Journal of Information Hiding and Multimedia Signal Processing ©2017 ISSN 2073-4212
Ubiquitous International Volume **8**, Number **5**, September **2017**

Research on the Video Advertising Detection Based on PNN and Text Detection Method

Xiao-ling Lan

The Art College Lu Dong University
No.186, HongQi Road, ZhiFu District, Yantai, 264025-China
848446636@qq.com

Jia Cui and Shu-tuan Zhang

Department of Control Engineering
Naval Aeronautical and Astronautical University
No.182, Two Ma Road, ZhiFu District, Yantai, 264001-China
beyond2002@163.com

Received July, 2016; revised June, 2017

Abstract. *Aimed to solve the false detection problem of video advertising detection because of low frequency in shot cut, the video advertising detection method based on the particle swarm optimization algorithm and text detection is introduced after analyzing and studying the feature of video advertising. Firstly, the video advertising detection based on the particle swarm optimization algorithm and neural network algorithm can be realized according to the shot cut rate. Secondly, the video advertising of the non-video advertising in the first test result can be acquired by using the text detection method. At last, the sum of the two parts video advertising detection results is the number of the video advertising in this video. Compared to the method only using the particle swarm optimization optimized neural network algorithm (PNN) the accuracy of video advertising detection based on the PNN algorithm and the text detection algorithm has improved. The experimental results show that the precision and recall are improved, and the method of video advertising detection can realize the accuracy detection*
Keywords: Video advertising detection; PSO algorithm; Neural networks

1. **Introduction.** With the development of multimedia technology, digital television and the Internet, the number and type distribution of video information continues to rapidly grow and expand. Among so much multimedia information, for example, the text, graphics, images, sound, video, animation, video information with vividness and intuition is highly favored which creates a huge number of video files. These videos have great amounts of data and information content and how to effectively store, sort, retrieve, and use them becomes an important and urgent issue.

At present, a variety of video ad resists in network or television that has a great impact on society. The video advertising detection is very important to strengthen advertising management and review. It also can reduce and eradicate illegal advertising in order to monitor advertising programs; we firstly should distinguish between TV programs and commercials. In order to successfully detect video advertising, we need to extract the unique feature of most advertising video because it changes fast and its characteristics is very complex. So, how to select the features of video advertising is essential. Generally the shot effect of video advertising relative to other programs is quite lively, dynamic and

it also has fast scene switch. So it can easily attract audiences. The advertising detection can be completed by taking the feature of high shot cut frequency. Literature [2] has proposed a novel feature representation method, which formed a new feature space, and trained SVM in this space to recognize the text; literature [3] combined fuzzy clustering with characteristics of camera motion, shot detection base on fuzzy clustering is reported in the paper. The algorithm needs not select threshold and reduces the impact of flash, subtitles and advertising on shot detection effectively. However, due to the complexity of the video itself, there are still missed and false detections in this algorithm and research is further needed. In literature [4], a slope-based detecting method of gradual change shot is presented. Literature [5] mainly studies an automatic seed algorithm, and applies it to iterative graph cut algorithm, finally completes text segmentation. Combining the research results, text detection method [6] can be used to detect advertising information to improve the precision of detection. According to the different feature of the video advertising, the different methods can be used to detect the video advertising by classification. The video shots segmentation methods can be used for those video advertising with high switching frequency. Text detection methods can be used for those video advertising with low switching frequency and commercial showing.

Aim at the accurate detection requirements of the video advertising, on the basis of original scholars research, the new method of the video advertising detection based on the particle swarm optimization algorithm and the neural network algorithm combining text detection is introduced in this paper, in order to improve the accuracy and recall of video advertising detection.

2. **Neural network and Particle Swarm algorithm.**

2.1. **Neural network algorithm.**

The Neural networks (NN) algorithm is an important branch of artificial intelligence, which is formed by the large number of neuron connection to simulate the behavior of complex network system of the human brain.

Feed forward network is also known as forward network, which is characterized in that the neurons are connected to each other, and there is no feedback between the two layers. Each neural network can receive multiple inputs from the previous layer, and generate an output that is passed to the next layer of individual neurons, and the information from the input layer begins to move forward sequentially. The feed forward network can include many layers, but we have proved that three layer feed forward network is sufficient to meet requirements for solving the problem The feed forward network is composed of three layers, which are input layer, hidden layer and output layer.

One of the classifiers used in this paper is the typical feed forward network, the back propagation network (BP network), and the topology of the network is shown in Figure 1.

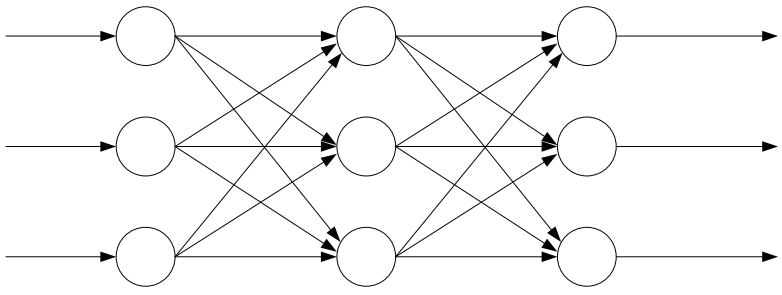

FIGURE 1. The structure figure of BP neural network

The excellent information processing ability of the neural network is from the connection weights between the neurons in the network. To simulate the human brain's information processing ability, we must make the neural network with learning function. The essence of learning is to adjust the weights of the connections between the neurons, so how to adjust the weights of the connections constitute the learning algorithm of the neural network.

Among the most typical forward networks, the BP neural network is the most typical one. BP network introduces the least squares learning algorithm. In the learning process, the output error of the network and the expected output s while, it obtain the minimal value of mean-square error between the actual output and the expected output of the network as spreading as correcting the connection strength (weighted coefficient). The learning process can be divided into two parts: forward network calculation and backward error propagationtwo parts of connection weight coefficient. These two parts are continuous and repeat until the error meets the requirements.

2.2. **Particle swarm optimization algorithm.** Particle swarm optimization [2] (PSO) is an optimization algorithm based on swarm intelligence theory proposed by J.Kennedy and R.C.Eberhart in 1995. This based on the simulation of the migration and aggregation of birds to forage. The algorithm has the characteristics of fast convergence speed, easy to implement and the less adjustment of the parameters which is a new research hotspot in the field of intelligent optimization and evolutionary computation.

The basic idea of PSO is that the potential solution of each optimization problem is to search the particles of space, and all particles have a fitness function to determine the fitness value, and each particle also has a velocity vector to determine the direction and distance of particle flight, and then the particles will follow the current best particles in the solution space to search. In PSO, each particle also the solution of optimization problem is considered as a point in the search space. All particles have a property representing the position of the current in the solution space $X_i = \{x_{i1}, x_{i2}, \cdots, x_{im}\}$; each particle also has the speed to determine the direction and distance of their movement $V_i = \{v_{i1}, v_{i2}, \cdots, v_{im}\}$t. In the solution space, the information of the particles is shared by the current best particles. First, the particle swarm is initialized to a group of random particles (random solution), then the optimal solution is obtained by iteration. In each round of iteration, the particles will follow the two extremes, one for the particle itself to find the optimal solution p_{best}, and the other for the whole population to find the optimal solution g_{best}.

The mathematical description is introduced in a D dimension search space The position of the i particles is $x_i = \{x_{i1}, x_{i2}, \cdots, x_{iD}\}^T$, the speed is $v_i = \{v_{i1}, v_{i2}, \cdots, v_{iD}\}^T$. Its individual extreme value is $p_i = \{p_{i1}, p_{i2}, \cdots, p_{iD}\}^T$; the global extreme value of the population is $p_g = \{p_{g1}, p_{g2}, \cdots, p_{gD}\}^T$. According to the principle of following the current optimal particle, the particle x_i will update its position and speed by the following formulas.

$$v_{id}(t+1) = \omega v_{id}(t) + c_1 r_1 \left[p_{id}(t) - x_{id}(t)\right] + c_2 r_2 \left[p_g(t) - x_{id}(t)\right] \tag{1}$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \tag{2}$$

Among them, $i = 1, 2, \cdots, m$, $d = 1, 2, \cdots, D$, m is for the population size, t is for the current evolutionary algebra, r_1 and r_2 are for the uniform distribution of the $[0, 1]$, c_1 and c_2 are for the learning factor (non negative constant), usually the value is 2. $x_{id}(t)$ is the current position of the i particles. p_{id} is the best position to search the entire particle swarm optimization. p_{gd} is the optimal location of the whole particle swarm search. $v_{id}(t)$ is the current speed of the particles. ω is for the inertia weight. The influence of the

particle velocity on the current generation is described, which makes the particles keep moving inertia and the trend of expanding the search space. The speed at each one of the particles is limited to a maximum speed $v_{\max}$. If the speed updated of each dimension exceeds the user's setting $v_{\max}$, the speed limit is defined as $v_{\max}$.

3. **Shot segmentation.** In general, the video can be used to frame (Frame), the shot (Shot), the scene (Scene) and the (Program), etc. The frame is the smallest unit of the video, where the key frame contains important information, the key frame extraction is the research hotspot; the camera is composed of continuous video frames, which describe a series of events or a series of continuous movements. As a whole, the video can be seen as the 3D space, the horizontal and vertical direction of the video image is two dimensional, and the third is the axis of time, and the content of the video image changes with the time axis. Video shot boundary detection is to segment the video from the time axis into some continuous video shots. Video shot boundary detection method is mainly in order to find location of the interrupt on the video content t.

The N number of frame in the time axis of the video stream is extracted t is the number of some frame on the edge of a candidate shot boundary. Take t as the center and select a window which size is $2 \times d$. Each frame in the window is a node. Connecting any two points can be constructed with the right map.

Connecting the weights w_{ij} of the edges of any two points reflects the similarity of the two frame images, and the similarity between the two images is larger, and the possibility that the two images belongs to the same camera is larger. When the distance between the two frame images is very large, they are less likely to belong to the same camera. The weight w_{ij} between any two frames i, j in the video is defined as the equation (3), H_k^i represents the color number of the level k of the color histogram of the HSV model of the frame i.

$$w_{ij} = \sum_k \min(H_k^i, H_k^j) \begin{cases} e^{\frac{-\|i-j\|^2}{\sigma^2}} & if\ |i-j| \le d \\ 0 & otherwise \end{cases} \tag{3}$$

The minimum and maximum classification discriminate function can be obtained by the following methods. If E represents the d frame before t and F represents the d frame after t, then the correlation between E and F as the equation (4), which is defined as inter class distance, the internal correlation definition of the E as the equation (5), that is, within the range of categories.

$$cut(E, F) = \sum\nolimits_{i \in E, j \in F} w_{ij} \tag{4}$$

$$assoc(E) = \sum\nolimits_{i,j \in E} w_{ij} \tag{5}$$

From the perspective of video segmentation, when the distance between the two classes is larger, the distance inner the two classes is smaller, the more likely the E and F two parts can be switched between the two parts. Therefore, the minimum and maximum classification discriminated function is defined as follows the equation (6).

$$Mcut(E, F) = \frac{cut(E, F)}{assoc(E)} + \frac{cut(E, F)}{assoc(F)} \tag{6}$$

According to minimum and maximum classification discriminated function, the value of the $N-1$ possible cut point can be calculated, such as the equation (7), so as to construct characteristic curve of video stream.

$$score(t) = Mcut(\{t-d+1, \cdots, t\}, \{t+1, \cdots, t+d\}) \tag{7}$$

4. **Shot boundary detection.**

4.1. **Cut boundary detection.** The shear shot is also known as cut. This means that a shot directly switches to the next shot, and there is no delay in the middle. Cut shot only occurs between two adjacent frames and there is a significant change in the video stream continuous signal curve. A continuous signal of video stream containing a shear shot boundary is constructed according to the equation (11). Hypothetically, the wave trough of the curve shape of shot boundary is regular and symmetry. Sheer shot boundary detection can be completed by identification the shape of the wave trough. Firstly trough region extraction is through threshold. Then the characteristics of wave trough are extracted to judge whether it is a symmetrical trough shape.

The boundary detection algorithm for shear shot is shown as follows:

Step1: The continuous signal of video stream is constructed by using graph partitioning model. Take the image frame of the video as the point of graph partitioning model. Connect arbitrary two points to construct the graph partitioning model.

Step2: Calculate the weight value. The weights of the edges of the two frames are calculated by the equation (3).

Step3: Calculate the $N-1$ possible cut-off value According to the equation (6) and equation (7), calculate the $N-1$ possible cut-off value, and constructs the continuous signal of the video stream.

Step4: Identify the candidate regions and extract their features. First, candidate region is extracted by using the threshold method, this paper adopts the th=0.7 threshold, the value is less than the threshold value of the wave is extracted; then, find the minimum value S_t of each candidate region, in order to show the minimum value as the center to construct the feature vector B_t^{ε} of each wave trough as shown in the equation (8). Because when a shear occurs between the t frame and the $t+1$ one, the D frame before the t frame and the D frame the after one is affected, so $r = d = 4$ is taken.

$$B_t^{\varepsilon} = (S_{t-r}, S_{t-r+1}, \cdots, S_{t+r}) \tag{8}$$

Step5: Using the particle swarm optimization algorithm and the neural network algorithm to realize the classification

The shear shots are classified based on the particle swarm optimization algorithm and the neural network algorithm according to the acquired feature B_t^{ε} value.

4.2. **Gradient shot boundary detection with monochrome frame.** The gradual transition process can occur in multiple video frames, and the change of the adjacent two frames is not obvious. This has brought a great difficulty to the detection of the boundary of the gradient shot. Gradual shot change containing monochrome frame includes: rinse out, fade in and monochrome frame containing a special camera. In the process of dredging on a fade, the shot gradually disappeared and the emergence of a series of monochrome frames, then the next shot appeared.

A picture with a monochrome frame is changed from one part to a monochrome frame, and then a part of the frame is replaced by the next shot. Literature [9] detects fade out and fade in and monochrome frames containing special shot boundary. But it is not necessary that shot cut happens when monochrome frames emerge in the video.

So this paper first uses image segmentation model to construct characteristic curve, and then extracts fade out and fade in and the candidate region containing monochrome frame shot dash switching through the threshold, when monochrome frame takes place in the video and they all have shot switch, video signal curve does not have obvious trough shape. So this method can remove the area of monochrome frame in the video without occurring shot switching and narrows the search range, finally determines the candidate region by using monochromatic frame feature.

A gradient shot boundary detection algorithm with a monochrome frame is as follows:

Step1: The continuous signal of video stream is constructed by using graph partitioning model. Take the image frame of the video as the point of graph partitioning model. Connect arbitrary two points to construct the graph partitioning model.

Step2: Calculate the weight value. The weights of the edges of the two frames are calculated by the equation (7).

Step3: Calculate the $N-1$ possible cut-off value According to the equation (10) and equation (11), calculate the $N-1$ possible cut-off value, and constructs the continuous signal of the video stream.

Step4: Identify the candidate regions and extract their features. Candidate region is extracted by using the threshold method, this paper adopts the th=0.8 threshold, the value is less than the threshold value of the wave is extracted; the key of pre-detection step is to ensure high recall of detection result. So it is necessary to set a higher threshold algorithm, and to remove the shot shear area.

Step5: Monochrome frame detection

Calculate the color histogram of the RGB of each candidate region, namely R, G, B three color histogram, set the number of points in the three histogram is greater than 192 sum or the points number less than 64 sum, if sum is greater than the threshold, then the frame is a monochrome frame. The candidate region containing the monochrome frame sentenced to fade or change shot containing monochromatic frame boundary.

4.3. **Detection of gradient shot boundary detection without monochromatic frame.** It is difficult to detect the change of the gradient shot with no monochromatic frame because of the large number of signal characteristics, which has no obvious signal characteristics. Some of the process of gradual shot change without monochromatic frame can even occur in nearly more than 60 video frames, which is reflected in the video signal on the characteristics of continuous wave trough is very weak. According to the characteristics of different shot switching process without monochrome frame, the feature extraction can be realized by taking the method of multiresolution analysis.

When the process of the gradient shots without a monochrome frame switch is long, this signal is not easy to detect in high resolution, but it is easy to be detected in the low resolution, so it can be used for the analysis of the low resolution of the gradient shot with no color frame.

The gradient shot boundary detection algorithm, which does not contain a monochrome frame, is as follows:

Step1: The continuous signal of video stream is constructed by using graph partitioning model. Take the image frame of the video as the point of graph partitioning model. Connect arbitrary two points to construct the graph partitioning model.

Step2: Calculate the weight value. The weights of the edges of the two frames are calculated by the equation (3). Then $N-1$ possible cut-off value calculation. According to the equation (6) and equation (7) calculate the $N-1$ possible cut-off value, and constructs the continuous signal of the video stream.

Step3: Identify the candidate regions and extract their features. First, candidate region is extracted by using the threshold method, this paper adopts the th=0.8 threshold, the value is less than the threshold value of the wave is extracted:

$$\begin{cases} S_{begin} \approx S_{end} = \max\{S_{begin}, S_{begin+1}, \cdots, S_{end}\} \\ S_{begin} \leq T_{GT} \end{cases} \tag{9}$$

$$S_t = \min\{S_{begin}, S_{begin+1}, \cdots, S_{end}\} \tag{10}$$

Then, find the minimum value S_t of each candidate region, take the minimum value as the center to construct the feature vector B_t^{ε} of each trough, here σ were taken1, 3, 5, namely to extract the feature vector of one candidate region in 3 resolution as shown, here $r = 10$

$$B_t^{\delta} = \left(S_{t-r\times\delta}, S_{t-(r-1)\times\delta}, S_t, \cdots, S_{t+r\times\delta}\right) \tag{11}$$

Step4: Neural network integration based on particle swarm optimization is used for classification.

Firstly, the training of neural network integration. Using a trough feature B_t^{δ} extracted from different resolution, respectively to train neural network ensemble fewer than three kinds of resolution particle swarm optimization. Secondly, the test of the neural network integration. If there is an output for the output of the three neural network integration, the candidate region is the gradient shot boundary.

4.4. **Text detection.** Video text for image content description provides a lot of semantic information, and compared with other semantic content, it is easier to extract, therefore, the research of video text extraction is widely used in advertising automatic statistics, news automatic monitoring and video database indexing. A lot of text detection method is used. The following describes a method based on the maximum gradient difference method [11] (Maximum Gradient Difference, MGD) video text detection process.

Step1: Using the template $[-1, 1]$ to move on the original image P (top and bottom of the two rows, as well as the left and the right of the two columns are not processed), then achieve the gradient image T_x in horizontal direction of the image. In the same way, use the template $[-1, 1]^T$ to obtain the vertical direction of the image gradient T_y.

Step2: For horizontal directional gradient image T_x, the MGD value of the center pixel is calculated in the direction of the window in the $1 \times n$ window. The selection n is related to the size of the text to be detected, usually slightly larger than the stroke width of the character, and then set a threshold value MGD_thd that will be lower than the value MGD of the threshold value MGD_thd to zero, thereby getting the horizontal direction MGD of the image M_x.

Step3: For the vertical direction of the gradient image T_y, the MGD value of the pixel is calculated in the column direction, so as to get the vertical direction of the image M_y.

Step4: For M_x and M_y all the pixels which MGD value greater than zero are evaluated 255 (white), the other pixels assigned to 0 (black). At last, we can acquire the image Q.

4.5. **Video advertising detection system design.** The characteristics of video advertising are complex and multiple. Only we extract the unique features of most of the advertisement detection, detection of video advertising can be realized through the analysis, the shot with high switching frequency is advertising, and switching low rate is defined as non advertising. Undetected may exists at that certain extent. Therefore, to the video advertising with low rate shot switch, we detect trademark information of advertising by the text detection. In order to improve the detection precision and recall of video advertising, this paper designs a detection system based on video advertising particle swarm optimization neural network and the combination of text detection, and use the method of text detection in trademark text detection to process the advertising which is undetected by shot switch method. That further improves the detection accuracy. The specific flow chart is shown in Figure 2.

The video advertising detection system mainly includes two parts: first in shot cut detection, the HSV color histogram is used to construct the video image content. Then, according to the feature of the image segmentation model, we construct the video continuous signal and use the threshold method to extract the candidate region. Finally,

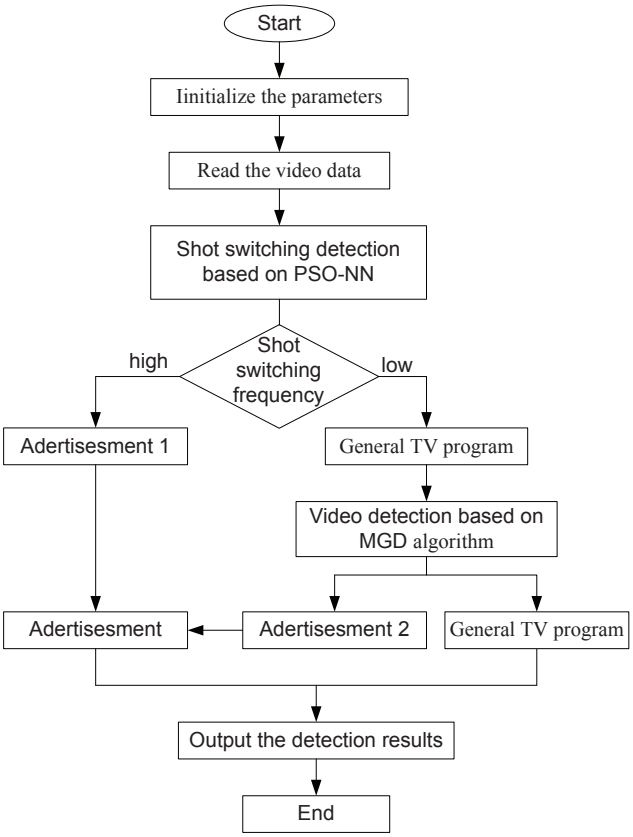

FIGURE 2. The flow chart of video advertising inspection system

according to the characteristics of the algorithm, the algorithm is realized by using the algorithm of the particle swarm optimization algorithm. Finally, we use the algorithm of particle swarm optimization. The algorithm realizes the decision of the shot boundary region of the candidate region. Finally, according to the characteristics of each shot, the shot boundary region is extracted. Through the use of the shot switch detection, we can obtain the shot switch frequency, high shot switch frequency is advertising and low one is non advertising. We first use maximum gradient difference method to do text detection for non advertising. Then we locate the text and mark it in the original image and the decision on advertising or non-advertising will be achieved. This system uses shot cut detection to achieve the frequency of shots and use text detection to obtain trademark information of the advertising. The combination of the two methods achieves accurate detection of advertising. The video segment of the shot switch frequency or the trademark information of the advertising meeting the conditions of the video will be determined advertising otherwise, it will be non advertising. Finally the results of the detection of video advertising are output.

Two performance metrics are usually adopted for video advertising detection: precision and recall. Precision is used to measure the ability of eliminating independent video advertising of the detection system; recall is used to measure the ability to retrieve relevant video advertising of the detection system. The definitions of these two indicators are as follows:

$$\text{precision}(\%) = \frac{N_C}{N_C + N_F} \times 100\%$$

$$\text{recall}(\%) = \frac{N_C}{N_C + N_M} \times 100\%$$

In the formula, N_C (correct, detection) is the advertising frames correctly picked out, N_M (missed, detection) is the missing detection of advertising frames, N_F (false, alarm) is non advertising which is false detection as advertising.

5. **Experiment.** In order to verify the correctness of the proposed method for the detection of video advertising, 60 segment TV stations in China are selected as the test sample to test the method. The video advertising, which have more gradual change shots from the CCTV and Shandong satellite TV, mainly include the football match gap sports advertising and entertainment program gap of the life supplies advertising. The sample consists of 36 segment video advertising and 24 segment non advertising programs. Each program clip length is different, but all in the 300 frames and the frame rate is 20 frames every second. In the process of testing, the test sample is divided into two groups with a small group and a big one for advertising detection which is used to verify the detection accuracy of the algorithm is affected by the sample or not. In group 1, the test swatch is made up of the 12 segment video advertising and 8 segment non advertising programs, the total frame of advertising and non advertising in the group is 12856; In group 2, the test swatch is made up of the remaining 24 video advertising and 16 non advertising programs, the total frame of advertising and non advertising in the group is 21768.

In order to compare the video advertising detection method adopted in this paper with only adopting the particle swarm optimization and the neural network algorithm, the two kinds of test methods are all used in the experiment. The test results are shown in Table 1.

Table 1. Test results

	Advertising detection method	Video advertising(frame)	precision%	Recall %
Group1	PNN algorithm	4286	78.4	89.5
	Algorithm in this paper		81.7	91.8
Group2	PNN algorithm	10534	77.8	87.3
	Algorithm in this paper		80.9	90.4

In the first groups of test samples, when only using the particle swarm optimization neural network algorithm, the precision and recall of the video advertising detection are 78.4% and 89.5%, but the precision and recall of the video advertising detection which adopted the detection algorithm in this paper are 81.7% and 91.8%. In the second groups of test samples, when only using the particle swarm optimization neural network algorithm, the precision and recall of the video advertising detection are 77.8% and 87.3%, but the precision and recall of the video advertising detection which adopted the detection algorithm in this paper are 80.9% and 90.4%.

The test results show that, compared with detection neural network of particle swarm optimization algorithm of video advertising, using this algorithm in this paper, the precision and recall of advertising detection have been improved to a certain extent. But with the increase of the number of samples, using two kinds of algorithm for video advertisement detection, its precision and recall are reduced. After further experiments showed that the increase of gradient shots without obvious characteristics, the precision and recall of advertisement detection will be reduced accordingly. Therefore, the research on gradual shot detection without obvious feature will help to further improve the accuracy of detection of video advertising.

6. **Conclusions.** This paper proposes a method of video advertising detection based on the combination of particle swarm optimization neural network algorithm and text detection. On the basis of video shot cut detection, further detection of missing video advertising is detected by text detection, which improves the detection accuracy of video advertising.

REFERENCES

[1] L. Meng, Y. L. Li, G. Chen, TV Advertisement Retrieval Based on Shot Change and Text Detection, *Journal of Jilin University(Science Edition)*, vol. 50, no. 2, pp. 1185-1191, 2012.

[2] C. Y. Lee, A. Bhardwaj, W. Di, V. Jagadeesh, R. Piramuthu, Region-based Discriminative Feature Pooling for Scene Text Recognition, *[C]. In CVPR*, pp.4050-4057, 2014.

[3] X. J. Yuan, Research on video shot boundary detection algorithms, *[D]. Wuhan Polytechnic University master degree thesis*, 2010.

[4] B. Xu, J. Sun, J. Zhou, Detecting Method of Gradual Change Shot for Video Advertisements Detection, *[J]. Videc Engineering*, vol. 31, no. 12, pp. 67-70, 2007.

[5] X. Jiang, Research of video advertising text extraction algorithms, *[D]. University of Electronic Science and Technology of China master degree thesis*, 2015.

[6] C. Y. Lee, A. Bhardwaj, W. Di, V. Jagadeesh, R. Piramuthu, Region-based Discriminative Feature Pooling for Scene Text Recognition, *[C]. In CVPR*, pp.4050-4057, 2014.

[7] U. Gargi, R. Kasturi and S. H. Strayer, Performance characterization of video-shot-change detection methods, *IEEE Tram. Circuits Syst Video Technol*, vol. 10, no. 1, pp. 1-13, 2000.

[8] J. H. Yuan, H. Y. Wang, L. Xiao, et al, A Forma Study of Shot Boundary Detection, IEEE transactions on circuits and systems for video technology, vol.2, no.17, 2007.

[9] J. H. Yuan, H. Y. Wang, L. Xiaoet al, A Forma Study of Shot Boundary Detection, IEEE transactions on circuits and systems for video technology, vol.17, no.2, pp. 168-186, 2007.

[10] K. Edward Wong, M. Chen, A new robust algorithm for video text extraction, *JPattern Recognition*, vol.36, pp. 1397-1406, 2003

[11] L. N. Tan, Research on the image segmentation and video advertising detection based on graph partitioning model *[D]. University of Jinan master degree thesis*, 2009.

[12] X. D. Huang, K. H. Liu and L. S. Zhu, Auto Scene Text Detection Based on Edge and Color Features, *[C]. International Conference on Systems and Informatics* ,pp.1882-1886, 2012.

[13] G. K. Yildirim, R. Achanta and S. Susstrunk, Text Recognition in Natural Images using Multiclass Hough Forests, *[C]. International Conference on Computer Vision Theory and Applications (VISAPP)*, pp.737-741, 2013.