

# Smoke Image Recognition Method Based on the optimization of SVM parameters with Improved Fruit Fly Algorithm

Jingwen Liu<sup>1</sup>, Junshan Tan<sup>1\*</sup>, Jiaohua Qin<sup>1</sup>, Xuyu Xiang<sup>1</sup>

<sup>1</sup> Central South University of Forestry and Technology  
Hunan Changsha, 410004 - CHINA

[e-mail: ljw844532422@foxmail.com, 837380161@qq.com.cn, qinjiaohua@163.com, xyuxiang@163.com]

\*Corresponding author: Junshan Tan

*Received March 28, 2020; revised July 11, 2020; revised July 18, 2020; accepted July 21, 2020;  
published August 31, 2020*

---

## Abstract

The traditional method of smoke image recognition has low accuracy. For this reason, we proposed an algorithm based on the good group of IMFOA which is GMFOA to optimize the parameters of SVM. Firstly, we divide the motion region by combining the three-frame difference algorithm and the ViBe algorithm. Then, we divide it into several parts and extract the histogram of oriented gradient and volume local binary patterns of each part. Finally, we use the GMFOA to optimize the parameters of SVM and multiple kernel learning algorithms to Classify smoke images. The experimental results show that the classification ability of our method is better than other methods, and it can better adapt to the complex environmental conditions.

---

**Keywords:** smoke image; fruit fly optimization algorithm; Multiple Kernel Learning

## 1. Introduction

**F**ire is a great threat to people's property safety and life safety. At present, various types of fire detectors on the market mainly include thermal sensing detector, light sensing detector, smoke sensing detector. These traditional detectors are cheap and accurate, but some deficiencies are difficult to solve. For space above a certain height or outdoor places where fires frequently occur, such detectors are no longer suitable. Smoke sensing detectors are usually used in some office buildings or factories. As time goes by, some dust and other particles will enter the air to cause corrosion to those detectors. At this time when a real fire occurs, those detectors can't alarm at all. To prevent this from happening, those detectors need to be regularly checked and repaired which will also bring extra manpower and material consumption. Therefore, most researchers attempt to combine some coding [1][2][3], classification [4], recognition [5], feature fusion model [6][7], and some other image methods [8][9][10] to design efficient smoke image detection methods.

Russo uses the approximate median filtering algorithm to subtracts the background region from the input frame and the shape-based filtering algorithm to find the motion region [11]. This method has a high smoke image recognition rate. Murat TOPTAŞ uses the YUV color space to detect smoke areas. After that, the gray level co-occurrence matrix (GLCM) is used to extract the features of the smoke image and the support vector machine (SVM) method is used to classify it [12]. However, the YUV color space of the smoke image is not obvious, so the recognition ability of this method is not good. Mengtao Huang takes the main direction of smoke movement as the dynamic features of indoor low-light smoke images and selects the texture features of the smoke images as the static features of indoor low-light smoke images to recognize the smoke images. It has good anti-interference ability and has good performance for smoke image recognition under low-light indoor conditions [13]. Lin Wang extracts the color features of the suspicious region from the RGB color space and the HSI color space [14] of the early smoke images. Then Lin Wang uses a two-dimensional discrete wavelet transform to extract background blur features and calculates a ratio of the number of pixels in the suspected smoke region to the number of pixels in the corresponding minimum enclosing rectangle to extract the contour irregularity features. Lin Wang uses the optical flow method to extract the features of the main direction of the smoke movement. This method can effectively improve the smoke image recognition rate, but the use of the optical flow method for the main direction of the smoke movement will result in low algorithm efficiency and cannot achieve real-time monitoring on smoke images detecting [15]. At present, the way of how to extract smoke images features, and the way of how to select SVM parameters based on the data model becomes the key issue. This paper intends to use the good group of improved fruit fly optimization algorithm to find the optimal kernel parameters and penalty terms of SVM, and combine the smoke images features to train SVM to recognize the smoke images under complex conditions.

## 2. Related Work

### 2.1 Three-frame difference algorithm

The three-frame difference algorithm is an extension of the inter-frame difference algorithm. It is an algorithm for obtaining motion regions by performing two-two different operations

on successive three frames of video image sequences. When abnormal target motion occurs in the monitoring camera, there will be a more obvious difference between the adjacent two frames. Then we subtract the two frames to get the absolute value of the difference in pixel values at the corresponding position and determine whether it is greater than a certain threshold, to extract the motion region. Fig. 1 shows the steps.

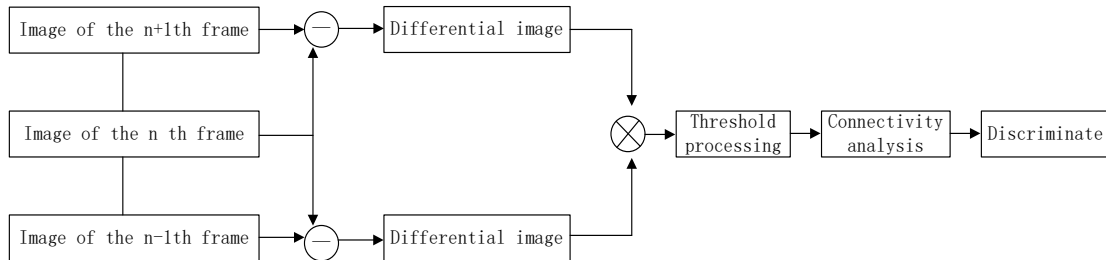


Fig. 1. Three-frame difference algorithm

## 2.2 ViBe algorithm

The ViBe algorithm is a motion region recognition algorithm based on background updates. The principle is to create a set of  $K$  sample pixel values for each pixel in the current area. If the number of the Euclidean distance between the new pixel and all  $K$  pixel values in the sample set is less than a threshold  $T_1$  is greater than another threshold  $T_2$ , then the new pixel point is considered to be a background point, otherwise, it is a pixel in the motion region. Fig. 2 shows the steps, where  $D$  is the radius and the current pixel  $V(x,y)$  is the center of the circle [16].

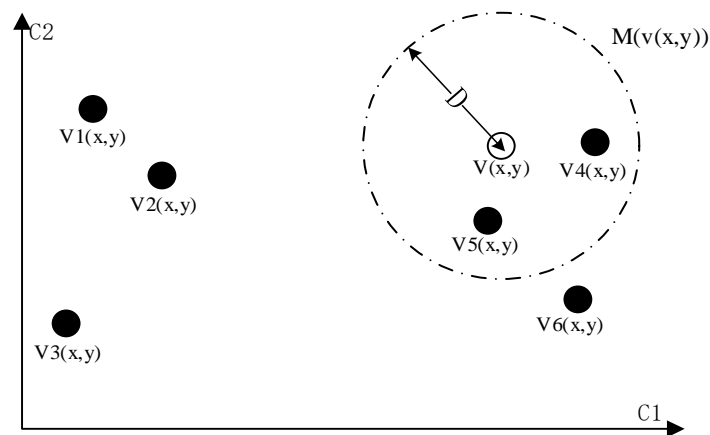


Fig. 2. ViBe algorithm

## 2.3 Gaussian mixture model for background modeling

The motion region detection method based on Gaussian mixture model is robust to the dynamic change of the scene. The basic idea is to define the distribution model of each pixel as a set consisting of multiple single Gaussian models. Then update the model parameters based on each new pixel value, according to some criteria, the pixel is determined as the background point or the pixel in the motion region [17].

## 2.4 IMFOA

IMFOA is proposed by Xuzhi Miao. The algorithm is a global optimization algorithm based on the process of fruit flies looking for food, which has the advantages of a few parameters and higher optimization accuracy [18].

Initialize the fruit fly population. The number of fruit fly groups is  $N$ , The number of iterations is  $G_{max}$ , randomly given group location:

$$X_{-axis}, Y_{-axis} \#(1)$$

Randomly give the direction and distance of the fruit fly's optimization,  $Rand()$  is defined as the optimization distance:

$$X_i = X_{-axis} + Rand() \#(2)$$

$$Y_i = Y_{-axis} + Rand() \#(3)$$

The source of the taste of the food is unknown.  $Dist_i$  is defined as the distance from fruit fly individuals to the origin.  $S_i$  is defined as the judgment value of the taste concentration of fruit flies.

$$Dist_i = \sqrt{X_i^2 + Y_i^2} \#(4)$$

$$S_i = \frac{1}{Dist_i} \#(5)$$

The taste concentration judgment value  $S_i$  is substituted into the taste concentration judgment function, and the taste concentration of the fruit fly individuals is defined as  $Smell_i$ . ( $Function()$  is the taste concentration judgment function, also called the fitness functions)

$$Smell_i = Function(S_i) \#(6)$$

Find the fruit fly with the best taste concentration in the fruit fly population.

$$[bestSmell \ bestindex] = \max(Smell_i) \#(7)$$

Based on the obtained  $Smell_i$  and  $bestSmell$  and the parameter  $a$ , where  $a$  is a random number with value (0,1). The fruit fly population is divided into two groups: good group and bad group. The formula is as follows:

$$a_i = \frac{Smell_i}{bestSmell} \geq a \#(8)$$

Fruit fly individuals that meet the above formula are assigned to the good group, and the remaining individuals are assigned to the bad group.

Record and preserve the best taste concentration  $bestSmell$  and its  $X$ ,  $Y$  coordinates, and other fruit fly individuals in the population use visual flight to this position.

$$Smell_{best} = bestSmell \#(9)$$

$$X_{-axis} = X(bestindex) \#(10)$$

$$Y_{-axis} = Y(bestindex) \#(11)$$

Finally, start iterative optimization and repeat steps 2) to 6), and when the taste concentration is no longer better than the previous taste concentration or when the number of repetitions reaches the maximum number of iterations, the cycle is stopped.

## 2.5 Histogram of Gradients features

The Histogram of Gradients features is also called HOG, it is based on the gradient direction histogram of the local region to form the feature, which was first proposed by Navneet Dalal for human body posture recognition [19].

## 2.6 Volume Local Binary Patterns

The Volume Local Binary Patterns is an extension of the local binary mode operator that is widely used in static texture analysis. The texture features of the image are analyzed by combining the motion and shape of the image. It was proposed by G. Zhao. and they also proposed a rotation-invariant Volume Local Binary Patterns [20].

## 3. Smoke recognition algorithm based on GMFOA optimized SVM

### 3.1 System framework

The system framework is shown in Fig. 3 below. The smoke video is obtained from the cameras, and then we use the mixed motion region extraction algorithm to extract the motion region of smoke video. After that, we extend the boundary of the motion region to an external rectangular region and divide the rectangular region into several regions of the same size, then we extract the HOG features and VLBP features of the several regions. Finally, we use GMFOA to optimize the parameters of the SVM and the synthetic kernel method of multiple kernel learning to classify the fused features.

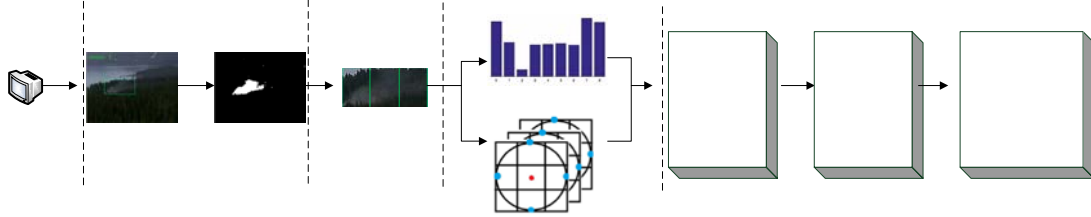


Fig. 3. The system framework

### 3.2 Mixed motion region extraction algorithm

Because the speed of smoke's movement is slow, it is difficult to extract a complete motion region of smoke video by using a single motion region extraction algorithm. Therefore, we compare the three-frame difference algorithm, ViBe algorithm, and Gaussian mixture model for background modeling to find the best combination.

The three-frame difference algorithm is an extension of the interframe difference method, and it is a method for obtaining a motion region by performing a pairwise difference operation on successive three frames of a video image sequence. Set the grayscale image of the current frame to be  $r_k$ , the grayscale image of the previous frame adjacent to it is  $r_{k-1}$ , the gray image of the next frame is  $r_{k+1}$ , then we make a difference between the pixel value of  $r_{k-1}$  and the corresponding pixel value of  $r_k$  to get a motion image  $S_k$  and a difference between the pixel value of  $r_k$  and the corresponding pixel value of  $r_{k+1}$  to get the motion region image  $S_{k+1}$ , we define the final motion region image as  $F_k = S_k \cap S_{k+1}$ . Finally, we form an external rectangular region based on the boundary extension of the motion region and divide this region into multiple motion regions of the same size.

$$F_k = \{F_{k1}, F_{k2}, F_{k3}, \dots\} \#(12)$$

The ViBe algorithm is a motion region extraction algorithm based on background updates. We define a region centered on pixel  $(x, y)$  and  $D$  as the radius,  $v(x, y)$  is the pixel's value. We make a sample set  $L(x, y) = \{v_1(x, y), v_2(x, y), \dots, v_N(x, y)\}$  of size  $N = 20$  to store the background point  $(x, y)$  of the previous frame and a new set  $M(v(x, y)) =$

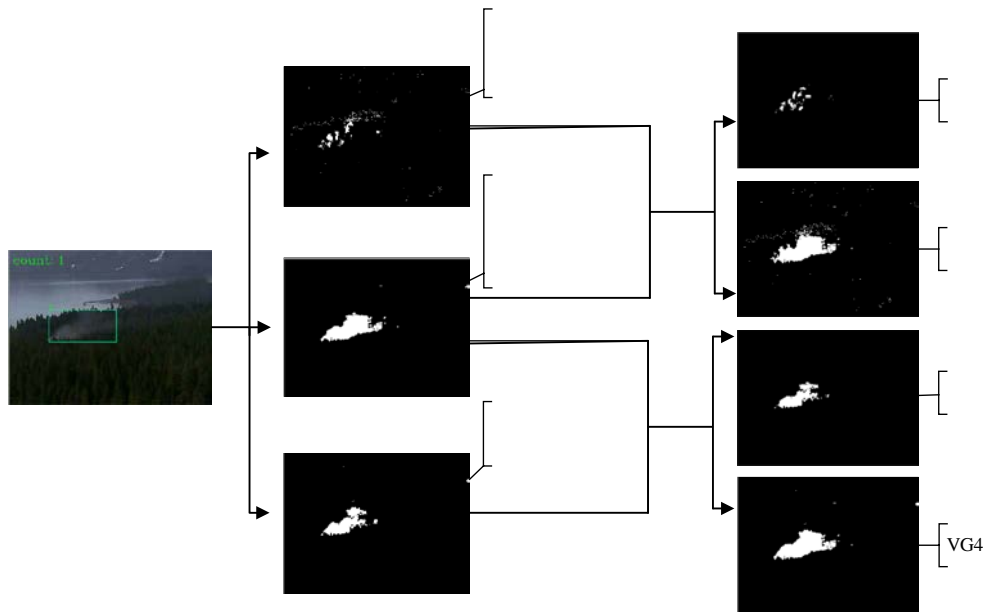
$\{v_1(x, y), v_2(x, y), \dots, v_N(x, y)\}$  of size  $N = 20$  to store the background point  $(x, y)$  of the next frame. Then we calculate the distance between each pixel in the new set and each pixel in the sample set, if the distance between those two pixels is greater than the threshold  $T_3$ , the pixel  $(x, y)$  which belongs to the new set is considered to be the motion pixel, otherwise the background pixel. We can get the motion region image  $V_k$  by this algorithm. Finally, we form an external rectangular region based on the boundary extension of the motion region and divide this region into multiple motion regions of the same size.

$$V_k = \{V_{k1}, V_{k2}, V_{k3}, \dots\} \#(13)$$

The Gaussian mixture model for background modeling is used to extract the motion region of smoke video. We define the RGB image of the current frame of smoke video is  $I_k$ , and its motion region which is extracted by the Gaussian mixture model for background modeling is  $G_k$ . Then we form an external rectangular region based on the boundary extension of the motion region and divide this region into multiple motion regions of the same size.

$$G_k = \{G_{k1}, G_{k2}, G_{k3}, \dots\} \#(14)$$

To select the best-mixed motion region extraction algorithm, we use four combinations of  $FG_1 = F_k \cap G_k$ ,  $FG_2 = F_k \cup G_k$ ,  $VG_3 = V_k \cap G_k$ , and  $VG_4 = V_k \cup G_k$  to extract the motion region of the same smoke video. The source of the smoke video is (<http://staff.ustc.edu.cn/~yfn/vsd.html>), and the experimental results are shown below.



**Fig. 4.** Mixed motion region extraction algorithm

It can be seen from **Fig. 4** that  $FG_1$  and  $VG_3$  can't completely extract the overall shape of the smoke, and the motion region extracted by  $FG_2$  has more background noises. The extracted motion region obtained by  $VG_4$  is more conducive to feature extraction of smoke images. Therefore, this paper uses  $VG_4$  methods to extract the motion region of the smoke video.

### 3.3 Feature extraction

#### 3.3.1 HOG feature extraction

The edges in the image contain abundant information. The HOG can effectively distinguish the difference between objects by counting the edge directions of different objects. Sobel operator is one of the most important operators in pixel image edge recognition, it combines Gaussian smoothing and differential derivation to calculate the approximate gradient of image gray function. The corresponding gradient vector or normal vector can be obtained by using the Sobel operator for each pixel in the image. The kernel of the Sobel operator is:

$$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, S_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

If  $I$  is the original image,  $G_x = S_x * I$  is the image detected by the lateral edge, and  $G_y = S_y * I$  is the image detected by the longitudinal edge. It is easy to know that the gradient of the image is  $[G_x, G_y]^T$ . Therefore, the gradient vector of each pixel is  $[G_x(x, y), G_y(x, y)]^T$ , and so we can infer its polar form is  $[m(x, y), \theta(x, y)]^T$ . Then we can easily get the amplitude  $m$  and the angle  $\theta$  of the gradient. The amplitude  $m$  is defined as follows:

$$m(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \quad (15)$$

The angle  $\theta$  is defined as follows:

$$\theta(x, y) = \arctan\left(\frac{G_x(x, y)}{G_y(x, y)}\right) \quad (16)$$

Divide 360 degrees into 12 bins, each bin contains 30 degrees, and the entire histogram contains 12 dimensions which are also the 12 bins. Then according to the gradient direction of each pixel, add its amplitude to the histogram by using bilinear interpolation.

#### 3.3.2 VLBP feature extraction

The basic LBP operator proposed by Ojala [21] for texture analysis can be defined as follows:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p, s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (17)$$

$g_c$  is the gray value of the center pixel,  $g_p$  is the gray value of the nearby pixel,  $R$  is the radius, and  $P$  is the total number of sampling points in the circular neighborhood. Fig. 5 is the basic calculation process for an LBP with  $P = 8$ ,  $R = 1$ . For any other value ( $P, R$ ), if a point on the circle is not in the image coordinates, we use his interpolation point to estimate the gray value of the neighbor at the center of the pixel by interpolation.

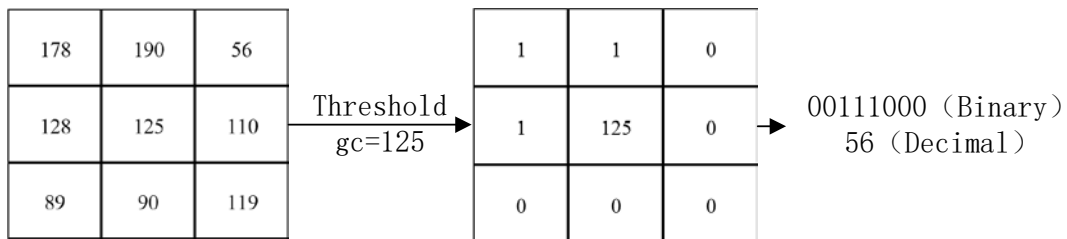


Fig. 5. The calculation process of LBP

The definition of VLBP is as follows:

$$VLBP_{L,P,R} = \sum_{p=0}^{3P+2} s(g_p - g_c) 2^p \quad (18)$$

$g_c$  corresponds to the gray value of the center pixel of the local neighborhood,  $g_p$  represents the gray value of  $3P+2$  pixels in the three frames having the interval  $L$ , and  $R$  is the radius of the circle in each frame image. Fig. 6 shows the calculation process of  $VLBP_{1,4,1}$ .

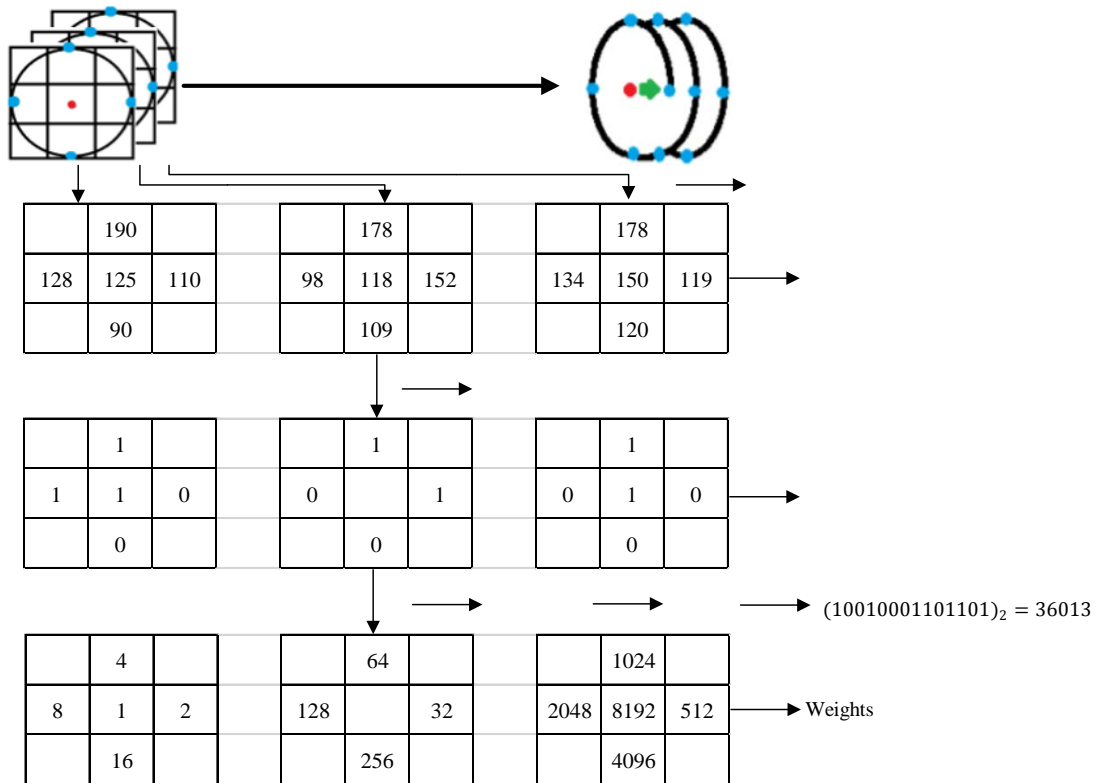


Fig. 6. The calculation process of  $VLBP_{1,4,1}$

We divide the recognition window into  $16 \times 16$  small cells and compare the gray value of 14 pixels in adjacent consecutive 3 frames with each pixel in each cell. This will produce a 14-bit binary number, which is the LBP value of the center pixel of the window. Then we calculate the histogram of each cell and normalize it. Then we connect the statistical histogram of each cell into a feature vector which is the VLBP texture feature vector of the entire image.

## 4. Optimization of SVM parameters based on GMVOA

### 4.1 SVM decision function and kernel function

When we use the RBF kernel as the kernel of the SVM, the kernel parameter  $\sigma$  and the penalty term  $c$  are the main influencing factors affecting the classification accuracy of the

SVM. The change of the parameter  $\sigma$  can affect the mapping function and change the distribution of the sample in the high-dimensional feature space, meanwhile, when the parameter  $\sigma$  constructs the optimal classification hyperplane in the high-dimensional feature space will directly affect the performance of the SVM. Therefore, the parameter  $\sigma$  has a great influence on the classification accuracy of SVM. The error penalty factor  $c$  is defined as the tolerance of the SVM for errors in the classification. With the increase of  $c$ , the classification accuracy of the training sample data set may be too high, which will lead to the low classification accuracy of the SVM on the test sample data set, and the generalization ability of the SVM becomes correspondingly low. On the contrary, with the decrease of  $c$ , the classification accuracy of the test sample data set may not reach the expectation, which will lead to the reduction of SVM classification ability. We can conclude that the prerequisite for ensuring the optimal classification result of SVM is to calculate the optimal SVM parameters [22][23].

#### 4.2 SVM parameter optimization method based on GMFOA

In IMFOA, iterative step  $Q_g$  of the good group in step (8) is:

$$Q_g = \frac{(Q_{max} + Q_{min}) - \frac{(Q_{max} - Q_{min})(O - 1)}{O_{max} - 1}}{2} \quad \#(19)$$

$Q_{max}$  is the maximum iteration step size,  $Q_{min}$  is the minimum iteration step size,  $O_{max}$  is the maximum number of iterations, and  $O$  is the current number of iterations. The iterative step  $Q_b$  of the bad group is:

$$Q_b = \frac{(Q_{max} + Q_{min}) + \frac{(Q_{max} - Q_{min})(O - 1)}{O_{max} - 1}}{2} \quad \#(20)$$

At the beginning of the algorithm, the moving step is  $(Q_{max} + Q_{min}) / 2$ . After the first iteration is completed, all flies are divided into two groups based on the clustering judgment formula (8). Then the two groups iterated until the number of iterations reached the maximum. After each iteration, all flies are regrouped. It can be seen from the formula (19) and the formula (20) that the search step size of the good group changes from large to small as the number of iterations increases, the search space becomes smaller, and the global optimization ability of the algorithm becomes weaker, the local optimization ability is enhanced to avoid missing the local optimal solution. On the contrary, the search step size of the bad group changes from small to large as the number of iterations increases, and the global optimization ability of the algorithm is enhanced, which can quickly search for the vicinity of the global optimal solution and avoid the algorithm falling into the local optimal solution. After each iteration, all flies are rearranged and then the two groups are optimized in each iteration step.

In the parameter optimization experiment of SVM in this paper, we use the IMFOA and discarding the iterative steps of the bad group to obtain a global or local optimal solution. In IMFOA, the advantages of the good group's global optimization ability and the advantage of the bad group's local optimization ability are weakened at the same time, which leads to the mediocrity of the good group and the bad group. And if the difference between the search step and the number of iterations increases from small to large, the search step can be simplified, and the overall optimization ability of the algorithm can be ensured.

**Algorithm 1** SVM classifier based on GMFOA

---

$N$ : Number of fruit fly population,  $X_{axis}$ : The X-axis coordinate of the fruit fly,  $Y_{axis}$ :  
 Y-axis coordinate of the fruit fly,  $Dist_i$ : The length between fruit fly and origin,  $S_i$ :  
 Reciprocal of  $Dist_i$   
 for  $gen = 1 : maxgen$   
     for  $i = 1 : sizepop$   
         if  $R(i) \geq \alpha$   
              $X(i) = X_{axis} + 2Q_g \text{ rand} - Q_g$ ;  
              $Y(i) = Y_{axis} + 2Q_g \text{ rand} - Q_g$ ;  
         end  
          $S(i) = 1 / D(i)$ ;  
          $smell(i) = \text{Function}(S_i)$ ;  
          $R(i) = smell(i) / bestsmell$ ;  
          $[bestsmell \text{ bestindex}] = \max(smell)$ ;  
         if  $smell(i) < bestsmell$   
              $bestsmell = smell(i)$ ;  
              $Smellbest = bestSmell$   
          $X_{axis} = X(bestIndex)$   
          $Y_{axis} = Y(bestIndex)$   
     end  
end

---

**4.3 GMFOA performance test**

To verify the classification ability of GMFOA. We use three optimization algorithms, GMFOA, IMFOA, and FOA to test three fitness functions. **Table 1** is the specific formula of the fitness functions given, and **Table 2** is the mean and standard deviation of the fitness functions obtained under the optimization of the three algorithms.

**Table 1.** Fitness functions

Function name	Function expression	Ranges
$f_1$	$f_1(x) = \sum_{i=1}^n x_i^2$	$[-100,100]$
$f_2$	$f_2(x) = \sum_{i=1}^n [x_i^2 + 10 \cos(2\pi x_i) + 10]$	$[-5.12,5.12]$
$f_3$	$f_3(x) = \sum_{i=1}^n (\sum_{j=1}^j x_j)^2$	$[-100,100]$

**Table 2.** Experimental comparison of different fitness functions

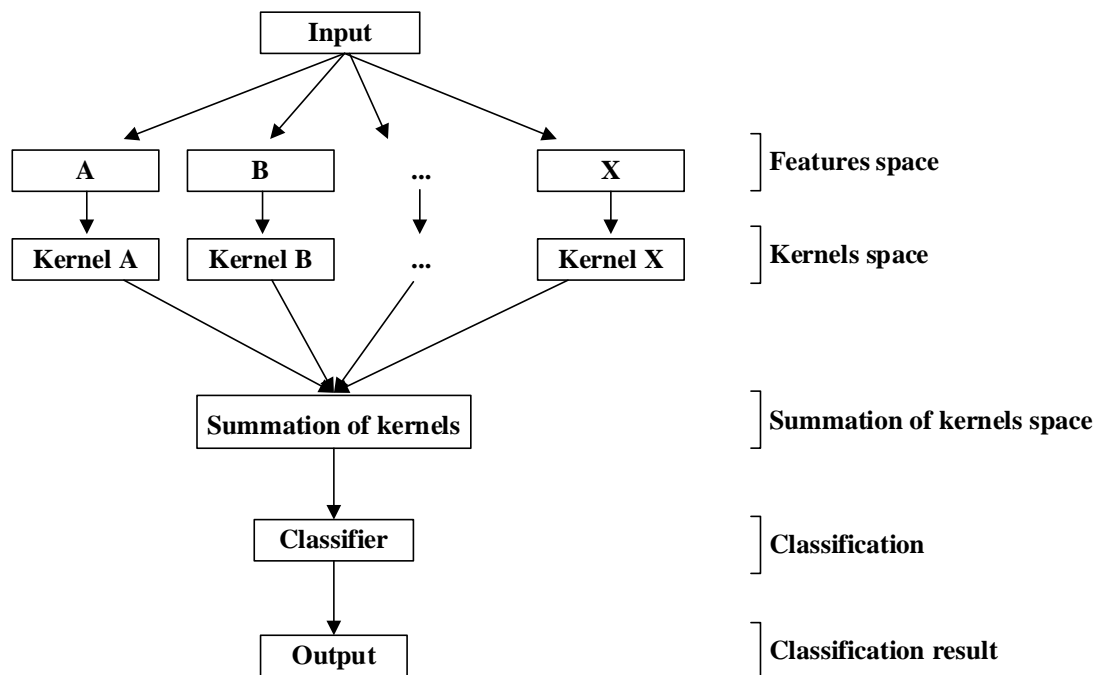
Function name		FOA	IMFOA	GMFOA
$f_1$	Mean	1.27e-07	6.51e-08	6.31e-08
	Std	8.61e-07	4.69e-07	4.18e-07
$f_2$	Mean	3.71	2.96	2.61
	Std	4.98	5.84	5.59
$f_3$	Mean	1.23e-07	8.89e-08	8.63e-08
	Std	6.97e-07	7.56e-07	7.32e-07

It can be seen from [Table 2](#) that among the three fitness functions, the accuracy of the mean and standard deviation of GMFOA is significantly better than the other two algorithms, and GMFOA has better convergence speed and optimization accuracy than the IMFOA algorithm.

#### 4.4 Multiple kernel learning method based on GMVOA optimized SVM

It is not ideal to use the traditional SVM method for classifying smoke images, because the features of samples from multiple sources have different characteristics. The studies show that using multiple kernel learning can improve the interpretability and performance of decision functions. When we use multiple kernel learning methods to fuse features, we are essentially combining several basic kernel matrices corresponding to a single feature. We can implement heterogeneous data fusion on the multiple kernel matrix synthesized based on multiple basic kernel matrices and use it to train the classifier.

The multiple kernel learning method [\[24\]\[25\]](#) has different classification methods based on different classification criteria. According to the construction methods and characteristics of multiple kernel functions, multiple kernel learning methods can be roughly divided into three categories: synthetic kernel learning method, multiscale kernel learning method, and infinite kernel learning method. The synthetic kernel learning method is to combine multiple kernel functions with different characteristics, so that we can get a multiple kernel function containing the characteristics of each kernel function. The multiple kernel learning method is more accurate and Stronger mapping capabilities than the traditional SVM method, in practical applications, the advantages of multiple kernel learning will be very obvious for the classification and regression of sample data with complex distribution structures. The synthetic multiple kernel learning method is shown in [Fig. 7](#).



**Fig. 7.** Synthetic multiple kernel learning method

In this paper, we compared the recognition rate of the four methods of  $Method_1 = VLBP/RBF, HOG/RBF$ ,  $Method_2 = VLBP/RBF, HOG/Sigmoid$ ,  $Method_3 = VLBP/Sigmoid, HOG/Sigmoid$ ,  $Method_4 = VLBP/RBF$  to the smoke video of two different scenes of forest and town. **Table 3** is the recognition rate of the different methods.

**Table 3.** the recognition rate of the different methods

	$Method_1$	$Method_2$	$Method_3$	$Method_4$
Vid1(Forest)	0.9266	0.9079	0.8943	0.8768
Vid2(Town)	0.9138	0.8991	0.8926	0.8421

We can see from **Table 3**, the recognition rate of the  $Method_1$  is higher than other methods, and the recognition rate of  $Method_1$ ,  $Method_2$ ,  $Method_3$  is higher than the  $Method_4$ . This shows that GMFOA for SVM parameter optimization is helpful to improve the recognition rate for the smoke image, and the recognition rate of multiple kernel learning method is higher than the traditional SVM method.

## 5. Analysis of results

To verify the effectiveness of the algorithm we proposed in this paper, we pick two interference videos and four smoke videos, as shown in **Fig. 8** and **Fig. 9**. The video frame rate used in the experiment was 25 fps and the video size was 320×240. The algorithm is based on the Windows 10 operating system. The computer is configured as Intel(R) Core(TM) i7-7700HQ CPU @ 2.80GHz, 8.00 GB memory, and the graphics card is NVIDIA GeForce GTX 1050 Ti.

The experimental results are evaluated by true positive (TP), false positive (FP), true negative (TN), and false negative (FN) [26]. The calculation formula of the recognition rate is  $TP/(TP+FN)$ , which means in all smoke videos, the proportion of frames where smoke appears among all the frames recognized as smoke. The recognition rate represents the probability that the system correctly recognizes smoke images. The calculation formula of the false alert rate is  $FN/(TN+FN)$ , which means in all interference videos, the number of frames where non-smoke appears is mistakenly recognized as the proportion of the number of frames where smoke appears in the total number of non-smoke frames. The false alert rate represents the probability of system failure.



Vid 3



Vid 4

**Fig. 8.** Two interference videos

**Fig. 9.** Four smoke videos

In these 6 videos, vid3 and vid4 are interference videos, and vid5 is a smoke video in the forest which the ignition point is far away from cameras, vid6 is a smoke video in the forest which the ignition point is closer to the cameras, vid7 is a smoke video in the forest with interference factors (smoke and clouds mixed), vid8 is a smoke video in town.

We use the three methods of GMFOA-SVM (*Method<sub>1</sub>*), IMFOA-SVM and TsPSO-SVM for the experiment [27], **Table 4** is the false alert rate of those three methods for interference videos, **Table 5** is the recognition rate of those three methods for smoke videos.

**Table 4.** Interference videos false alert rate of three methods

Video sequence	GMFOA-SVM	IMFOA-SVM	TsPSO-SVM
Vid3	0.1322	0.1439	0.2014
Vid4	0.1476	0.1589	0.2427

**Table 5.** Smoke videos recognition rate of three methods

Video sequence	GMFOA-SVM	IMFOA-SVM	TsPSO-SVM
Vid5	0.9231	0.9216	0.9132
Vid6	0.9331	0.9252	0.9164
Vid7	0.9034	0.9027	0.8969
Vid8	0.9145	0.9123	0.9013

It can be seen from **Table 4** that the false alert rate of the GMFOA-SVM method is slightly lower than the IMFOA-SVM method and the TsPSO-SVM method. Because the smoke near the camera has obvious HOG features and VLBP features. It can be seen from **Table 5** that the recognition rate of the GMFOA-SVM method for Vid 6 is higher than Vid 5. When smoke and clouds are mixed, the recognition rate of forest smoke by those three

methods is generally low, which indicates that the HOG features and VLBP features of smoke are not ideal for classification. We can conclude the experiment that the GMFOA algorithm has fewer parameters which is better than the TsPSO algorithm when optimizing SVM parameters. The GMFOA-SVM method has the highest recognition rate of smoke images in those three methods.

## 6. Conclusion

In this paper, we use the good group of improved fruit fly optimization algorithm to optimize the parameter of SVM, then, we detect the smoke image based on the obtained smoke image features. When we optimized the parameters of SVM, by improving the fruit fly optimization algorithm, the recognition accuracy of smoke images was improved. The algorithm proposed in this paper has important practical significance for the research and application of fire smoke monitoring based on video and pictures. However, the algorithm proposed in this paper is slightly inadequate in real-time monitoring, and there will be a high rate of misjudgment for the places where the interference factors such as smoke and non-smoke exist at the same time. When under the interference of strong light and strong wind, the recognition rate of the algorithm we proposed in this paper will be further reduced, which still needs further improvement.

## Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 61772561; in part by the Key Research and Development Plan of Hunan Province under Grant 2018NK2012; in part by the Science Research Projects of Hunan Provincial Education Department under Grant 18C0262; in part by the Degree & Postgraduate Education Reform Project of Hunan Province under Grant 2019JGYB154; in part by the Postgraduate Excellent teaching team Project of Hunan Province under Grant [2019]370-133; and in part by the Natural Science Foundation of Hunan Province (No.2020JJ4140, 2020JJ4141).

## References

- [1] Y. Tan, J. H. Qin, X. Y. Xiang, W. T. Ma, W. Y. Pan, N. N. Xiong, "A Robust Watermarking Scheme in YCbCr Color Space Based on Channel Coding," *IEEE Access*, Vol. 7, No. 1, pp. 25026-25036, 2019. [Article \(CrossRef Link\)](#)
- [2] W. P. Lu, X. Zhang, H. M. Lu, F. F. Li. "Deep hierarchical encoding model for sentence semantic matching," *Journal of Visual Communication and Image Representation*, Vol. 71, 102794, 2020. [Article \(CrossRef Link\)](#)
- [3] L. Xiang, Y. Li, W. Hao, P. Yang, and X. Shen. "Reversible Natural Language Watermarking Using Synonym Substitution and Arithmetic Coding," *CMC: Computers, Materials & Continua*, Vol. 55, No. 3, pp. 541-559, 2018. [Article \(CrossRef Link\)](#)
- [4] J. H. Qin, W. Y. Pan, X. Y. Xiang, Y. Tan, G. M. Hou, "A biological image classification method based on improved CNN," *Ecological Informatics*, Vol.58, No. 2020, pp. 1-8, 2020. [Article \(CrossRef Link\)](#)
- [5] J. Wang, J. H. Qin, X. Y. Xiang, Y. Tan, N. Pan, "CAPTCHA recognition based on deep convolutional neural network," *Mathematical Biosciences and Engineering*, Vol. 16, No. 5, pp. 5851-5861, 2019. [Article \(CrossRef Link\)](#)

- [6] Y. T. Chen, J. J. Tao, L. W. Liu, J. Xiong, R. L. Xia, J. B. Xie, Q. Zhang, K. Yang, "Research of improving semantic image segmentation based on a feature fusion model," *Journal of Ambient Intelligence and Humanized Computing*, 2020. [Article \(CrossRef Link\)](#)
- [7] Y. T. Chen, J. Wang, S. J. Liu, X. Chen, J. Xiong, J. B. Xie, K. Yang, "Multiscale fast correlation filtering tracking algorithm based on a feature fusion model," *Concurrency and Computation-Practice and Experience*, 2019. [Article \(CrossRef Link\)](#)
- [8] Y. J. Luo, J. H. Qin, X. Y. Xiang, Y. Tan, Q. Liu, L. Y. Xiang, "Coverless real-time image information hiding based on image block matching and Dense Convolutional Network," *Journal of Real-Time Image Processing*, Vol. 17, No. 1, pp. 125-135, 2020. [Article \(CrossRef Link\)](#)
- [9] Q. Liu, X. Y. Xiang, J. H. Qin, Y. Tan, J. S. Tan, Y. J. Luo, "Coverless steganography based on image retrieval of DenseNet features and DWT sequence mapping," *Knowledge-Based Systems*, Vol. 192, No. 2020, pp. 105375-105389, 2020. [Article \(CrossRef Link\)](#)
- [10] H. Li, J. H. Qin, X. Y. Xiang, L. L. Pan, W. T. Ma, N. N. Xiong, "An efficient image matching algorithm based on adaptive threshold and RANSAC," *IEEE Access*, Vol. 6, No. 1, pp. 66963-66971, 2018. [Article \(CrossRef Link\)](#)
- [11] A. U. Russo, K. Deb, S. C. Tista and A. Islam, "Smoke Recognition Method Based on LBP and SVM from Surveillance Camera," in *Proc. of 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*, Rajshahi, pp. 1-4, 2018. [Article \(CrossRef Link\)](#)
- [12] M. Toptaş and D. Hanbay, "Smoke recognition using texture and color analysis in videos," in *Proc. of 2017 International Artificial Intelligence and Data Processing Symposium (IDAP)*, Malatya, pp. 1-4, 2017. [Article \(CrossRef Link\)](#)
- [13] M. Huang, Y. Wang and Y. Hu, "Smoke identification of low-light indoor video based on support vector machine," in *Proc. of 2017 3rd IEEE International Conference on Computer and Communications (ICCC)*, Chengdu, pp. 2045-2049, 2017. [Article \(CrossRef Link\)](#)
- [14] Y. T. Chen, J. J. Tao, Q. Zhang, K. Yang, X. Chen, J. Xiong, R. L. Xia, J. B. Xie, "Saliency detection via improved hierarchical principle component analysis method," *Wireless Communications and Mobile Computing*, Vol. 2020, Article ID 8822777, 2020. [Article \(CrossRef Link\)](#)
- [15] L. Wang and A. Li, "Early fire recognition based on multi-feature fusion of video smoke," in *Proc. of 2017 36th Chinese Control Conference (CCC)*, Dalian, pp. 5318-5323, 2017. [Article \(CrossRef Link\)](#)
- [16] O. Barnich and M. Van Droogenbroeck, "ViBE: A powerful random technique to estimate the background in video sequences," in *Proc. of 2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, pp. 945-948, 2009. [Article \(CrossRef Link\)](#)
- [17] C. Zhang and Q. Xu, "An Improved GMM based Video Foreground Separation," in *Proc. of 2019 Chinese Control And Decision Conference (CCDC)*, Nanchang, China, pp. 1371-1374, 2019. [Article \(CrossRef Link\)](#)
- [18] X. Z. Miao, F. M. Bi, X. Y. Fu, et al, "Mine Fire Image Recognition Based on Improved FOA-SVM," *Computer Engineering*, Vol. 45, No. 4, pp. 267-274, 2019. [Article \(CrossRef Link\)](#)
- [19] N. Dalal and B. Triggs, "Histograms of oriented gradients for human recognition," in *Proc. of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, CA, USA, Vol. 1, pp. 886-893, 2005. [Article \(CrossRef Link\)](#)
- [20] G. Zhao and M. Pietikainen, "Dynamic Texture Recognition Using Volume Local Binary Patterns with an Application to Facial Expressions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 6, pp. 915-928, June 2007. [Article \(CrossRef Link\)](#)
- [21] T. Ojala, M. Pietikäinen, and D. Harwood, "A Comparative Study of Texture Measures with Classification Based on Feature Distributions," *Pattern Recognition*, Vol. 29, pp. 51-59, 1996. [Article \(CrossRef Link\)](#)
- [22] Mamouni El Mamoun, Zennaki Mahmoud and Sadouni Kaddour, "SVM Model Selection Using PSO for Learning Handwritten Arabic Characters," *Computers, Materials & Continua*, Vol. 61, No. 3, pp. 995-1008, 2019. [Article \(CrossRef Link\)](#)

- [23] L. Xiang, J. Yu, C. Yang, D. Zeng and X. Shen, "A Word-embedding-based Steganalysis Method for Linguistic Steganography via Synonym-substitution," *IEEE Access*, Vol. 6, pp. 64131-64141, 2018. [Article \(CrossRef Link\)](#)
- [24] M. Alioscha-Perez, M. C. Oveneke and H. Sahli, "SVRG-MKL: A Fast and Scalable Multiple Kernel Learning Solution for Features Combination in Multi-Class Classification Problems," *IEEE Transactions on Neural Networks and Learning Systems*, 2019. [Article \(CrossRef Link\)](#)
- [25] L. Xiang, G. Zhao, Q. Li, W. Hao and F. Li, "TUMK-ELM: A fast unsupervised heterogeneous data learning approach," *IEEE Access*, Vol. 6, pp. 35305-35315, 2018. [Article \(CrossRef Link\)](#)
- [26] A. Filonenko, L. Kurnianggoro and K. Jo, "Comparative study of modern convolutional neural networks for smoke recognition on image data," in *Proc. of 2017 10th International Conference on Human System Interactions (HSI)*, Ulsan, pp. 64-68, 2017. [Article \(CrossRef Link\)](#)
- [27] X. Yishu and M. Lin, "A tri-level programming model based on TsPSO for cabin layout selection in civil airlines," in *Proc. of 2017 2nd International Conference on System Reliability and Safety (ICSRS)*, Milan, pp. 526-530, 2017. [Article \(CrossRef Link\)](#)



**Jingwen Liu** received his B.S. in Information and Computing Science from Shaanxi University of Science & Technology, China, in 2016. He is currently pursuing his M.S. in Information and Communication Engineering at the School of computer and Information Engineering of Central South University of Forestry and Technology, China. His research interests include machine learning, image processing, and pattern recognition.



**Junshan Tan** received the B.S. degree in physics from Central China Normal University, China, in 1985, the M.S. degree in industrial automation from Central South University, China, in 1997, and the Ph.D. degree in ecology from Central South University Of Forestry and Technology, China, in 2010. He is currently a professor with the College of Computer Science and Information Technology, Central South University of Forestry and Technology, China. His research interests are information management and computer application technology.

E-mail: tan\_junshan@yahoo.com.cn



**Jiaohua Qin** received the B.S. degree in mathematics from the Hunan University of Science and Technology, China, in 1996, the M.S. degree in computer science and technology from the National University of Defense Technology, China, in 2001, and the Ph.D. degree in computing science from Hunan University, China, in 2009. She was a Visiting Professor with the University of Alabama, Tuscaloosa, AL, USA, from 2016 to 2017. She is currently a Professor with the College of Computer Science and Information Technology, Central South University of Forestry and Technology, China. Her research interests include network and information security, machine learning and image processing.

E-mail: qinjiaohua@163.com



**XUYI XIANG** received his B.S. in mathematics from Hunan Normal University, China, in 1996, M.S. degree in computer science and technology from National University of Defense Technology, China, in 2003, and PhD in computing science from Hunan University, China, in 2010. He is a professor with the College of Computer Science and Information Technology, Central South University of Forestry and Technology, China. His research interests include network and information security, image processing and machine learning.

E-mail: xyuxiang@163.com