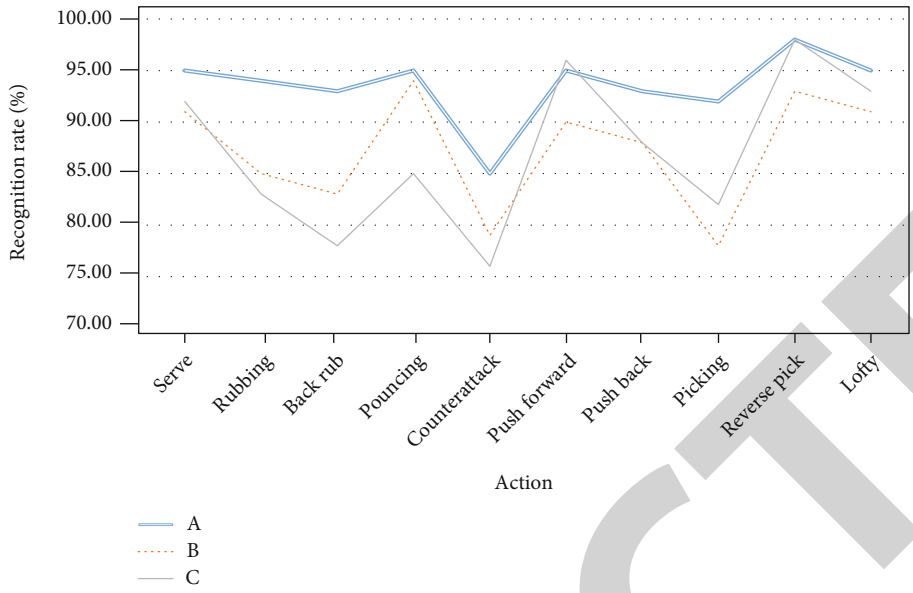

Figure 8: Recognition rates of different segmentation methods.

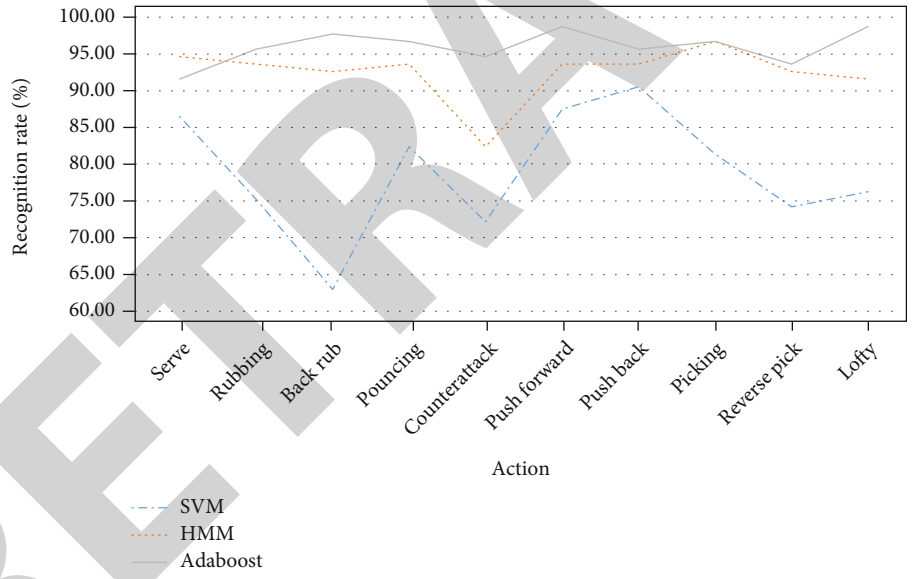

Figure 9: Recognition rates of different algorithms.

the method based on the hitting time window is used to extract the players' hitting actions. For each hitting point, the peak value of hitting time is detected, and it is superior.

3.2. Recognition Rates of Different Algorithms. The comparison of recognition rates of three different training and recognition models is shown in Figure 9.

Figure 9 shows that the recognition rate of AdaBoost for each action is greater than 90%, and the average recognition accuracy is greater than 95% compared with SVM and HMM. The average action recognition rate of SVM is only 78%, and that of HMM is between 90% and 95%. This shows that the effect of AdaBoost is much higher than that of traditional algorithms.

3.3. The Image Recognition Rate of the Same and Different Players. The comparison of the recognition rates of the same and different players is shown in Figure 10.

Figure 10 shows that AdaBoost's recognition rate of each action of the players is greater than 90%, and it has a more accurate recognition rate for the actions of one player. Its recognition rate for the same player is more than 99%, and its lowest rate is 92%. For the action recognition rate of different players, its recognition rate is also up to 99%, and the

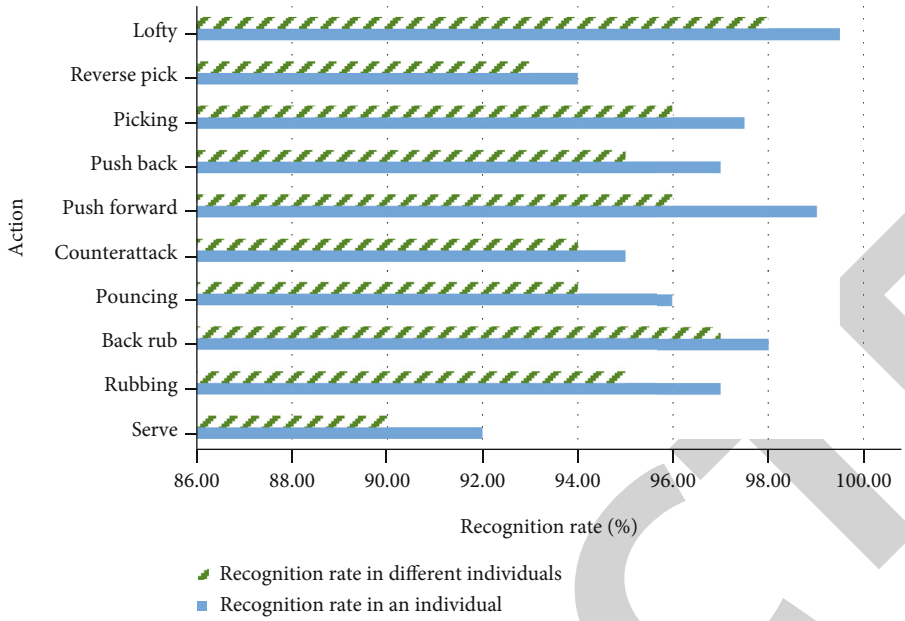

Figure 10: Recognition rates of players.

lowest is 90%. Therefore, in badminton teaching action decomposition, it is best to use the action image of the same player.

4. Conclusion

Based on the intelligent image recognition of football robots, a vision sensing system suitable for badminton teaching action decomposition based on DL and machine vision is proposed. The image data are collected, and Haar-like is used for feature extraction. The data of badminton players' actions is preprocessed, and the dataset is constructed. A new model is implemented and trained by Haar-like and AdaBoost, and the performance of the constructed recognition algorithm is tested and analyzed. The experiment shows that the badminton teaching action recognition technology based on Haar-like and AdaBoost can successfully capture and classify images intelligently, improving the quality of badminton teaching.

The study achieves the expected research results and draw valuable conclusions, but there are still deficiencies: (1) the algorithm may have some difficulty in recognizing players' actions in professional competitions, and more action data in formal competitions will be collected in the future; (2) the action recognition rate of different players is not ideal, and more action data of different players will be collected for training afterward.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors report no declarations of interest.

Acknowledgments

This work was supported by the Key Project of Quality Engineering Project Teaching and Research in Anhui Province, Research on the "Trinity" Health Literacy Cultivation Model of College Students—A Case Study of Bozhou University (Project No. 2019jyxm0538).

References

[1] A. Wodecki and A. Wodecki, "Artificial intelligence methods and techniques," *Journal of Fundamental and Applied Sciences*, vol. 8, no. 2, pp. 71–132, 2019.

[2] S. Ghosal, D. Blystone, A. K. Singh, B. Ganapathysubramanian, A. Singh, and S. Sarkar, "An explainable deep machine vision framework for plant stress phenotyping," *Proceedings of the National Academy of Sciences*, vol. 115, no. 18, pp. 4613–4618, 2018.

[3] J. Radcliffe, J. Cox, and D. M. Bulanon, "Machine vision for orchard navigation," *Computers in Industry*, vol. 98, no. 1, pp. 165–171, 2018.

[4] R. A. Hamzah, A. F. Kadmin, S. Ghani, M. S. Hamid, and S. Salam, "Disparity refinement process based on RANSAC plane fitting for machine vision applications," *Journal of Fundamental and Applied Sciences*, vol. 9, no. 4S, p. 226, 2018.

[5] H. Kim, "CMOS image sensor for wide dynamic range feature extraction in machine vision," *Electronics Letters*, vol. 57, no. 5, pp. 206–208, 2021.

[6] M. Fahimipirehgalin, E. Trunzer, M. Odenweller, and B. Vogel-Heuser, "Automatic visual leakage detection and localization from pipelines in chemical process plants using

machine vision techniques," *Engineering*, vol. 7, no. 6, pp. 758–776, 2021.

[7] Y. Zhou, C. Chen, M. Cheng et al., "Comparison of machine learning methods in SEMG signal processing for shoulder motion recognition," *Biomedical Signal Processing and Control*, vol. 68, no. 1, article 102577, 2021.

[8] G. Storey, R. Jiang, S. Keogh, A. Bouridane, and C. T. Li, "3DPalsyNet: a facial palsy grading and motion recognition framework using fully 3D convolutional neural networks," *IEEE Access*, vol. 7, no. 2, pp. 121655–121664, 2019.

[9] H. Zhang, Z. Fu, and K. Shu, "Recognizing ping-pong motions using inertial data based on machine learning classification algorithms," *IEEE Access*, vol. 7, no. 6, pp. 167055–167064, 2019.

[10] A. Koirala, K. B. Walsh, and Z. Wang, "Attempting to estimate the unseen—correction for occluded fruit in tree fruit load estimation by machine vision with deep learning," *Agronomy*, vol. 11, no. 2, article 347, 2021.

[11] B. Nair, S. Krishnamoorthy, and S. Rao, "Machine vision based flood monitoring system using deep learning techniques and fuzzy logic on crowdsourced image data," *Intelligent Decision Technologies*, vol. 1, no. 9, pp. 1–14, 2021.

[12] S. Harnsoongnoen and N. Jaroensuk, "The grades and freshness assessment of eggs based on density detection using machine vision and weighing sensor," *Scientific Reports*, vol. 11, no. 1, pp. 46–49, 2021.

[13] Q. Liao, C. Wei, Y. Li, L.'. Guo, and H. Ouyang, "Developing a machine vision system equipped with UV light to predict fish freshness based on fish-surface color," *Food and Nutrition Sciences*, vol. 12, no. 3, pp. 239–248, 2021.

[14] P. Bossuyt, S. Vermeire, and R. Bisschops, "Assessing disease activity in ulcerative colitis using artificial intelligence: can equally good be seen as better," *Gastroenterology*, vol. 159, no. 4, pp. 1625-1626, 2020.

[15] R. Hirayama, T. Iwata, S. Yamada et al., "COT-16 development of automatic lesion extraction application using artificial intelligence for the purpose of simplifying tumor volume measurement of meningioma," *Neuro-Oncology Advances*, vol. 6, no. 6, pp. 46–49, 2021.

[16] N. Gillet, A. Mesinger, B. Greig, A. Liu, and G. Ucci, "Deep learning from 21-cm tomography of the cosmic dawn and reionization," *Monthly Notices of the Royal Astronomical Society*, vol. 2, pp. 1–5, 2019.

[17] D. S. Kermany, M. Goldbaum, W. Cai et al., "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.

[18] M. M. U. Rathore, A. Paul, A. Ahmad, B. W. Chen, B. Huang, and W. Ji, "Real-time big data analytical architecture for remote sensing application," *Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 10, pp. 4610–4621, 2017.

[19] H. Roth, L. Lu, N. Lay et al., "Spatial aggregation of holistically-nested convolutional neural networks for automated pancreas localization and segmentation," *Medical Image Analysis*, vol. 45, no. 3, pp. 94–107, 2018.

[20] W. Ni, A. Stoffelen, and K. Ren, "Hurricane eye morphology extraction from SAR images by texture analysis," *Frontiers of Earth Science*, vol. 11, no. 8, pp. 1–16, 2022.

[21] K. Y. Yeung and W. L. Ruzzo, "Principal component analysis for clustering gene expression data," *Bioinformatics*, vol. 17, no. 9, pp. 763–774, 2001.

[22] Y. Miftahuddin, N. F. Fahrudin, and M. F. Prayoga, "Algoritma Scale Invariant Feature Transform (SIFT) pada Deteksi Kendaraan Bermotor di Jalan Raya," *Mind Journal*, vol. 5, no. 1, pp. 54–65, 2021.

[23] S. Pal, "Human face detection technique using Haar-like features," *International Journal of Computer Applications*, vol. 175, no. 32, pp. 56–60, 2020.

[24] H. C. Husada and A. S. Paramita, "Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma Support Vector Machine," *Teknika*, vol. 10, no. 1, pp. 18–26, 2021.

[25] B. Naik, "Extra-tree learning based socio-economic factor analysis and multi-class adaptive boosting meta-estimator for prediction of agricultural productivity," *Indian Journal of Science and Technology*, vol. 13, no. 29, pp. 2081–2101, 2020.

[26] G. N. Pham, S. H. Lee, O. H. Kwon, and K. R. Kwon, "A watermarking method for 3D printing based on menger curvature and K-mean clustering," *Symmetry*, vol. 10, no. 4, article 97, 2018.

[27] A. Karine, A. Toumi, A. Khenchaf, and M. El Hassouni, "Radar target recognition using salient keypoint descriptors and multitask sparse representation," *Remote Sensing*, vol. 10, no. 6, pp. 843-844, 2018.

[28] H. Kawasaki, N. Soma, and R. Kretsinger, "Molecular dynamics study of the changes in conformation of calmodulin with calcium binding and/or target recognition," *Scientific Reports*, vol. 9, no. 1, pp. 24–29, 2019.

[29] J. K. Deter and Y. Rybarczyk, "Hidden Markov model approach for the assessment of tele-rehabilitation exercises," *International Journal of Artificial Intelligence*, vol. 16, pp. 1–19, 2018.