Smoke Recognition based on Dictionary and BP Neural Network

Yuanbin Wang, Yu Duan, Yuanyuan Li, Huaying Wu

Abstract—In order to improve the video-based smoke detection rate in a low illumination environment, a segmentation and recognition method of smoke images based on the combination of the dictionary and backpropagation (BP) neural network is proposed. Firstly, an image enhancement model is established for low-frequency and high-frequency images obtained by wavelet decomposition on the input image. The neural network predicts the coefficients of low-frequency and high-frequency images in the enhancement model to improve the image.e. Typical smoke images are collected to construct the smoke data matrix, and then background subtraction is employed to extract the motion region. The candidate smoke area is segmented further by principal component analysis and smoke data dictionary matrix. The feature vector is constructed using the statistics of the gray level co-occurrence matrix after the wavelet decomposition of each layer image. Finally, smoke is recognized by BP neural network. The experimental results show that the segmentation of the smoke image is relatively more complete, and the recognition rate is higher.

Index Terms—Wavelet transform, neural network, Image segmentation, smoke data dictionary matrix, smoke recognition

I. INTRODUCTION

FIRE poses severe risks to both people and the environment. Early smoke detection relies on different types of sensors, and the traditional methods used for detection have sensitivity and response time issues. However, the above issues can be avoided with video-based smoke detection, which is also less susceptible to environmental factors and spatial distances, including airflow, dust, and electromagnetic interference. Therefore, video-based smoke detection has been a research focus for academics both domestically and internationally.

In general, static and dynamic features serve as the foundation for video-based smoke detection. These smoke characteristics, including color, texture, wavelet transform features, movement features, diffusion features, etc., have all been extensively used by scholars to extract candidate smoke regions. To locate fire regions in video frames, Chen [1] integrated the information from color, space, and motion. The Gaussian mixture model was used to establish a fire color model for detecting candidate smoke regions. Spurious fire-like regions were removed according to combined spatial and temporal features. Finally, some missing fire regions were located by the region growing method. Gao [2] developed a new algorithm, which combines the ViBe algorithm with the MSER algorithm based on Bayesian theory to generate complete shapes of candidate smoke areas. It reduces the leak detection rate at full scale. Xu [3] extracted the dynamic target in the smoke video through background subtraction. Then, image features such as the color and texture of smoke are obtained to train the classifier.

For the texture feature extraction of smoke, gray level co-occurrence matrix (GLCM), wavelet, and other methods are widely used. Yuan [4] extracted the statistical features of smoke based on GLCM and described the local characteristics of smoke by LBP. Finally, smoke was detected by the Fisher classifier. Tong [5] employed discrete wavelet transform and discrete cosine transform to extract smoke features. And then they recognized smoke by the AdaBoost algorithm. Based on the static and dynamic features, Chen [6] proposed a cascade detection framework that has a good effect on complex scenes. For the color and diffusion trait of smoke, Santosa [7] extracted features by utilizing the gray co-occurrence matrix and selected the GLCM features by genetic algorithm. Wu [8] rearranged the continuous frames into a new large image according to the pixel-block rules and then utilized the dual tree-complex wavelet transform to decompose the new large image several times to extract smoke changes in different directions from these subbands.

Due to the uncertainty of fire location and smoke shape, forest fires are easily confused with clouds and trees. According to the human visual attention mechanism, smoldering smoke could be regarded as the turbulent and gray area in the smoke video. Based on this, Wang [9] adopted a smoke area extraction method based on saliency detection and the SURF-VIBE model. Sun [10] acquired the moving area by background subtraction and extracted the candidate smoke area by histogram background projection. Smoke-free images were eliminated through different statistical characteristics of smoke and smoke-free images. According to the smoke color, Chen [11] proposed a smoke judgment rule that combined color-based static and diffusion-based dynamic feature decision rules. In general, the significance test is related to the color, but the smoke color is not prominent enough, so there are still some limitations. Jia [12] conducted the significance test to detect candidate smoke. The rectangular sliding window with innovation was

Yuanbin Wang is an Associate Professor of Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an 710054, China (e-mail: 13379232752@163.com).

Yu Duan is a postgraduate student in Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an 710054, China (e-mail: 307073082@qq.com).

Yuanyuan Li is a postgraduate student in Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an 710054, China (e-mail: 2764845392@qq.com).

Huaying Wu is a postgraduate student in Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an 710054, China (e-mail: 1661349504@qq.com).

employed to compare the feature distribution and calculate the significance of smoke detection.

Local binary patterns (LBP), pyramid local binary patterns, and GLCM are widely applied to obtain texture features in smoke recognition. In [13], the LBP operator was extracted by the new smoke texture, which is obtained from the missing code of LBP. In [14], the smoke detection algorithm is presented based on the dynamic update of the background and dark channel prior. Favorskaya [15] extracted the luminance difference mapping, luminance gradient mapping, and Laplace mapping. These mappings are encoded by the spatial-temporal local binary pattern as texture features. However, although it fused many features, it disregards other factors like illumination, weather, and others. Yuan [16] proposed a method based on the Gabor kernel and LBP. It can more clearly express smoke characteristics and is appropriate for smoke identification. Emmy [17] determined the smoke region according to wavelet frequency and dynamic texture. And GLCM is introduced to capture the relation of pixels. Liu [18] utilized the k-means to remove the non-smoking color interference pixels to obtain candidate smoke regions. Confidences were evaluated after extracting area growth characteristics, such as spatial energy, LBP histogram, and histogram of oriented gradient (HOG) features. The calculated feature confidences were input into the dynamic score combination to determine whether the region contained smoke.

In the approaches mentioned above, static and dynamic features of smoke were extracted, and smoke was detected by the threshold setting. However, there are some difficulties, including the selection of smoke characteristics, the collection of smoke samples, and the manual setting of thresholds. Therefore, this kind of method is limited in general applicability.

In all, the video-based smoke detection system has taken great strides, but there are still some disadvantages with practical application. How to raise the accuracy rate for smoke detection while lowering the false alarm rate and missing alarm rate is still significant.

The smoke segmentation and recognition method that combines the smoke dictionary and BP neural network is proposed to aim at the issues of incomplete segmentation and low fire recognition rate. In this paper, smoke segmentation and recognition are conducted on the data dictionary based on 1500 typical smoke images. The contributions and improvements of this article are as follows:

1) An image enhancement model based on wavelet decomposition and coefficient prediction is proposed for image enhancement.

2) A smoke region segmentation method based on a smoke dictionary matrix is presented to catch candidate smoke areas.

3) We proposed a smoke recognition method based on wavelet transform and GLCM for smoke recognition.

II. FLOW CHART OF THE PROPOSED METHOD

The flow chart of the proposed method is shown in Fig. 1, and the specific process is as follows:

1) Firstly, the original image is used as the input of the image enhancement model, which is established based on the low-frequency and high-frequency images decomposed by wavelet transform. The parameters in this enhancement model are predicted by the neural network.

2) Compare the enhanced image with the smoke dictionary matrix to obtain the candidate smoke area.

3) Finally, wavelet decomposition is utilized to extract the features of the candidate smoke area. Moreover, these features are input into the neural network for smoke recognition.
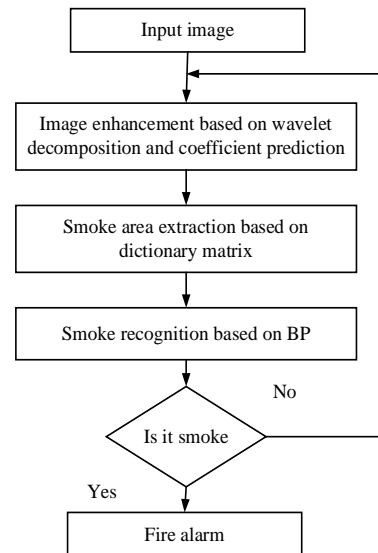

Fig. 1. Flow chart of smoke segmentation and recognition.

III. IMAGE ENHANCEMENT PREPROCESSING

A. Modeling for Image Enhancement

There are a few popular algorithms for improving image quality in the spatial domain, including histogram equalization, exponential transform, and others. However, these algorithms perform non-selective processing on pixels when enhancing images, which resulting in problems such as the reduction of effective signal contrast and lack of clarity in target details.

In the unsharp masking, the input image is subtracted from the fuzzy image obtained by low-pass filtering to obtain the image with high-frequency components retained. Then an amplification coefficient is added to the high-frequency image and superimposed on the input image to enhance the image. The application is simple due to the introduction of fewer parameters, and can accurately capture image edge details. The principle is shown as formula (1).

$$g(x, y) = k\left[f(x, y) - f'(x, y)\right] + f(x, y) \qquad (1)$$

Where $f(x, y)$ is the original image, and $f'(x, y)$ is a passivated fuzzy image after low-pass filtering, k is the amplification coefficient.

When unsharp masking is performed, the low-frequency portion of the image remains unchanged, which only enhances the image edge. Therefore, a new image enhancement model is proposed based on unsharp masking and wavelet decomposition. The image is enhanced from detail and edge, as shown in formula (2).

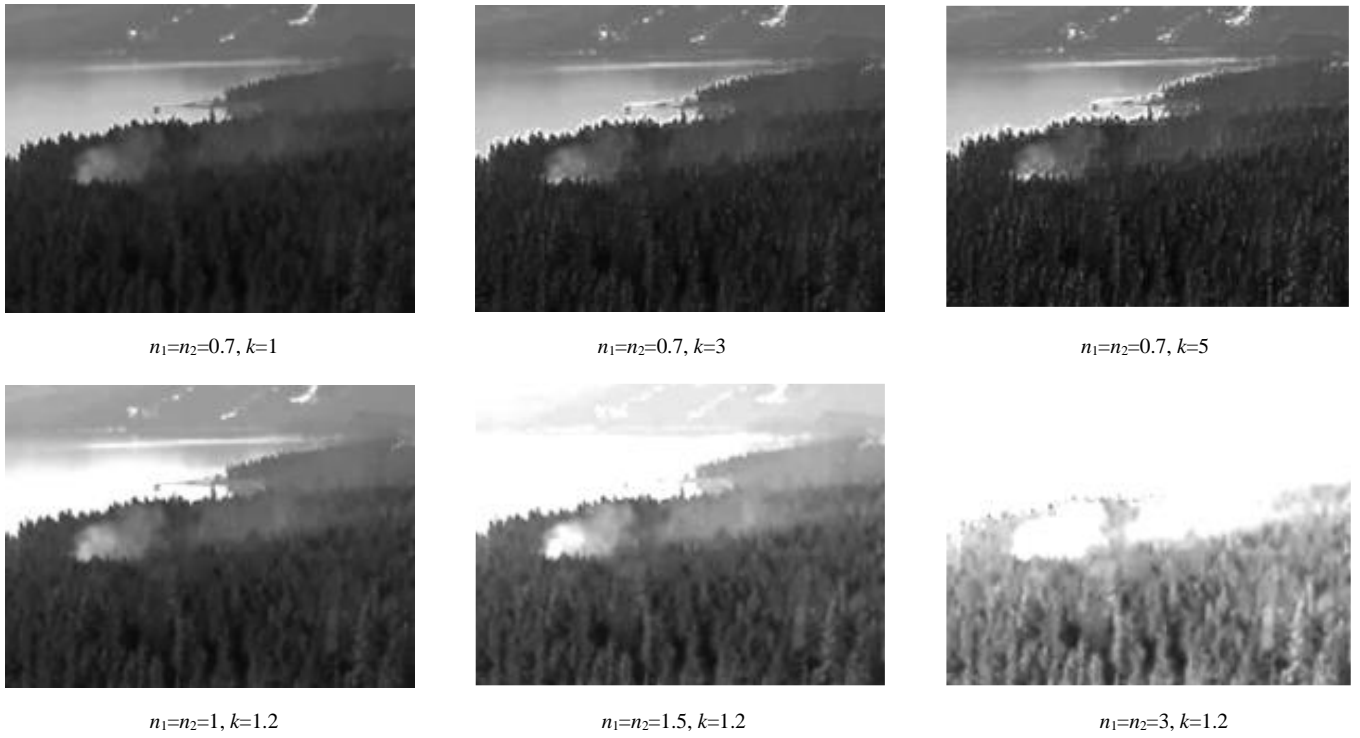$$g(x, y) = k\sum H + n_1 L_1 + n_2 L_2 \qquad (2)$$

$n_1=n_2=0.7, k=1$ $n_1=n_2=0.7, k=3$ $n_1=n_2=0.7, k=5$

$n_1=n_2=1, k=1.2$ $n_1=n_2=1.5, k=1.2$ $n_1=n_2=3, k=1.2$

Fig. 2. Effects of n_1, n_2, and k on the image.

Where $g(x, y)$ is the enhanced image, L_1 and L_2 denote low-frequency images of the first and second layers after wavelet decomposition, respectively. k and n_i are the coefficients of high-frequency and low-frequency images, respectively. $\sum H$ is the sum of all high-frequency images after the decomposition of two layers by wavelet.

In this model, $\sum H$ denotes the edge information, and the edge will be more prominent after linear multiplication. L_1 and L_2 represent low-frequency information. The details will be more distinct after linear multiplication. The influence of coefficients k, n_1, and n_2 on the image is shown in Fig. 2. The edge information of the image is affected by the value of k when n_1 and n_2 remain unchanged. The larger the value of k, the more prominent the edge will be enhanced. The image edge will be over-enhanced if the value of k is too large. The brightness is so intense that details are lost when n_1 and n_2 are both greater than 1. Through experimental analysis, the value range of n_1 and n_2 are both set from 0.1 to 1, and k is set from 1 to 5.

To summarize, enhancement results of edge and detail information are related to the settings of k and n_i, so how to adaptively select coefficients k and n_i plays an important role in image enhancement.

B. Coefficient Prediction Based on BP Neural Network

In the reason of the nonlinear relation between coefficient n_i and k, coefficients in the image enhancement model can be predicted by the BP neural network, which made up of several nonlinear structural elements. The mean value, variance, and entropy of the image are selected as the input of the BP neural network, and the expressions are shown in formulas (3), (4), and (5).

$$\bar{f}(x, y) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} f(x_i, y_j) \tag{3}$$

$$D(x, y) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} [f(x_i, y_j) - \bar{f}(x, y)]^2 \tag{4}$$

$$H = \sum_{i=0}^{255} p_i \log p_i \tag{5}$$

Where P_i denotes the proportion of pixels with the gray value i in the image, and $f(x, y)$ is the gray value in the image (x, y).

The coefficients in the model need to be adjusted to optimize the enhanced image quality, and the enhancement degree is judged by the energy gradient function in the formula (6). The smaller the value of the function is, the blurrier the image is. Vice versa.

$$D(f) = \sum_{y} \sum_{x} \left(\left| f(x+1, y) - f(x, y) \right|^2 + \left| f(x, y+1) - f(x, y) \right|^2 \right) \tag{6}$$

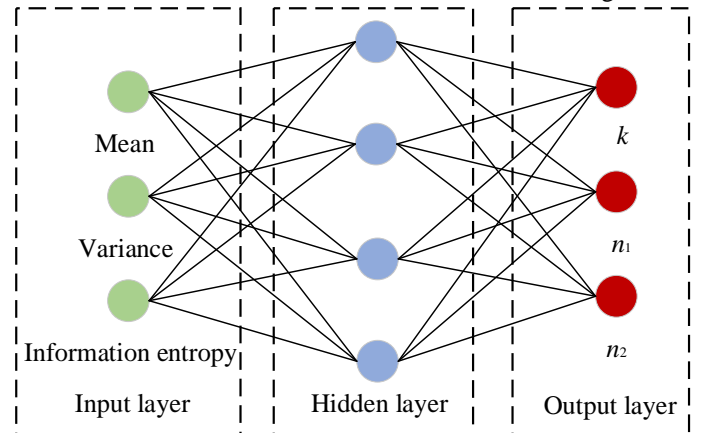The structure of the BP neural network is shown in Fig. 3.

Fig. 3. The structure of the forecast network.

1) **The input of the BP neural network:** 100 frames of images are input to calculate values of the mean, variance, and information entropy, respectively, as the input of the BP neural network.

2) **Network training:** There are three components in the network, including the input layer with three neurons, the

hidden layer with four neurons, and the output layer with three neurons. And the mean square error is selected as the loss function, and the number of training times is set to 500.

3) **The output of the BP neural network:** When the energy gradient function of the images reaches the maximum value, the coefficients of each image frame are recorded as the output of the BP neural network.

The change curve of the loss function during the neural network training is shown in Fig. 4.
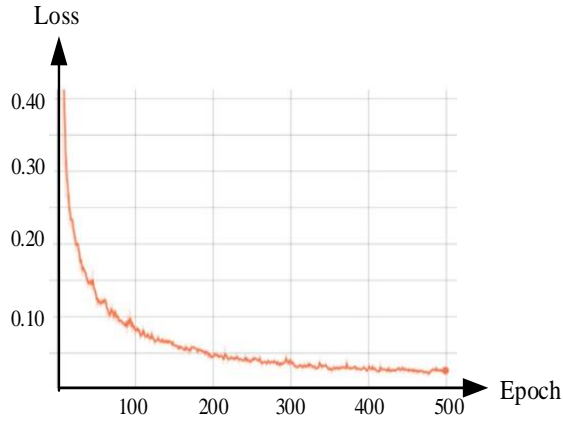

Fig. 4. The curve of the loss function.

Fig. 4 shows the error value of prediction in the ordinate and the iteration times in the abscissa. The final prediction error order is "10-2" as the iteration times increase. Since coefficients k, n_1, and n_2 are all greater than "10-1", the error value has little impact on the enhancement result. The real coefficients and predicted coefficients of the two graphs in Fig. 5(e) are shown in Table I. It can be seen that the predicted results are close to the actual results, and the model can improve work efficiency effectively.

TABLE I
COMPARISON OF REAL COEFFICIENTS AND PREDICTIVE COEFFICIENTS

Figure	Predicted value	Real value
	k	1.31
Figure 5(e) left	n_1	0.54
	n_2	0.86
	1.30	1.32
Figure 5(e) right	0.50	0.64
	0.90	0.68

To test the effectiveness of the proposed algorithm, experiments of the proposed and other classical algorithms are performed in Fig. 5. Histogram equalization enhances the brightness of the overall image effectively, but over-enhancement and under-enhancement appear locally. The image after linear transformation is bright as a whole, but the details are blurred. Although some features are amplified by the logarithmic modification, the overall image is very bright.

At the same time, the above comparison methods have the problems of low efficiency of artificial parameter selection and poor image enhancement effect. However, the proposed algorithm adaptively determines the parameters through the BP neural network, which can achieve better image enhancement effects.

In addition, the enhanced image contrast of the proposed algorithm is better than that of other algorithms, which retains more image details and avoids the situation of excessive enhancement, making the smoke target more prominent and convenient for subsequent segmentation.

(a) original image

(b) Histogram equalization

(c) Linear transformation

(d) Logarithmic transformation

(e) The proposed algorithm

Fig. 5. Enhancement results based on different algorithms.

IV. CANDIDATE SMOKE AREA EXTRACTION BASED ON SMOKE DICTIONARY MATRIX

A. Candidate Smoke Area Extraction

After the enhancement of the input image, the background model is built to obtain the segmented image of the motion region by comparing it with the enhanced image. The background model in this paper is constructed using the mean method, as stated in formula (7):

$$B(i, j) = \frac{1}{N} \sum_{M=0}^{N} f_M(i, j) \qquad (7)$$

Where $f_M(i, j)$ is the pixel value (i, j) of the frame M in the image, and $B(i, j)$ is the background model.

We select the scene with a slight change and establish the background model based on 100 frames of the static image.

The moving region can be obtained by making a background difference between the enhanced image and the background model. If the distance between the enhanced image and the pixel corresponding to the background image is more significant than than a certain threshold, it can be regarded as foreground. The output is "1". Otherwise, it is considered to be the background, and the output is "0". According to the actual situation, the threshold is adjusted several times so that the binary image obtained can extract the smoke target fully. In this paper, the threshold is selected as 15. The extraction results of the motion area as shown in Fig. 6.

(a) Background model

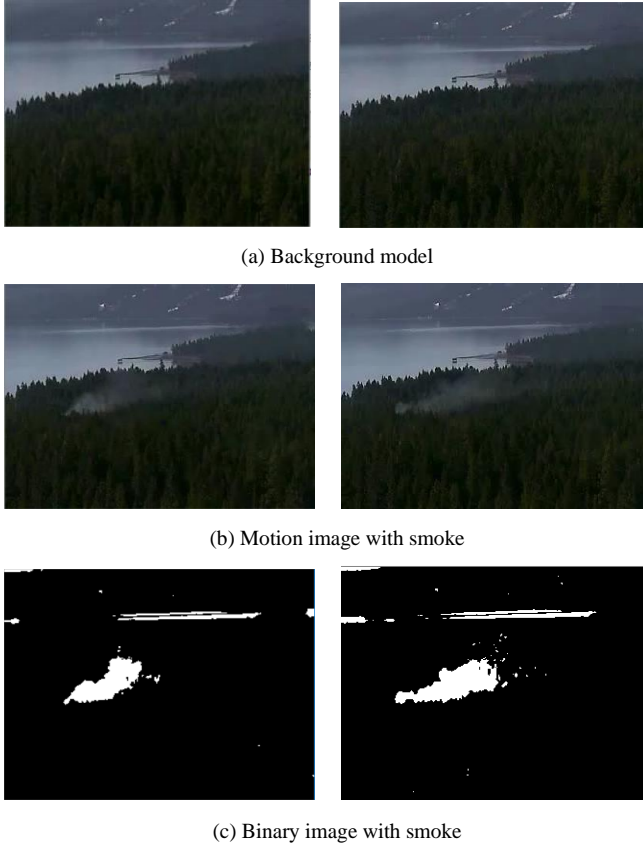(b) Motion image with smoke

(c) Binary image with smoke

Fig. 6. Binary image of motion area.

It can be seen from Figure 6 that there are still some non-smoking areas in the binary image of motion area obtained by subtracting background, as shown in blocks B1 and B3 of Fig. 7. We need further optimization to extract exact smoke regions. In [11], Chen found in statistics that the range of gray value of smoke in the RGB channel was [80, 220]. Therefore, we can judge whether blocks B_1 B_2 and B_3 conform to the smoke characteristics according to this rule, and remove the non-conforming areas.
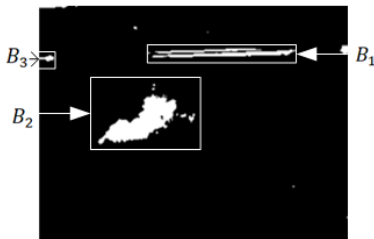

Fig.7. Binary image segmentation of motion regions with smoke.

Due to the uncertain shape of smoke, such as airflow, it is not easy to segment the image accurately. In this paper, the typical smoke image is collected and made into the smoke dictionary, and then the input image is compared with the smoke dictionary to determine whether there is smoke in the input image.

The smoke segmentation based on the smoke dictionary is mainly divided into two stages. The first stage is the representation of the smoke dictionary. Without losing information, Principal Component Analysis (PCA) is used to extract the principal components of the collected typical smoke images, and the smoke dictionary matrix Ω is determined by transforming them into the same space. In the second stage, the input image is compared with the smoke dictionary matrix to obtain the smoke target.

The smoke dictionary matrix Ω is determined based on the collected 1500 typical smoke images as follows:

1) Create a sample matrix X based on the training samples, where x_i is the typical smoke image collected.

$$X = (x_1, x_2, ..., x_{1500})^{\mathrm{T}} \tag{8}$$

2) Calculate the average value of each line, that is, the mean value of pixels on each frame:

$$\varphi = \frac{1}{1500} \sum_{i=1}^{1500} x_i \tag{9}$$

3) Calculate the difference between the collected smoke image of each frame and the average value of each line:

$$d_i = x_i - \varphi \tag{10}$$

4) Construct a covariance matrix:

$$C = \frac{1}{1500} \sum_{i=1}^{1500} d_i d_i^{\mathrm{T}} \tag{11}$$

5) Find the eigenvalues and eigenvectors of C. As in [16], in order to reduce data loss of training samples, the top p eigenvalues and their corresponding eigenvectors w are selected to ensure that the projections on the top p eigenvector sets have 90% energy [19].

6) According to formula (12), the difference vector d_i is projected into the smoke space and constructed as a smoke dictionary matrix Ω = [Ω₁, Ω₂, …, Ω₁₅₀₀].

$$\Omega_i = w^{\mathrm{T}} d_i \tag{12}$$

The input image is compared with the smoke data dictionary. Each image block B_j of the moving region in the binary image is transformed into the smoke space with the corresponding part of the enhanced image to obtain the resulting image y_j. Then, the minimum value from y_j to Ω_i is calculated as shown in formula (13), and the image block with the minimum value is chosen as the candidate smoke image.

$$\min \left\{ \left\| y_j - \Omega_i \right\|_2^2 \right\} \tag{13}$$

Where i is the collected typical smoke images, and j is the number of image blocks.

The proposed algorithm and the classical foreground detection algorithm, Gaussian Mixture Model (GMM), are used to extract the candidate smoke regions. The extraction results are shown in FIG. 8. Since GMM does not utilize the context information of smoke image frames, the segmented smoke regions are incomplete, and most of the targets are not extracted. The algorithm compares the input image with the smoke dictionary matrix. It uses the context information combined with smoke gray value features to complete the segmentation of smoke features, eliminating redundant background interference.

(a) Binary image of the motion region with smoke

(b) Segmentation by GMM

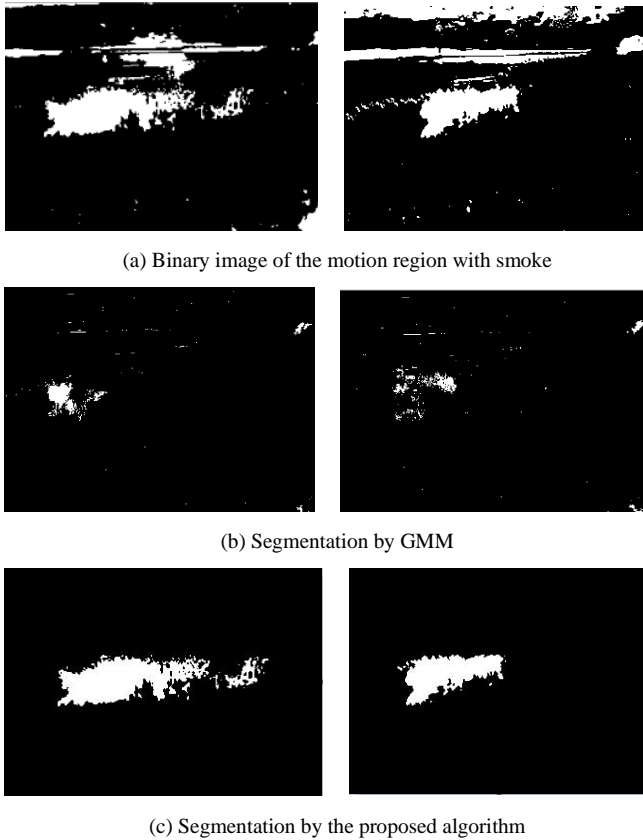(c) Segmentation by the proposed algorithm

Fig.8. Smoke segmentation results.

B. Segmentation Evaluation

In this section, the segmentation results are evaluated from the aspect of accuracy and effectiveness.

In [20], Matthews correlation coefficient (*MCC*), F1 parameters, and *GT* (ground truth) image are adopted to evaluate the algorithm. *MCC* is a balanced evaluation index which often selected for machine learning as an evaluation index for binary classification results. It not only considers accuracy but also essentially describes reality. The range of the correlation coefficient between actual classification and prediction classification is [-1, 1]. The value of "1" represents a perfect prediction, and "0" is considered that the prediction results are worse than random prediction results. "-1" means that the predicted classification is entirely inconsistent with the actual classification. *MCC* is calculated as follows:

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TN + FN)(TN + FP)(TP + FN)(TP + FP)}} \tag{14}$$

In formula (14), the meaning of the parameters is shown in Table II.

TABLE II
MEANING OF PARAMETERS IN THE MCC INDICATOR

Parameter	Target classification	Actual classification
TP	Smoke pixel	Smoke pixel
FN	Smoke pixel	Non-smoke pixel
FP	Non-smoke pixel	Smoke pixel
TN	Non-smoke pixel	Non-smoke pixel

F_1 is a categorical indicator that considers both the precision and the recall rate. It refers to the harmonic mean of the precision and recall, that is:

$$F_1 = 2 \times \frac{\text{Precision} + \text{Recall}}{\text{Precision} \times \text{Recall}} \tag{15}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{16}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{17}$$

The segmentation based on the *GT* image is calculated by comparing the number of actual smoke pixel points with the number of smoke pixel points to obtain segmentation precision, over-segmentation rate, and under-segmentation rate. The calculation formula is shown in Table III.

In this paper, the *MCC*, F_1, and *GT* images, including three evaluation indexes, are used to evaluate the proposed algorithm, as shown in Table IV.

From Table IV, the proposed algorithm performs better on *MCC*, F_1, and *GT*-based images than the Gaussian mixture model. The subsequent smoke recognition, which can be further researched, will not be impacted by the slight discrepancy between smoke regions segmented by the proposed algorithm and all smoke areas.

V. SMOKE RECOGNITION

In practical application, there are many things similar to smoke, such as clouds, water mist, and so on. Therefore, further extracting image features from the segmented smoke region is necessary for determining whether it is smoke. In this paper, we propose a smoke recognition method based on wavelet transform and GLCM for smoke recognition.

1) Firstly, the candidate smoke regions are decomposed by the wavelet.

2) Then, the statistic of the GLCM from each layer image is calculated to construct the feature vectors.

3) Finally, the constructed feature vectors are input into the BP neural network for smoke recognition.

TABLE III
SEGMENTATION ACCURACY FORMULAS

Evaluation index	Formula	Meaning of each parameter
Segmentation accuracy (*SA*)	$SA = (1 - \frac{\lvert R_s - T_s \rvert}{R_s}) \times 100\%$	R_s is the actual number of smoke pixels T_s is the number of pixels in the segmented image.
Over-segmentation rate (*OR*)	$OR = \frac{O_s}{R_s + O_s}$	O_s is the number of smoke pixels that should not appear in the segmentation result but appear.
Under-segmentation rate (*UR*)	$UR = \frac{U_s}{R_s + O_s}$	U_s is the number of pixels that should appear in the segmentation result but do not appear.

TABLE IV
SEGMENTATION ACCURACY FORMULAS

Method	MCC	F1	SA	OR	UR
The proposed algorithm	0.893	0.912	89.3%	1.2%	5.7%
GMM	0.868	0.846	76.8%	0.8%	7.6%

After the segmentation, the candidate smoke region is gradually approximated by the wavelet function with scale 2, and the result is shown in Fig. 9.

*HL*₁

*LH*₁   *HH*₁

(a) Smoke sample 1
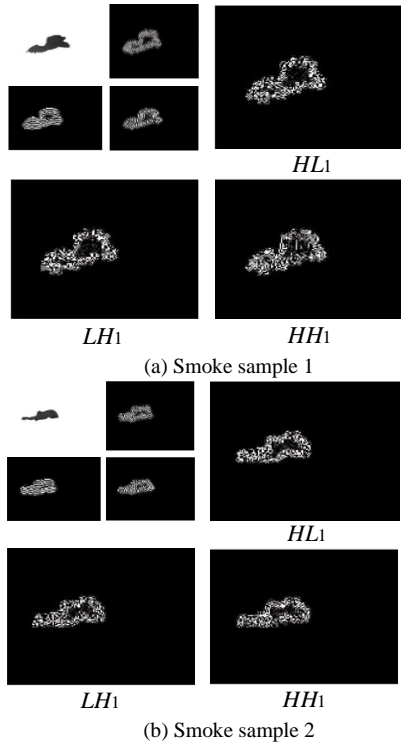
*HL*₁

*LH*₁   *HH*₁

(b) Smoke sample 2

Fig. 9. Levels of the wavelet decomposition.

In Fig. 9, the image shows different characteristics in different layers after wavelet decomposition, including high-frequency subband coefficients HL_1, LH_1, and HH_1 in the first layer. The upper-left corner of the figure is the second layer of wavelet decomposition, and the position of each high-frequency subband coefficient is consistent with the first layer.

After wavelet decomposition, the statistics features can be calculated, such as gray distribution unevenness, gray mean square error, gradient entropy, mixed entropy, inertia, deficit moment, correlation, and variance. The texture feature vector is set as P= [P_1, $P_{i,1}$, $P_{i,2}$, $P_{i,3}$]. Here P_1 is the gradation of the GLCM of the original picture. In $P_{i,n}$, i is the i-th layer image obtained by wavelet decomposition. The value of n can be taken as 1, 2, and 3, which are the gray co-occurrence matrix statistics of HH, LH, and HL in the high-frequency subgraph of layer i, respectively.

The validation data includes 90 frames of smoke, five frames of water mist, and five frames of cloud image. The texture feature vector P is the input vector of the BP neural network, and the predicted result is used as the output. BP neural network contains a hidden layer composed of 7 hidden layer neurons. The ReLU function is selected as the activation function of the BP neural network, and the L_1 normalized cross-entropy loss is applied as the loss function. The optimization algorithm adopts stochastic gradient descent (SGD). We set the maximum number of iterations to 2000.

The accuracy curve and loss function curve are shown in Fig. 10 and Fig. 11, respectively.
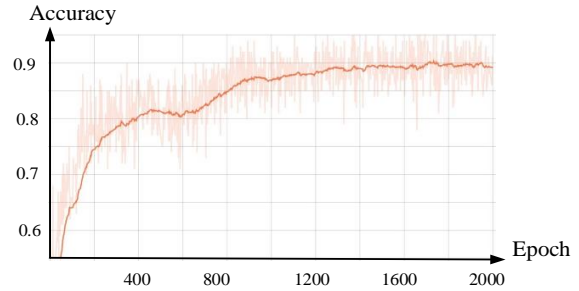

Fig. 10. The variation of the accuracy curve.

In Fig. 10, abscissa and ordinate denote the iteration times and recognition accuracy during the verification process, respectively. It can be seen that at the beginning, the learning speed of the model is relatively fast, and the recognition accuracy increases rapidly. However, as the learning ability of the model is gradually saturated, the ascent rate slows down, and the final recognition accuracy reaches 90.1%.
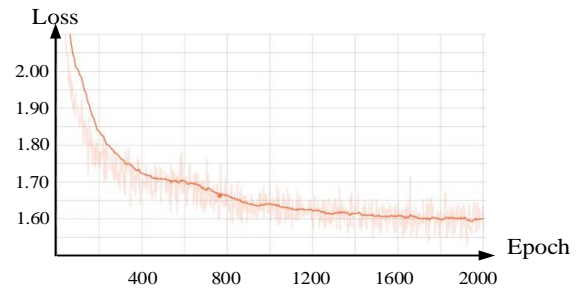

Fig. 11. The variation of the loss function value.

In Fig. 11, the ordinate represents the change in the loss function value that corresponds to the increase of verification iterations on the abscissa. The loss value decreases rapidly at first and then decreases slowly as the model becomes saturated until the loss function value drops to 1.60.

As the iteration time increases, the parameters in the neural network model gradually adapt to smoke recognition. When the recognition rate and loss function curves no longer change, the neural network model reaches the limit of smoke recognition, and the recognition accuracy is at the highest level, reaching 90.1%.

The test set contains 100 frames of smoke, water mist, and cloud images in a ratio of 8: 1: 1. The best model retained in the above verification process is adopted for testing. In order to verify the superiority of this method, we compare it with the other four methods and select the recognition accuracy and false alarm rate as objective evaluation indexes in Table V. It can be seen from Table V that the proposed method has better performance, with a recognition accuracy of 90 % and a false alarm rate of 4.3 %.

TABLE V
COMPARISON OF TEST RESULTS

Methods	Accuracy	False alarm rate
Based on the literature [21]	78.3%	8.4%
Based on the literature [22]	83.6%	5.7%
Based on the literature [23]	88.5%	6.7%
Based on the literature [24]	86.7%	5.1%
The proposed method	90.0%	4.3%

VI. CONCLUSION

In this paper, aiming to improve the image contrast in the low illumination environment, wavelet decomposition is utilized to enhance the high-frequency and low-frequency of the image, respectively, which makes the smoke target display more evident. In view of the different characteristics between smoke and background texture, the smoke target region is segmented based on the motion block and smoke data dictionary matrix. At the same time, the method based on wavelet decomposition and BP neural network is intended to increase recognition accuracy. Furthermore, texture vectors are extracted according to the GLCM, which reduces the computational complexity. Experimental results show that the smoke recognition accuracy of the proposed method reaches 90.0 %. Compared with other comparison algorithms, it has the lowest false alarm rate of 4.3 %, which can effectively realize smoke detection so as to take corresponding treatment measures in time.

Although the accuracy of this paper is significantly better than other comparison algorithms, it is still limited by the performance of the model. In addition, the selection of the smoking threshold is based on the actual situation of many experiments. Therefore, subsequent work can continue around the improvement of neural network models and the selection of threshold.

REFERENCES

[1] L. H. Chen and W. C. Huang, "Fire detection using spatial temporal analysis," *Lecture Notes in Engineering and Computer Science*, vol. 2206, no. 1, pp. 2222–2225, 2013.

[2] Y. Gao and P. Cheng, "Full-Scale Video-Based Detection of Smoke from Forest Fires Combining ViBe and MSER Algorithms," *Fire Technology*, vol. 57, no. 6, 2021.

[3] X. B. Xu, G. X. Tang, J. Y. Wu, and C. Geng. "Research on Fire Smoke Detection Method Based on Video Image," *In 2021 the 5th International Conference on Graphics and Signal Processing (ICGSP 2021)*, June 25–27, 2021, Nagoya, Japan.

[4] J. Yuan, W. Yuan, Y. Jia, T. Fang, J. J. Wang, and Y. M. Zhang, "An Active Infrared Smoke Recognition Method Based on Texture Features," *Journal of Safety and Environment*, vol. 16, no. 02, pp. 86-89, 2016.

[5] B. B. Tong and S. T. Wang, "Video smoke detection using two-layer classification algorithm," *Journal of Computer Science*, vol. 42, no. 03, pp. 301-306, 2015.

[6] T. Z. Chen, Z. J. Wang, H. X. Chen, and L. Y. Zhao, "Video dynamic smoke detection based on concatenated convolutional neural network," *Journal of University of Electronic Science and Technology of China*, vol. 45, no. 06, pp. 992-996, 2016.

[7] S. Santosa, R. A. Pramunendar, D. P. Prabowo, and Y. P. Santosa, "Wood types classification using back-propagation neural network based on genetic algorithm with gray level co-occurrence matrix for features extraction," *IAENG International Journal of Computer Science*, vol. 46, no.2, pp. 149-155, 2019.

[8] X. H. Wu, Y. C. Cao, X. B. Lu, and H. Leung. "Patchwise dictionary learning for video forest fire smoke detection in wavelet domain," *Neural Computing and Applications*, vol. 33, no. 13, pp. 7965-7977, 2021.

[9] X. Wang, K. Z. Wu, Z. R. YU, and H. Zheng, "Smoke Extraction Algorithm for Forest Fire in Video Based on PooLNet Saliency and SURF-VIBE Model," *Journal of Nanchang Hangkong University (Natural Science)*, vol. 34, no. 2, pp. 94-100, 2020.

[10] J. K. Sun and R. Y. Yang, "Smoke detection in video based on color histogram and wavelets," *Computer Science*, vol. 41, no. 12, pp. 251-254+287, 2014.

[11] T. H. Chen, Y. H. Yin, and S. F. Huang, "The smoke detection for early fire-alarming system base on video processing," *Proceedings of International Conference on Intelligent Information Hidingand Multimedia Signal Processing, Pasadena, USA: IEEE Computer Society*, pp. 427-430, 2006.

[12] Y. Jia, J. Yuan, and J. J. Wang, "A saliency-based method for early smoke detection in video sequences," *Fire Technology*, vol. 52, no. 5, pp. 1271-1292, 2016.

[13] F. N. Yuan, J. Shi, X. Xia, and Y. Yang. "Sub oriented histograms of local binary patterns for smoke detection and texture classification," *KSII Transactions on Internet and Information Systems*, vol. 10, no. 4, pp. 1807-1823, 2016.

[14] L. Zhao, Y. M. Luo, and X. Y. Luo, "Based on dynamic background update and dark channel prior of smoke detection algorithm," *Application Research of Computers*, vol. 34, no. 3, pp. 957-960, 2017.

[15] M. Favorskaya, A. Pyataeva, and A. Popov, "Verification of smoke detection in video sequences based on spatio-temporal local binary patterns," *Procedia Computer Science*, vol. 60, pp. 671-680, 2015.

[16] F. N. Yuan, G. Li, X. Xia, L. Zhang, and Y. Zhou, "Smoke recognition by combining aggregated Gabor kernels and local binary patterns," *Journal of Chinese Computer Systems*, vol. 40, no. 4, pp. 828-833, 2019.

[17] C. Emmy Prema, S. S. Vinsley, and S. Suresh, "Multi Feature Analysis of Smoke in YUV Color Space for Early Forest Fire Detection," *Fire Technology*, vol. 52, no. 5, pp. 1319-1342, 2016.

[18] H. Liu, S. Fang, L. Zhe, and X. Wei, "The field of smoke detection based on background feedback and k-means," *China Science Paper*, vol. 14, no. 03, pp. 261-267, 2019.

[19] D. H. Liu, Y. S. Cui, Y. S. Zhao, Y. T. Song, and J. H. Wang, "A teaching image retrieval method based on improved convolutional neural network," *Journal of Xidian University*, vol. 4, no. 11, pp. 1-7, 2019.

[20] T. Toulouse, L. Rossi, and M. Akhloufi, "Benchmarking of wildland fire colour segmentation algorithms," *Image Processing Iet*, vol. 9, no. 12, pp.1064-1072, 2015.

[21] Z. Wang, P. Liu, and T. Cui, "Research on Forest Flame Recognition Algorithm Based on Image Feature," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 925-928, 2017.

[22] N. Zhao, X. W. Wang, and S. L. Yin, "Research of fire smoke detection algorithm based on video," *International Journal of Electronics and Information Engineering*, vol. 13, no. 1, pp. 1-9, 2021.

[23] M. J. Sousa, A. Moutinho, M. Almeida, "Wildfire detection using transfer learning on augmented datasets," *Expert Systems with Applications*, vol.142, pp.1-13, 2020.

[24] F. M. Bi, X. Y. Fu and C. Wei, "Fire detection method based on improved fruit fly optimization-based SVM," *CMC: Computers, Materials & Continua*, vol. 62, no. 1, pp. 199-216, 2020.