Video Shot Boundary Detection and Sports Video Classification Algorithm Based on Particle Filter

Dongsheng Chen[1] and Zhen Ni [2,*]

[1] Guangxi College for Preschool Education, Guangxi 530022, China
[2] School of Physical Education & Health, Nanning Normal University, Nanning, Guangxi 530001, China

Abstract

INTRODUCTION: Sports video is an essential information resource. Classifying sports videos with high accuracy can effectively improve users' browsing and query effect. This project intends to study a motion video classification algorithm based on deep-learning particle filters to solve the problems of solid subjectivity and low accuracy of existing motion video classification algorithms. A critical box extraction method based on similarity is proposed. The moving video classification algorithm is studied based on a deep learning coding model. Examples of various types of sports videos are analyzed. The overall performance of the motion video classification algorithm proposed in this paper is much better than other existing motion video classification algorithms. This algorithm can significantly improve the classification performance of motion video.

*Corresponding author. Email: 160012@nnnu.edu.cn

1. Introduction

With the rapid development of multimedia and Internet technology, video data on the Internet shows an explosive growth trend. For a considerable amount of video, it is impossible to annotate it artificially. In addition, because of the subjective nature of artificial marks, text-based information retrieval has been unable to meet users' needs. Therefore, it is essential to realize the effective processing and query of video and to classify it automatically. In medical care, network supervision, intelligent video surveillance, and other fields, automatic video classification are also of great significance: for example, classifying videos to screen out harmful videos to meet network supervision needs and ensure online video quality [1]. An image classification algorithm is used for wireless endoscopy images in medical treatment. The image is divided into four parts: entrance, stomach, small intestine and colon, which can reduce the workload of medical staff. In 2001, a group of researchers proposed a new approach to multimodal monitoring. This method proposes an algorithm for pedestrian detection and classification based on multi-view and applies it to long-distance environments.

Fisher first proposed the content-based classification of video types in 1995. It divides video into categories such as news, sports, commercial advertising and animation and puts forward three steps to classify video: extracting visual and audio features of video frames [2]. The image segmentation, camera motion and target motion obtained in the first step are analyzed in the second step. The third step is to classify the image according to its features. Since then, researchers have continuously improved and innovated in several fields, such as classifiers and video features, which have enabled the development of classification-based videos. Some scholars extract the text based on the data generated by users. Some researchers have recently proposed an acyclic support vector machine

(SVM) that fuses multiple image information [3]. Its standard features are editing, color, and movement characteristics. Previous studies have combined the Mel frequency band's cepstrum value with the image's color statistical histogram to classify different types of videos. Some people identified four types of videos in their doctoral dissertations. Firstly, football, badminton, basketball and table tennis are classified on the color information of the court. Then, distinguish basketball and table tennis according to the way of movement.

At last, football and badminton are identified by the texture feature of the image. Some researchers have proposed image recognition algorithms based on temporal patterns [4]. Firstly, the different actions that can take place in the video are defined. Different videos are classified using the machine learning method. Finally, a video classification algorithm is proposed based on a time series pattern.

However, the existing video classification algorithms have three common problems: 1) Video classification methods often design classification rules according to prior knowledge. It has certain limitations. 2) Use the classifier to learn and classify the complete image. This will inevitably bring about errors, which will affect the accuracy of recognition. 3) The classification method is too ideal. Simultaneously, an image file can include different kinds of image fragments. In general, this method can only classify one type of image. Given the shortcomings of the existing motion video recognition algorithms, this project intends to study the particle filter algorithm of motion video recognition [5]. The results are expected to have some reference significance for classifying motion video and even images.

2. Feature extraction of motion video based on deep learning

2.1 Feature Extraction

A feature represents an object, some attribute that can be quantified. There are two elements in motion video recording: universal and specific domain features. Considering the extraction speed of keyframes in moving video, the image characteristics of moving video are set as color histogram and color branch descriptor [6]. Color characterization of moving images is a color space problem. This paper presents a method to determine the similarity of keyframes using the edge of the shot. When the image is in the HSV color space, the index of the histogram is performed first. The similarity between the two frames f_a and f_b of the motion video histogram $\Xi(h,s,v)$ can be expressed as follows:

$$sim(f_a, f_b) = \sum_{h=0}^{15}\sum_{s=0}^{3}\sum_{v=0}^{3} \min\left\{\Xi_a(h,s,v), \Xi_b(h,s,v)\right\} \quad (1)$$

0 indicates a significant color column difference between

the two figures. One indicates that the color column difference between the two figures is the same. Color segmentation represents color regions in moving images [7]. The process of feature extraction is to segment the moving image. The γ, H, l component in 64 pixels is transformed by Discovery, and three sets of coefficients are obtained. Finally, the discrete cosine coefficients obtained are analyzed by Siegszag's method. Select fewer parts and construct new branch descriptors. Set to $Q\gamma, Q\delta, Q\varepsilon$ in order. The interval between frames f_a and f_b is set to:

$$Dist = \sqrt{\sum_{h=1}^{64} \varpi\gamma h\left[Q\gamma_{jh} - Q\gamma_{jh}\right]^2} + \sqrt{\sum_{h=1}^{64} \varpi Hh\left[QH_{jh} - QH_{jh}\right]^2} + \sqrt{\sum_{h=1}^{64} \varpi lh\left[Ql_{jh} - Ql_{jh}\right]^2} \quad (2)$$

$Q\gamma_{jh}$ represents the h term of the discrete cosine factor of the γ component in frame f_a. ϖ indicates the weight.

2.2 Edge detection of motion video

The lens edge factor is set according to the motion video field transformation characteristics. Suppose the width of the field window is $2N+1$. So, for frame a, the difference between adjacent Windows is

$$Q_{sw}(a) = \sum_{b=1}^{N} \frac{N-b+1}{G} Q(a-b, a+b) \quad (3)$$

$G = G(N+1)/2$. Changes in shots are higher than changes between shots. If the spacing between two arbitrary frames is Q_s, and the spacing between the two frames is Q_b, then $Q_s < Q_b$. If a lens mutation occurs between frames $N=3$ and G and G+1, then the constant sequence $K_{sw}(b)$ is (1, 3, 6, 6, 3, 1). Set the a frame lens boundary similarity factor to:

$$P_{sbs}(a) = \frac{\sum_{b=-N+1}^{N}\left[Q_{sw}(a+b)K_{sw}(N+S)\right]}{\sqrt{\sum_{b=-N=1}^{N}\left[Q_{sw}(a+b)\right]^2}\sqrt{\sum_{b=-N=1}^{N}\left[K_{sw}(N+b)\right]^2}} \quad (4)$$

When converting between adjacent frames, if the similarity coefficient of the lens edge is close to 1, in the rest, it is a minimal value from 0 to 1.

2.3 Keyframe sequence clustering

The same sports video will appear in the same momentum,

which leads to the repetition of keyframes. K-means clustering is performed on the keyframe sequence to reduce the repeatability of the final keyframe sequence [8]. Finally, the K value is determined by the clustering efficiency method. The characteristic parameters of clustering are:

$$F = i_a Scat(a) + dis(a) \qquad (5)$$

F represents the cluster result of the critical frame sequence in the motion video. $Scat(a), dis(a)$ specifies the sequential class of keys. i_a, i_b indicates the spacing between groups in sequence. A weighting factor is set since there is a large difference in the range of values between the two. d_{max} represents the preset maximum number of clusters. When the value reaches the minimum, d is the optimal number of clusters.

2.4 Depth coding method of motion video image

The critical frame information of the moving video is extracted, and the moving video is classified by deep learning of particle filter. The specific process is as follows:

1) Set the important frame feature database of the moving image to $F = \{f_1, f_2, \text{L}, f_N\}$. N is the number of features.

2) FF is coded using an unsupervised, finite Boltzmann machine (RBM). The encoding book of convergent parameters is obtained by using this method.

3) Carry out error non-forward propagation of the obtained label information. The supervised fine adjustment of the RBM neural network is realized. The F-code is encoded, the optimal image dictionary is obtained, and the image features are extracted.

4) The SVM classifier is described by the depth coding vector for training motion video images to complete the motion video classification.

3. Improve the AdaBoost algorithm based on the particle filter method

AdaBoost method is an up-and-coming method. The theoretical basis of this method is strict, and the prediction accuracy is high. This method will enhance the ability to distinguish incorrect samples in the previous cycle with each cycle, but the primary classifier will produce greater sensitivity to most categories of data after several cycles [9]. Therefore, the final classification algorithm can not predict the user well under the unbalanced sample set. The classic AdaBoost algorithm is improved. The following will give the specific steps to amend the AdaBoost algorithm. The flow of the modified AdaBoost algorithm is shown in Figure 1.
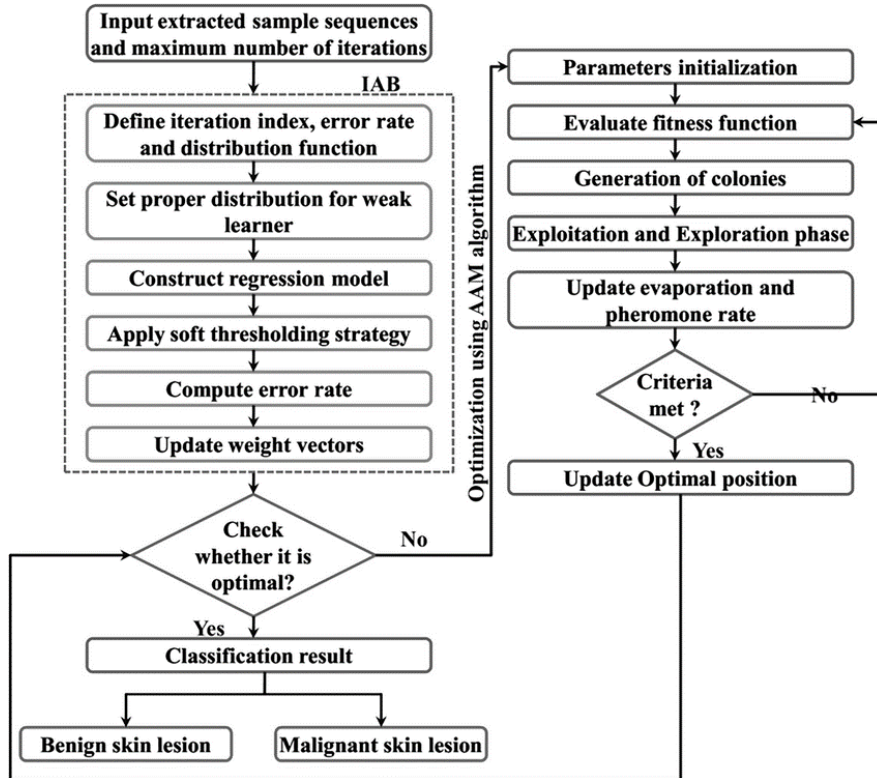

Figure 1. Flow chart of improved AdaBoost algorithm

1) Initialize the weighted distribution of all training samples:

$$S_1 = (w_{11}, w_{12}, \cdots, w_{1i}, \cdots, w_{1M}) \qquad (5)$$

$w_{1i} = \dfrac{1}{M}, i = 1, 2, \cdots, M$. S_1 represents the original weighted distribution. w_{1i} represents the initial weight of the sample i. M represents the number of samples. When the weighted values of two samples are equal, it indicates that each sample has equal importance in the classifier.

2) Repeated iterations. The F_n classifier is obtained by training the samples whose use value is assigned as S_n. Class error ε_n was calculated for classifier F_n.

$$\varepsilon_n = H(F_n(x_i) \neq y_i) = \sum_{i=1}^{M} w_{mi} U(F_n(x_i) \neq y_j) \quad (6)$$

The method aims at the most minor classification error ε_n and is based on the most minor classification error in each cycle. Here x_i is the actual fraction of sample i. y_i represents the classifier value predicted by the classifier for sample i. w_{mi} represents the weight of the i sample in the F_n cycle. The classification error ε_n of F_n is described by the weighted sum of the incorrect samples selected by F_n. The results show that the weight reassignment of classifier F_n on the training set has a great influence on its classification effect.

3) Calculate the weighting factor α_n for F_n. α_n represents the importance of $F_n(x)$ in the final classifier:

$$\alpha_n = \frac{1}{2} \log \frac{1 - \varepsilon_n}{\varepsilon_n} \qquad (7)$$

$\varepsilon_n \leq 0.5$ and $\alpha_n \geq 0$ can be derived from formula (3). The α_n value of taxon F_n increases with the decrease of ε_n value.

4) When class error ε_n is lower than the threshold or reaches the maximum number of repeats, the repeat terminates.

5) Update the weight distribution in the training data set, then go back to step 2.

$$S_{n+1} = (w_{n+1,1}, \cdots, w_{n+1,i}, \cdots, w_{n+1,N}) \qquad (8)$$

In formula (4), the AdaBoost algorithm increases the weight of the wrong or wrong classification samples. Finally, the weighted function S_{n+1} is obtained by modifying the weighted function [10]. The concept of cost sensitivity is used to improve the method, and the cost sensitivity coefficient is introduced. Modified sampling weighted modified formula. Here is the detailed calculation method:

$$w_{n+1,i} = \frac{w_{ni}}{Z_n} \exp(-\alpha_n y_i F_n(x_i) \times \zeta_a), \qquad (9)$$

$$F_n(x_i) = a; y_i \,!= a, a = 1 \cdots Y$$

$$Z_n = \sum_{i=1}^{M} w_{mi} \exp(-\alpha_n y_i F_n(x_i)) \qquad (10)$$

Here Z_n is the normalization factor, so S_{n+1} is a probability distribution. Where w_{mi} is the weighted value of n iterative sampling points i. Where $w_{n+1,i}$ is the weighted value obtained after sampling i for $n+1$ iterations. Where ζ_a is a cost-sensitive factor. Y is the sample fraction. It can be seen from formula (5) that when the sampling weight is updated in each iteration, the modified AdaBoost algorithm further analyzes the sampling fraction to distinguish between correct sample classification and incorrect sample classification. Multiply the true sample score x_i of the misclassified sample by the cost-sensitive factor ζ_a. The weights of different misclassified samples are increased in different proportions. The various types of error categories are weighted according to a certain proportion. The classification algorithm can improve classification performance effectively by repeated training. Finally, the effectiveness of the proposed method is verified by experiments. A method of generating a weakened classifier based on wavelet transform is proposed. This method has the same spatiotemporal complexity as the existing AdaBoost algorithm. A classification method based on a single feature is proposed. Where K represents the number of training samples. Where d is the number of sampling features. Where T is the number of poor classifiers. So, this method is relatively fast.

The cost sensitivity factor ζ_a has a great influence on the final classification effect. When the cost sensitivity coefficient is iterated, the particle swarm optimization method is used to modify the cost sensitivity coefficient ζ_a. The particle swarm optimization method is simple to construct and easy to control. This method is a heuristic global optimal solution. Set the fitness function. In this algorithm, the particles are co-optimized according to the fitness values of the population and the particles themselves. This new solution method can ensure fast convergence. A single exception does not affect the global optimal solution of PSO. It's very stable. Here are the specific formulas

$$\varphi_i(t+1) = \varphi_i(t) + \zeta_1(pbest_i - X_i) + \zeta_2(gbest_i - X_i) \,(11)$$

$$X_i(t+1) = X_i(t) + \varphi_i(t+1) \qquad (12)$$

The t in formulas (7) and (8) represents the number of

repetitions. φ_i is the particle's speed, which determines the direction and distance of the particle's movement. X_i represents the location of this particle. ζ_1 and ζ_2 are the two values of the normal distribution. $pbest_i$ is the overall best solution for this iteration. Where $gbest_i$ is the historically optimal solution. The method is based on random selection of a set of cost-sensitive factors. The algorithm tracks two extreme points in each cycle, namely a single extreme point $pbest_i$ and a global extreme point $gbest_i$. Each particle adjusts its search direction by adjusting (1), (2) and other equations to find the best cost

sensitivity.

4. Experimental results and analysis

This paper takes Windows XP as the test platform. This paper uses Matlab 2010 as the implementation platform of the system. Three sets of data are used in the experiment to check the correctness and validity of the algorithm. Its basic features are listed in Table 1. Each category in Data Set A contains more than 66 videos. The Data Set has 18 videos per category. Data Set C is a collection of videos that combine various movements.

Table 1. Experimental data

Data set name	Video content	Use	Number	Duration /min	Size /GB
A	Basketball, badminton, football, table tennis, snooker, tennis, volleyball, a total of seven categories of video	Verify the effectiveness of the convolutional neural network algorithm and BP neural network algorithm	1049	3125	11.67
B	There are three types of videos: basketball, football and tennis	Verify the effectiveness of the convolutional neural network algorithm and BP neural network algorithm	47	260	1.25
C	A combination of two or three videos randomly selected from a single video	Detect the effect of the BP neural network algorithm on mixed video	48	302	1.46

Recall rate and accuracy are measured, where a is the number of positive cases. b represents the number of positive cases in the test sample. c is the number of positive examples obtained by classification.

4.1 Video image classification

The classification results of sports video frames based on the particle filter model are presented. This training database has seven categories of motion videos, including two marker-shot motion videos and nine non-type marker-shot videos. The training frame is 9x100,900. The measured video frame of the picture is 9x40, a total of

360 frames.

Figure 2 shows the classification of footage of players with and without class-labelled shots based on particle filtering (image cited in Sports Videos in the Wild (SVW): A Video Data Set for Sports Analysis). The diagonal line in Figure 2 is the recall rate of each video frame classification. The other row, i, column j, lists the possibility of a video frame being incorrectly identified as type i. Overall, the lens video picture with a category mark is better than the one without a category mark. This is due to the variety of lenses without model marks. The class-labelled lens is thinner than the non-class-labelled lens.
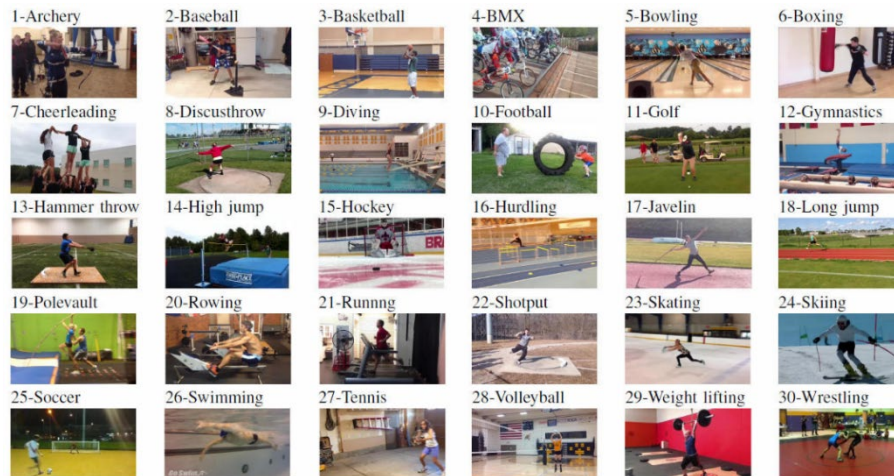
Figure 2. Video frame classification results

4.2 Tests to Exclude Irrelevant Frames

It can be seen from Figure 2 that the representation frame classification based on the particle filter model has achieved good results, but there are still errors in some representation frame classifiers. The correctness of the method is tested by randomly selecting a segment from the data set for different types of motion videos. Finally, recall rate and accuracy were used to evaluate the effect of eliminating outlier frames in time series. The results of the experiment are shown in Table 2.

Table 2. Test results without independent frames

Video type	Number of isolated frames			R_s^1 / %	R_s^2 / %
	A	B	C		
badminton	7	5	5	99.38	99.70
basketball	25	10	10	76.04	87.81
Table tennis	17	15	16	86.04	101.88
volleyball	14	11	11	89.17	101.88
tennis	20	10	15	94.90	99.79
soccer	35	17	20	67.50	81.25
snooker	3	2	2	97.68	99.90

In Table 2, A represents all the frames classified. B refers to the number of isolated frames correctly removed by the convolutional neural network algorithm. C represents the number of frames of interference eliminated in the convolutional neural network algorithm. R_s^1 is the recall rate of a video frame classifier in a typical frame classification. R_s^2 is to eliminate the lone wave frame in the sample frame by convolutional neural network in the process of sample frame classification to improve the sample recall degree. The experimental results show that this method can improve the recall rate of the image feature frame classifier. The algorithm effectively reduces the influence of interframe identification errors on video classification.

4.3 Single sports video classification experiment

Single sports video is a particular case of mixed sports video, so this method is also suitable for classifying single-moving images. In this paper, recall rate and accuracy are used to evaluate the performance of convolutional neural networks. The recall rate, accuracy and other indexes are used as BP neural network performance indexes. The experimental results show that combining the convolutional and BP neural networks influences single-moving image training. This experiment used the data set as a test video collection. The threshold for the BP neural network is T=8. The results of image classification thus obtained are shown in Table 3.

Table 3. Results of single video classification

Video type	Number	Convolutional neural network algorithm		BP neural network algorithm	
		R_s / %	P_s / %	R_m / %	P_m / %
badminton	113	98	103	96.46	95.25
basketball	259	99	102	96.67	98.38
ping-pong	92	99	103	99.48	99.48
soccer	272	99	103	96.98	98.18
volleyball	128	103	92	97.68	96.15
tennis	61	102	92	97.08	97.40
snooker	124	102	101	99.29	92.22

The convolutional neural network and BP neural network algorithms were used to classify several sports vi: basketball, badminton, football, table tennis, snooker, tennis and volleyball. And he got good results. For a single-motion video, the BP neural network's recall rate and accuracy are slightly lower than the convolutional neural network method.

The algorithm in this paper is tested on Data Set A and Data Set B, compared with the results in the literature [11]. In this paper, four types of sports videos are classified and tested. There's badminton, there's table tennis, there's basketball, there's football. Figure 3 shows the results of the comparison. Literature [12] provides recall rates for classifying basketball, football and tennis sports videos, so this paper tests the convolutional neural network algorithm on Data Set B. The recall rate of three types of videos was compared.

It can be seen from Figure 3 that for video test set data SetA and data SetB, both convolutional neural network and BP neural network have achieved good results. However, because there are a large number of interference shots in the tennis video on Data Set B, the error detection rate of these interference shots is relatively high. Therefore, compared with basketball and football, the classification effect of tennis video using convolutional neural network algorithm is slightly inferior [13]. At the same time, it also shows that the interference lens in the image will have a great impact on the image recognition.
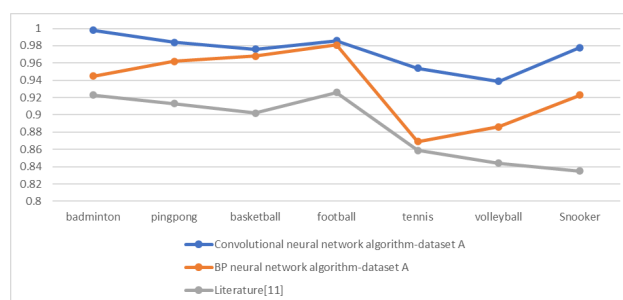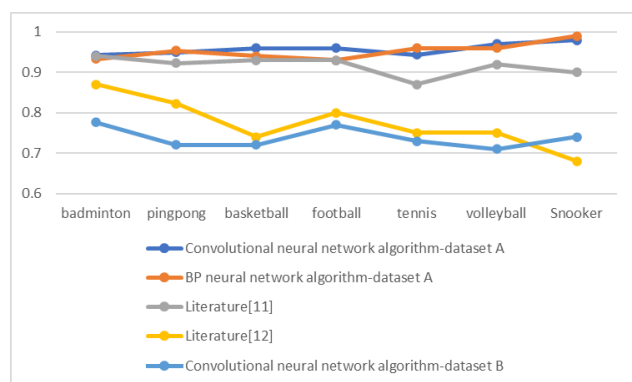
Figure 3. Comparison of the effect of a video exercise test

4.4 Mixed Sports Video Classification Experiment

Using the SetC data set as a test sample, the classification method of hybrid motion video based on the BP neural network is experimentally studied. Because the BP neural network has some dependence on parameter T, the effect of the BP neural network on T change is studied. The results of the experiment are shown in Figure 4. As can be seen from Figure 4, the BP neural network has the best classification results when T=8 and 9.
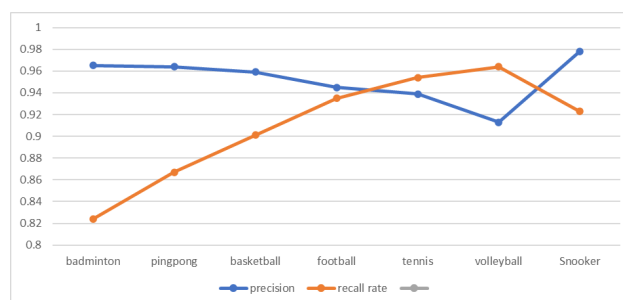
Figure 4. The effect of the classification method of hybrid motion video based on BP neural network under T-value

When classifying a single moving image, the convolutional neural network has better classification results than the BP neural network for mixed moving images, but for mixed moving images, the classification results of the convolutional neural network are not ideal. The learning method of the BP neural network is suitable for a single motion picture for a variety of motion pictures. which has better versatility than a convolutional neural network. However, these two methods have advantages and disadvantages in practical application, so there are some differences. This method effectively classifies single-moving images, making it suitable for knowing them in advance [14]. The BP neural network learning method is more suitable for learning unknown information. Sports videos classified can include a single video or multiple videos [15]. Chakraborty et al. proposed a new shot boundary detection method to optimize the weights of feedforward neural networks (FNNs). In order to improve the performance of the system, the output of the hybrid technology is analyzed again by forming a continuity matrix (phi) [16]. Video annotation, as the foundation of video indexing and retrieval, has important application prospects and research value. In semantic detection, the detection and tracking of moving targets are the foundation [17].

4.5 Experiment on image recognition effect of interference lens

Two control experiments are done in this paper: only the type flag shot class is in the video frame training library of the group. The other is the training base of an existing type labelled video frame. Through the example analysis of the method, it is proved that adding an unmarked lens category to the method can effectively improve the recall and accuracy of the method. In this paper, the data set SetA is used as a test sample to study the influence of disturbed scenes on the performance of convolutional neural networks. Recall rate and accuracy rate were used as evaluation indexes. The results of the comparison between the two test groups are shown in Figure 5. It can be seen from FIG. 5 that adding class-free shot categories to the training library of a video frame can significantly improve the classification performance of the proposed algorithm.
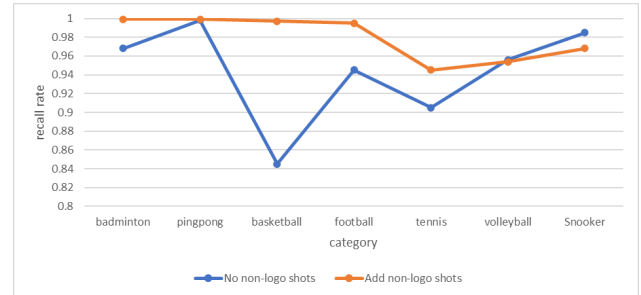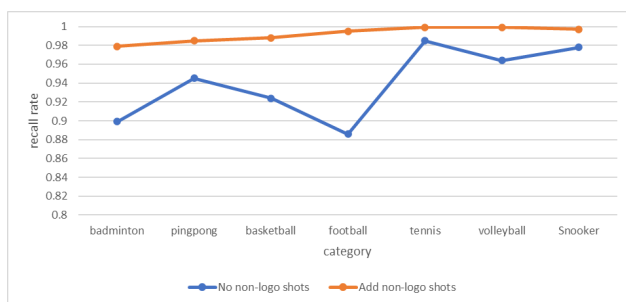

Figure 5. The role of interference lens in video recognition

5. Conclusion

The single-motion video and the mixed-motion video are classified. Adds a class of video frames with classless flags to the training program. The outlier frame removal algorithm and particle filter algorithm are used to represent the result sequence of frame classification to improve the accuracy of the representation frame classifier. Experiments show that the proposed algorithm performs better than other automatic video classification algorithms based on moving images. The method has the advantages of simple implementation, fast processing speed and high precision. In addition, when the types of moving video increase, the classification of moving video images can be realized only by adding training samples of video images corresponding to different types of marked scenes. It has good adaptability and popularization performance.

Data availability statement
The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest
The authors declare that there are no conflicts of interest regarding the publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

[1] Yuan, Y., Lu, Z., Yang, Z., Jian, M., Wu, L., Li, Z., & Liu, X. Key frame extraction based on global motion statistics for team-sport videos. Multimedia Systems,2022; 28(2):387-401.

[2] Shingrakhia, H., & Patel, H. Cricket Video Highlight Generation Methods: A Review. ELCVIA Electronic Letters on Computer Vision and Image Analysis, 2022; 21(2):1-22.

[3] Soundararajan, K. Sports highlight recognition and event detection using rule inference system. Concurrent Engineering, 2022; 30(2):206-213.

[4] Fu, Q., Ma, L., Li, C., Li, Z., Zhu, Z., & Lin, Z. Detection method of sports scene conversion for MPEG compressed video based on fuzzy logic. Journal of Intelligent & Fuzzy Systems, 2021; 40(2):3107-3115.

[5] Guntuboina, C., Porwal, A., Jain, P., & Shingrakhia, H. Deep learning based automated sports video summarization using YOLO. ELCVIA Electronic Letters on Computer Vision and Image Analysis, 2021; 20(1):99-116.

[6] Chakraborty, S., Thounaojam, D. M., & Sinha, N. A shot boundary detection technique based on visual colour information. Multimedia Tools and Applications, 2021;80(3):4007-4022.

[7] Chakraborty, S., Singh, A., & Thounaojam, D. M. A novel bifold-stage shot boundary detection algorithm: invariant to motion and illumination. The Visual Computer, 2022; 38(2):445-456.

[8] Chakraborty, S., & Thounaojam, D. M. SBD-Duo: a dual stage shot boundary detection technique robust to motion and illumination effect. Multimedia Tools and Applications, 2021; 80(2):3071-3087.

[9] Cao, Y., & Mao, H. High-dimensional multi-objective optimization strategy based on directional search in decision space and sports training data simulation. Alexandria Engineering Journal, 2022; 61(1):159-173.

[10] Tiwari, V., & Bhatnagar, C. A survey of recent work on video summarization: approaches and techniques. Multimedia Tools and Applications, 2021; 80(18):27187-27221.

[11] Raval, K. R., & Goyani, M. M. A survey on event detection based video summarization for cricket. Multimedia Tools and Applications, 2022; 81(20):29253-29281.

[12] Shingrakhia, H., & Patel, H. SGRNN-AM and HRF-DBN: a hybrid machine learning model for cricket video summarization. The Visual Computer, 2022; 38(7):2285-2301.

[13] Liu, Y., Dhakal, S., & Hao, B. Multimedia image and video retrieval based on an improved HMM. Multimedia Systems, 2022; 28(6):2093-2103.

[14] Singh, A., Singh, T. D., & Bandyopadhyay, S. V2t: video to text framework using a novel automatic shot boundary detection algorithm. Multimedia Tools and Applications, 2022; 81(13):17989-18009.

[15] Adly, A. S., Hegazy, I., Elarif, T., & Abdelwahab, M. S. Indexed Data Set from YouTube for a Content-Based Video Search Engine. International Journal of Intelligent Computing and Information Sciences, 2021;21(1):196-215.

[16] Chakraborty S, Thounaojam D M. A novel shot boundary detection system using hybrid optimization technique. Applied Intelligence, 2019, 49(9): 3207-3220.

[17] Zhang B. Moving target detection and trackingbased on camshift algorithm and kalmanfilter in sport video. International Journal of Performability Engineering, 2019, 15(1): 288.