IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 1, 2022, accepted April 18, 2022, date of publication April 25, 2022, date of current version May 6, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3170464

An Optimized Multisource Bilinear Convolutional Neural Network Model for Flame Image Identification of Coal Mine

LI ZHANG[1], YUQIN ZHU[1], HAO WU[2], AND KUN LI[1,3]

[1]State Key Laboratory of Coal Mine Safety Technology, China Coal Technology & Engineering Group Shenyang Research Institute, Shenfu Demonstration Zone, Fushun 113122, China
[2]College of Control Science and Engineering, Bohai University, Jinzhou 121013, China
[3]Faculty of Electrical and Control Engineering, Liaoning Technical University, Huludao 125105, China

Corresponding author: Kun Li (phdlikun@163.com)

This work was supported in part by the Joint Open Fund Project of the State Key Laboratory of Coal Mine Safety Technology of Liaoning Province under Grant 2020-KF-13-04.

ABSTRACT Underground fire monitoring is an important tool to improve coal mine production safety. In this paper, a multi-source information identification method based on bilinear convolutional neural network (B-CNN) is proposed, which consists of construction of multi-source image acquisition system, B-CNN and integrated decision making based on multi-source B-CNN. Aiming at the problem that Softmax loss function based on the gradient descent in B-CNN is easy falling into the local optimum, an improved Grasshopper Optimization Algorithm (GOA) is proposed to optimally selected two parameters of $\boldsymbol{W}$ and $\boldsymbol{\theta}$; the method of initial solution generation based on sine mapping and the method of accepting bad solution with certain probability are respectively adopted. In order to solve high computational complexity in the stage of model training and integrated recognition by multi-source B-CNN, an image feature preprocessing method is proposed in this paper. Several feature vectors of color feature, shape feature and texture feature of the collected image are extracted and used as input vectors of B-CNN to complete model training and integrated recognition. In simulation experiments, firstly, four Benchmark functions are used to verify the performance of the improved GOA; then, by scaling, expanding and rotating the image to simulate the results of image acquisition at multiple positions and angles, different information sources can be formed to complete the integrated recognition by B-CNN. Three performance indexes of Accuracy, Precision and Recall are used to evaluate the simulation result of different comparative models, which show that the proposed method has better recognition effects.

INDEX TERMS Bilinear convolutional neural network, coal mine, data modeling, feature extraction, image processing.

I. INTRODUCTION

In underground coal mine fire disaster, spontaneous combustion fire is the focus of prevention and control, mainly caused by own oxidation and combustion of coal or other combustible heat accumulation. In recent years, most coal mines pay more and more attention to the construction of underground production monitoring system, through the installation of various types of data acquisition equipment and the construction of network transmission system, so that the production management personnel can as quickly as possible to grasp the underground production dynamic and abnormal changes. At present, the underground coal mine safety monitoring mainly focuses on the following aspects:

(1) Data acquisition and analysis based on advanced sensors. Laser gas analyzer, infrared thermometer, optical fiber temperature measurement and other devices or equipment are used to monitor the underground production environment, and the data is transmitted to the ground monitoring center through the cable network, Internet of Things, Zigbee wireless network, etc., and the parameters are input into the database for comparative analysis, and then feedback to the underground production personnel.

The associate editor coordinating the review of this manuscript and approving it for publication was Shunfeng Cheng.

(2) Production disaster prediction based on multi-parameter. For production disasters such as fire, on the basis of mechanism analysis and production technology, several production parameters which are easy to be measured are directly used as auxiliary variables, and then linear or nonlinear mathematical models are established to predict different types of production disasters.

(3) Underground monitoring based on image processing. Video monitoring equipment is installed in key areas such as coal conveyance belt, electrical equipment and personnel operation position, to take pictures of the working status of production system and staff movement. Then, the image processing algorithm is used to reduce noise and enhance the image, so as to improve the readability of the image, and better grasp the underground working conditions.

At present, the combination of image processes and machine learning to realize underground coal mine fire monitoring has attracted more and more attention of researchers. The image contains rich information, good continuity, and the image processing method is low cost, operability. The automatic recognition of flame image can be realized by using machine learning method and training model from enough samples. In recent years, support vector machine (SVM) [1]–[7], feed-forward artificial neural network [8], wavelet neural network [9], fuzzy neural network [10]–[11], probabilistic neural network (PNN) [12], RBF neural network [13], BP neural network [14], extreme learning machine (ELM) [15], etc. have been applied in the fire image recognition. However, the generalization performance of SVM is poor and it is prone to over-fitting, and the prediction accuracy of the selected kernel function in SVM is greatly affected by data distribution characteristics; in RBF and TWSVM networks, problems such as poor generalization ability and easy over-fitting will also occur due to the use of high-dimensional data mapping; ELM relies on the accurate selection of the number of hidden layer nodes, and excessive number of hidden layer nodes will increase the calculation amount of the model, and it is easy over-fitting; the convergence rate of BP is slow and it is easy to fall into the local extremum; the selection of wavelet function in wavelet neural network is subjective and its prediction accuracy is greatly affected by data characteristics; the choice of membership function in fuzzy neural network is subjective and has great influence on the prediction accuracy of model. Nowadays, with the continuous development of deep learning technology, flame image recognition based on deep learning algorithm has been widely concerned. Muhammad *et al.*. [16] constructed an early fire detection framework by adding depth features of convolutional neural network (CNN); Saeed *et al.*. [17] proposed a CNN mixed model for early fire detection; Sharma *et al.*. [15] mixed two models of CNN and ELM to realize fire image detection; Pereira *et al.*. [18] proposed active fire detection method in satellite images, by introducing large scale active fire detection data set and using different CNN architectures; Ayala *et al.*. [19] proposed a new deep learning architecture for fire image recognition, which adopts inverted residual blocks, deep convolutions and octave to reduce the computational cost of the model, and at the same time, it can ensure high calculation accuracy; Zhu and Ren [20] proposed a flame image recognition method based on deep learning and particle algorithm, which introduced RGB and HIS systems to realize multi-feature fusion flame recognition; Sun *et al.*. [21] proposed an improved CNN to realize the rapid detection of forest fire smoke; Cao *et al.*. [22] proposed an attention enhanced bidirectional long short-term memory network (ABi-LSTM) for smoke identification of forest fires in video; Pundir and Raman [23] proposed a robust smoke detection method based on dual deep learning framework, the first deep learning framework extracted image-based features from smoke patches, and the second deep learning framework was used to extract motion-based features, which were then input into CNN classifier to complete classification; Yang *et al.*. [24] proposed a neural network model combining lightweight CNN and SRU, which could reduce the influence of strong interference, such as bright light flicker or high brightness background on single-frame fire image recognition. Bilinear convolutional neural network (B-CNN) [25] is a deep learning model based on weakly supervised learning. Two convolutional neural networks are adopted, in which, one network extracts the location information of images, while the other network extracts the appearance information of images. The classification effect of B-CNN is better than the image classification method based on image filtering. Zhou *et al.*. [26] trained deep convolutional neural network to analyze multidimensional emotional facial expressions, and bilinear pools were used to encode second-order features, which proved that B-CNN model had better performance; Dong *et al.*. [27] proposed combining B-CNN and pair-wise difference pooling (PDP) for texture classification of fine-grained images, which could not only obtain pairwise difference between two groups of features, but also encoded the difference between each pair of features; Tang *et al.*. [28] proposed a new spatial attention bilinear convolutional neural network (SA-BCNN) to detect defects in casting X-ray images, by combining spatial attention mechanism with bilinear pooling. Although the research on B-CNN has achieved some fruitful results, it can be further improved in optimizing B-CNN model, reducing the complexity of training process and carrying out multi-source information fusion. In this paper, B-CNN is used to establish an underground flame image recognition model, and an improved Grasshopper Optimization Algorithm (GOA) is proposed to optimize the parameters $\boldsymbol{W}$ and θ of the Softmax loss function in B-CNN; the color feature, shape feature and texture feature of flame image are used as input variables to train the model, so as to improve the processing speed of B-CNN. In addition, in order to reduce the influence of various interference factors on single source information processing, this paper adopts multi-source image acquisition method and multi-source information integration decision method to realize the coal mine flame image recognition. GOA is a

novel swarm intelligence optimization algorithm [29]. Compared with Particle Swarm Optimization Algorithm (PSO), Pigeon-Inspired Optimization Algorithm (PIO), Grey Wolf Optimization Algorithm (GWO), Bat Algorithm (BA), fruit fly optimization Algorithm (FOA), Brain Storm Optimization Algorithm (BSO), etc., it has advantages of simple structure, few parameters, larger operation space, and so on. However, it still has room for further improvement in convergence accuracy and global search capability of GOA. So, in this paper, some works about its method of initial solution generation and global search capability are conducted to further improve it; and it is applied to optimally select two parameters $\boldsymbol{W}$ and θ of Softmax loss function in B-CNN.

The main contributions of this paper are summarized as follows:

(1) An improved GOA was proposed to optimally select $\boldsymbol{W}$ and θ of Softmax loss function, which has better performance in initial solution generation and global search capability;

(2) A multi-information source image acquisition system in underground working face of coal mine was designed, in which several image acquisition devices in different positions were set up to acquire multi-angle and multi-directional images;

(3) A flame image recognition model based on multi-information source bilinear convolutional neural network was proposed, which was formed by image feature extraction unit, model training unit and, recognition and detection unit;

(4) In order to decrease the high computational complexity in B-CNN training and integration recognition stage, an image feature extraction method based on preprocessing was proposed.

The remaining part of this paper are arranged as follows: Section 2 introduces the research background of this paper, that is, remote image transmission system in underground coal mine and multi-information source image acquisition system in underground working face; in Section 3, bilinear convolutional neural network, multi-information source image processing methods based on bilinear convolutional neural network are briefly introduced; in section 4, the existing problems and proposed solutions are presented, firstly, an improved GOA is given, and then, the extracted color features, shape features and texture feature vectors are introduced; Section 5 introduces the flame image recognition model based on multi-information source bilinear convolutional neural network, including image feature extraction unit, model training unit and recognition and detection unit; in Section 6, different simulation experiments are designed to verify the effectiveness of the proposed method; Section 7 summarizes the main contributions of this paper.

II. RESEARCH BACKGROUND

The image acquisition instrument is placed in a specific position underground, to monitor objects in the target area. The mine remote image transmission system is mainly composed of image acquisition instrument, the underground control microcomputer (MCU), power supply, communication cable, ground MCU, wireless communication module, antenna, etc. The collected images are transmitted to wellhead by cable communication transmission, and then transmitted to the ground monitoring center by wireless communication unit, so as to realize the continuous acquisition of underground coal mine image. The main structure of remote image transmission system in coal mine is shown in Fig. 1.

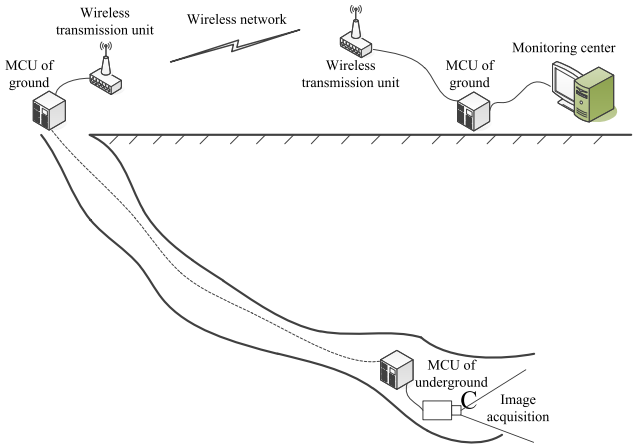


FIGURE 1. Main structure of remote image transmission system in coal mine.

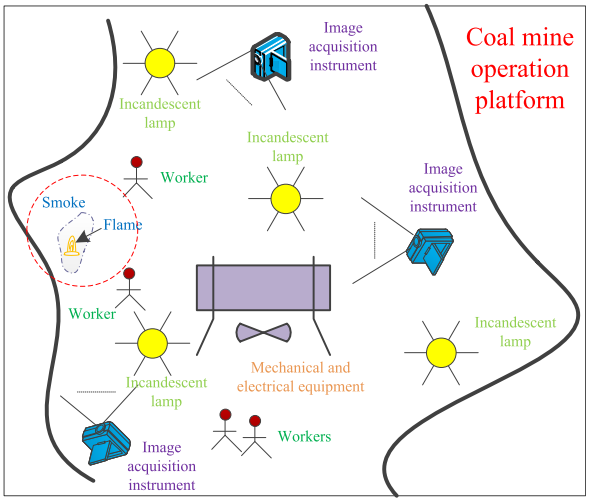


FIGURE 2. Structure diagram of multi-information source image acquisition system in underground working face.

Early fire has weak characteristics such as high temperature in the local area, a small amount of smoke, smaller flame, etc. and because of the distance, mine dust, environment temperature, worker, incandescent, mechanical and electrical equipment, the influence of early fire image is not easy to identify, make the existing method based on a single source of information processing has higher mistake rate and false negative rate. Therefore, based on the structure of remote image transmission system in underground coal mine as shown in Fig. 1, this paper constructs a multi-information source image acquisition system in underground working face site, as shown in Fig. 2.

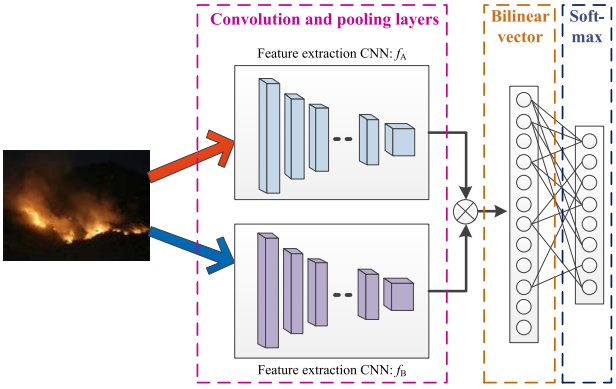


FIGURE 3. Network architecture of B-CNN.

According to Fig. 2, multi-information source image acquisition method is adopted in this paper to solve the problem that flame image features are easily interfered by various complex factors, such as: underground incandescent light, activities of works, lighting and heating of mechanical and electrical equipment. By setting up image acquisition devices in different positions, multi-angle and multi-directional image acquisition is carried out for the working face area, and then the integrated identification of flame features is carried out by multi-source image information processing method. Compared with the single information source processing method, the multi-source method can obtain the most useful information and reduce the influence of various interference factors as much as possible, so as to improve the accuracy of flame image recognition.

III. METHODOLOGY

In this paper, B-CNN is used for flame image recognition and processing. Firstly, a certain number of early flame images are collected to establish the training sample library. Then, the B-CNN based deep learning method is selected to design the operation mechanism, and the structure of the model is trained by the training sample library. Finally, the trained model is used to identify the enhanced images to determine whether the fire occurs.

A. BILINEAR CONVOLUTIONAL NEURAL NETWORK

The network architecture of B-CNN is shown in Fig. 3.

The mathematical representation of B-CNN is as follows:

$$B=(f_A, f_B, P, C) \tag{1}$$

where f_A is the feature function of the first convolutional network to extract convolutional network features, and f_B is the feature function of the other convolutional network to extract convolutional network features; P represents bilinear pooling function, which is the confluence of two convolution image features extracted from f_A and f_B; C is a network classification function used to map the space of sample markers and bilinear pooling image features. The feature extraction function can be regarded as a function mapping, which maps the input image I and position region l to a $C \times D$ dimension feature. The bilinear feature of the input image at a certain position I can be expressed by the following formula:

$$bilinear(l, I, f_A, f_B) = f_A(l, I)^T f_B(l, I) \tag{2}$$

Then, the pooling function P is used to accumulate bilinear features at all positions into one bilinear feature, and this feature is used to describe the features of the input image:

$$\phi(I) = \sum_{l \in L} bilinear(l, I, f_A, f_B) = \sum_{l \in L} f_A(l, I)^T f_B(l, I) \tag{3}$$

Set $A \in R_{\mathrm{m}} \times l_1$ and $B \in R_{\mathrm{m}} \times l_2$ respectively represent the two-dimensional matrices output by f_A and f_B:

$$\phi(I) = A^T B \tag{4}$$

Normalize the bilinear feature after vectorization as follows:

$$Y = \sin(\phi)\sqrt{|\phi|} \tag{5}$$

$$Y' = \frac{Y}{\|Y\|_2} \tag{6}$$

The calculation results of (6) are taken as the final representation of image features, and Softmax is used to classify images.

During the training of bilinear convolutional neural network, two convolutional neural branch networks can be trained at the same time, and the gradient of the pooling layer of bilinear convolutional neural network can be calculated by using the chain rule, and then the error is propagated back to obtain the gradient update. Assuming that for each position l, the output of feature extraction functions f_A and f_B are f_1 and f_2 respectively, then the bilinear feature at l is:$x = \mathbf{f}_1^T \mathbf{f}_2$. d$l$ /dx is used to represent the gradient value of the loss function at x, and the gradient value of the output of the loss function to the two networks A and B can be obtained by the chain rule:

$$\begin{cases} \dfrac{dl}{dA} = B\left(\dfrac{dl}{dx}\right)^T \\ \dfrac{dl}{dB} = A\left(\dfrac{dl}{dx}\right)^T \end{cases} \tag{7}$$

To sum up, the function of network A is to locate the object, and network B is used to extract the features of the object detected by network A. The two networks coordinate with each other to accomplish the tasks of region detection and feature extraction.

B. MULTI-INFORMATION SOURCE IMAGE PROCESSING BASED ON B-CNN

According to the multi-information source image acquisition method in underground coal mine shown in Fig. 2, and combined with the bilinear convolutional neural network shown in Fig. 3, the system structure of multi-information source bilinear convolutional neural network constructed in this paper is shown in Fig. 4.

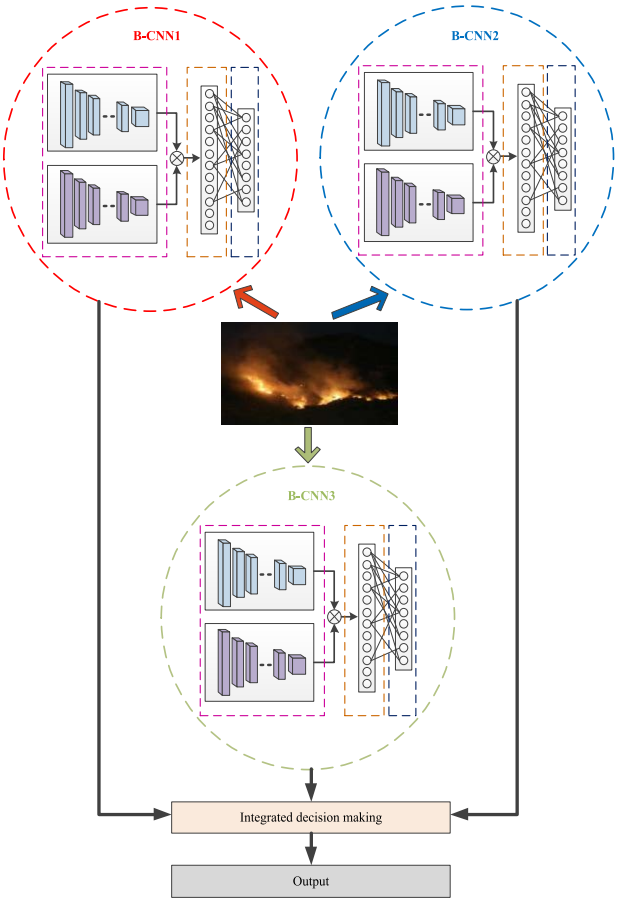


FIGURE 4. Schematic diagram of multi-information source B-CNN system.

As shown in Fig. 4, multiple B-CNN networks are used to process the images collected from different angles and positions. Different images are substituted into the B-CNN network as input, and each network identified the images according to their different positions and appearance information. Then, different recognition results are integrated and the final results are output.

IV. PROBLEMS AND SOLUTIONS

A. PROBLEM DESCRIPTION

According to the multi-information source B-CNN system shown in Fig. 4, in the model training stage, a certain number of flame images are input into the B-CNN respectively. Firstly, feature extraction and pooling calculation are completed respectively. Then, the bilinear feature after vectorization is normalized to obtain the image feature vector. Finally, the model is trained by the image feature vector and the defined image classification label.

A trained multi-information source B-CNN system is used for flame image recognition. Several flame images collected at different positions underground are respectively input into the trained B-CNN, and then two-channel feature extraction and pooling calculation are respectively completed. The image feature vector is obtained by the normalized operation, which are then used for classification calculation by the position of B-CNN. Finally, the final classification results are got by the integrated decision-making for several classification results, so as to determine whether the fire occurs.

According to the above training and classification methods of multi-information source bilinear convolutional neural network system, there are mainly two problems as follows:

(1) **Problem 1:** Parameter optimization in the B-CNN training. The Softmax loss function is used in B-CNN, and its output represents the probability that the input vector belongs to each category. For n samples, the training set is defined as: $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, in which $\{x_1, x_2, \ldots, x_n\}$ represents the input vector, and $\{y_1, y_2, \ldots, y_n\}$ represents the training label. So, when the input sample is $\boldsymbol{x}_i$, its m estimated probabilities can be expressed as

$$p_\theta(\mathbf{x}_i) = \begin{bmatrix} p(y_i = 1 \,|\mathbf{x}_i;\, \mathbf{W}) \\ p(y_i = 2 \,|\mathbf{x}_i;\, \mathbf{W}) \\ \vdots \\ p(y_i = k \,|\mathbf{x}_i;\, \mathbf{W}) \end{bmatrix} = \frac{1}{\sum_{j=1}^{m} e^{\mathbf{W}_j^T \mathbf{x}_i}} \begin{bmatrix} e^{\mathbf{W}_1^T \mathbf{x}_i} \\ e^{\mathbf{W}_2^T \mathbf{x}_i} \\ \vdots \\ e^{\mathbf{W}_m^T \mathbf{x}_i} \end{bmatrix} \tag{8}$$

where $\boldsymbol{Wm}$ represents parameters of the network model; $1/\sum_{j=1}^{m} e^{\mathbf{W}_j^T \mathbf{x}_i}$ represents normalization of output vectors.

The Softmax loss function can be defined by

$$Loss_{\text{softmax}} = -\frac{1}{n}\left[\sum_{i=1}^{n} \ln \frac{e^{\mathbf{W}_j^T \mathbf{x}_i}}{\sum_{j=1}^{m} e^{\mathbf{W}_j^T \mathbf{x}_i}}\right] \tag{9}$$

In the process of model training, the gradient descent method is generally used to optimize the Softmax loss function, and updating formula of the parameter is as follows:

$$\begin{cases} \mathbf{W}^{t+1} = \mathbf{W}^t - \lambda^t \dfrac{\partial Loss_{\text{softmax}}^t}{\partial \mathbf{W}^t} \\ \theta^{t+1} = \theta^t - \lambda^t \sum\limits_{i=1}^{n} \dfrac{\partial Loss_{\text{softmax}}^t}{\partial \mathbf{x}_i^t} \cdot \dfrac{\partial \mathbf{x}_i^t}{\partial \theta^t} \end{cases} \tag{10}$$

where t represents the number of iterations; λ represents the momentum factor.

However, when the model parameters $\boldsymbol{W}$ and θ are updated by (10), the gradient descent method is adopted for learning, so the whole search solution space cannot be traversed and the local minimum is easily trapped, thus reducing the fitting accuracy and generalization ability of the model.

(2) **Problem 2:** Computational complexity of the B-CNN model training. It can be seen that when a certain number of images are used to train two-channel CNN in B-CNN, two-channel feature extraction, pooling calculation and other operations have brought high computational complexity. When the trained model is used for flame image recognition, the flame images collected at each position need to be input into the B-CNN respectively, and the operations such as two-channel feature extraction and pooling calculation are carried out respectively, which will greatly increase the

computational complexity of flame recognition. Such high computational complexity will limit the application of multi-information source B-CNN system in real-time identification of underground flame images.

B. SOLVING METHODS

In order to solve the above two problems, this paper intends to adopt the following solutions.

1) OPTIMIZATION OF MODEL PARAMETERS BASED ON AN IMPROVED GOA ALGORITHM

The bionic principle of GOA is to map the small range movement behavior of larva to the local development of short step size, and map the large range movement behavior of adult to the global exploration of long step size, so as to conduct optimization in a similar way of "stepping collaborative." The mathematical model of GOA is expressed by

$$x_i^{t+1} = S_i^t + G_i^t + A_i^t \tag{11}$$

where t represents the current number of iterations; x_i represents location of the ith grasshopper; S_i represents the interaction between grasshoppers; G_i represents the force of gravity; A_i represents the force of wind.

The interaction S_i^t between grasshoppers is defined by

$$\begin{cases} S_i^t = \sum\limits_{\substack{j=1 \\ j\neq i}}^{popsize} s(d_{ij}^t) \cdot \dfrac{x_j^t - x_i^t}{d_{ij}^t} \\ d_{ij}^t = \left| x_j^t - x_i^t \right| \end{cases} \tag{12}$$

where popsize represents the population size; d_{ij}^t represents the distance between grasshoppers; $s(d_{ij}^t)$ represents the influence function of the interaction force between grasshoppers and other individuals in the population, defined by

$$s(d_{ij}^t) = f \cdot e^{\left(-\frac{d_{ij}^t}{l}\right)} - e^{(-d_{ij}^t)} \tag{13}$$

where f and l represent the attraction intensity and attraction scale, respectively.

The force of gravity G_i and the force of wind A_i are respectively defined by

$$\begin{cases} G_i^t = -\mathrm{g} \cdot e_g \\ A_i^t = u \cdot e_w \end{cases} \tag{14}$$

where g represents the gravitational constant; u represents the constant drift factor; e_g and e_w represent the unit vector of the force of gravity and the force of wind on the grasshopper.

Substitute S_i^t, G_i^t and A_i^t into (11), and get the updated formula for grasshopper location is as follows:

$$X_i^{t+1} = \sum_{\substack{j=1 \\ j\neq i}}^{popsize} s\left(\left|x_j^t - x_i^t\right|\right) \cdot \frac{x_j^t - x_i^t}{\left|x_j^t - x_i^t\right|} - \mathrm{g} \cdot e_g + u \cdot e_w \tag{15}$$

According to (15), after several iterations, grasshopper populations will gather in local comfort zones and lose population diversity. In this regard, the following changes to the updating formula in (15) are made,

$$X_i^d = c \cdot \left(\sum_{\substack{j=1 \\ j\neq i}}^{popsize} c \cdot \frac{ub_d - lb_d}{2} \cdot s\left(\left|x_j^d - x_i^d\right|\right) \cdot \frac{x_j - x_i}{|x_j - x_i|} \right) + T_d \tag{16}$$

where X_i^d represents the position of the ith grasshopper in the d-dimension on the $t+1$ iteration; x_i^d and x_j^d represents the position of the ith and jth grasshopper in the d-th dimension on the t-th iteration, respectively; ub_d and lb_d represent upper and lower bounds of the d-dimensional position; $\widehat{T}_d$ represents the position of the current optimal grasshopper individual in the d-dimension; c is the linear decreasing coefficient to avoid excessive grasshopper population aggregation and thus reduce the probability of the algorithm falling into local optimal, which is defined as follows:

$$c = c_{\max} - t \cdot \frac{c_{\max} - c_{\min}}{t_{\max}} \tag{17}$$

where $c_{\max}$ and $c_{\min}$ represent the maximum and minimum of c; $t_{\max}$ represents the maximum number of iteration.

In the traditional GOA, the initial solution of grasshopper population is randomly generated. However, although this method can improve the diversity of population, the over-disorderly random solution will reduce the convergence speed of the algorithm. In this paper, an initial solution generation method based on sine mapping is adopted. The sine mapping is a classical chaotic mapping system [30], and its calculation formula is as follows:

$$x(k+1) = 0.25 \cdot a \cdot \sin(\pi x(k)) \tag{18}$$

where $x(k) \in [0,1]$; a is the control parameter of the sine mapping.

Thus, the method of generating the initial solution in the improved GOA (IGOA) is as follows:

Step 1: The initial positions of all individuals in the grasshopper population are randomly generated by

$$X_i^d(0) = lb_d + (ub_d - lb_d) \cdot \mathrm{rand}(1, d) \tag{19}$$

Step 2: The normalization operation for the initial position $X_i^d(0)$ is as follows:

$$\tilde{X}_i^d(0) = \frac{X_i^d(0) - lb_d}{ub_d - lb_d} \tag{20}$$

Step 3: The sine chaotic mapping for the $\tilde{X}_i^d(0)$ is performed as follows:

$$\tilde{X}_i^d(1) = 0.25 \cdot a \cdot \sin(\pi \tilde{X}_i^d(0)) \tag{21}$$

where $\tilde{X}_i^d(1)$ represents the population position after the sine chaos mapping.

Step 4: Inverse normalization for $\tilde{X}_i^d(1)$ and obtain new ordered initial locations of grasshopper populations, as follows:

$$X_i^d(1) = lb_d + \tilde{X}_i^d(1) \cdot (ub_d - lb_d) \tag{22}$$

In addition, it can be seen from (16) that although the location updating method of grasshopper population has been improved to improve the diversity of the population, it still pays too much attention to the search of individual local location, which makes the solution process easily fall into local optimal. In this regard, in our IGOA algorithm, on the basis of (16), the location update mode of grasshopper population is further improved, as follows:

$$X_i^d = \begin{cases} c \cdot \left(\sum_{\substack{j=1 \\ j \neq i}}^{popsize} c \cdot \frac{ub_d - lb_d}{2} \cdot s\left(\left|x_j^d - x_i^d\right|\right) \cdot \frac{x_j - x_i}{|x_j - x_i|} \right) \\ +T_d, \text{rand}(1) \geq p_s \\ c \cdot \left(\sum_{\substack{j=1 \\ j \neq i}}^{popsize} c \cdot \frac{ub_d - lb_d}{2} \cdot s\left(\left|x_j^d - x_i^d\right|\right) \cdot \frac{x_j - x_i}{|x_j - x_i|} \right) \\ +\frac{(T_d + W_d)}{2}, \text{rand}(1) < p_s \end{cases} \tag{23}$$

where $\boldsymbol{W}_d$ represents the position of the worst grasshopper individual in the d dimension; p_s represents the select probability, normally taken by $p_s \in [0,0.3]$.

According to (23), the worst solution is considered with a small probability in the updating process of individual positions in the grasshopper population, which can jump out of the local optimal region to a certain extent, and improve the diversity of the population in a more comprehensive way.

2) INPUT PREPROCESSING BASED ON IMAGE FEATURE EXTRACTION

First, feature extraction is carried out on the collected images, and 19 feature vectors are extracted under the three types of color feature, shape feature and texture feature of the flame image. Then, these feature vectors are input into the B-CNN, and the two-channel feature extraction and pooling calculation are completed respectively. Finally, the model is trained and identified by the normalized image feature vector. In this way, the extracted digital image features are directly input into the B-CNN by prior knowledge, which can greatly reduce the computational complexity brought by image input. At the same time, the digital image features include color feature, shape feature and texture feature, which can basically reflect the flame feature information of the image.

The 19 feature vectors are extracted as follows:

(1) Proportion of RGB components

The R, G and B channels are used to calculate the proportion of RGB components, which are taken as three feature vectors of flame image recognition. The calculation is as follows:

$$\begin{cases} W_R = \dfrac{\sum\limits_{(i,j) \in fire\_region} R(i,j)}{\sum\limits_{(i,j) \in fire\_region} (R(i,j) + G(i,j) + B(i,j))} \\ W_G = \dfrac{\sum\limits_{(i,j) \in fire\_region} G(i,j)}{\sum\limits_{(i,j) \in fire\_region} (R(i,j) + G(i,j) + B(i,j))} \\ W_B = \dfrac{\sum\limits_{(i,j) \in fire\_region} B(i,j)}{\sum\limits_{(i,j) \in fire\_region} (R(i,j) + G(i,j) + B(i,j))} \end{cases} \tag{24}$$

where W_R, W_G and W_B represents the proportion of R, G and B components, respectively; *fire_region* represents area of flame after image segmentation; $R(i,j)$, $G(i,j)$ and $B(i,j)$ represent pixel values of R, G and B channels in RGB space.

(2) Moment characteristics of color [31]

The first moment μ_i and second moment σ_i of flame pixels in the flame region are used to represent the color information of the flame, which are defined as follows:

$$\begin{cases} \mu_i = \frac{1}{N} \sum\limits_{j=1}^{N} p_{i,j} \\ \sigma_i = \sqrt{\frac{1}{N} \sum\limits_{j=1}^{N} (p_{i,j} - \mu_i)^2} \end{cases} \tag{25}$$

where $p_{i,j}$ represents the probability of occurrence of pixel value j in color image channel i; N represents the total number of pixels in the image.

(3) Circularity characteristic [32], [33]

The circularity characteristic is defined by

$$C = \frac{L^2}{4\pi \cdot S} \tag{26}$$

where S represents the area of the region where the flame pixel is located; L is the perimeter of the area where the flame pixel is located.

(4) Rectangularity characteristic [32], [33]

The rectangularity characteristic is defined by

$$R = \frac{S}{S_R} \tag{27}$$

where S_R represents the area of the smallest rectangle which contains the object.

(5) Barycentric height coefficient [32], [33]

The barycentric height coefficient of the flame is defined by

$$R_C = \frac{H_C}{H} \tag{28}$$

where H_C represents height of center of mass of the flame pixel region; H represents total height of the flame pixel region.

(6) Texture feature

The texture feature of flame burning is an important criterion to distinguish flame from non-flame object. The gray co-occurrence matrix of flame pixel region is firstly calculated, and then the texture feature of image is calculated from the gray co-occurrence matrix [34], [35]. Define $f(x, y) = i$ and $f(x + \Delta x, y + \Delta y) = j$ as the grayscale of two related pixels, in which $i, j = 0,1,2,\ldots L$, L represents the grey scale; x and y are the coordinates of pixels, and Δx and Δy are the space. Thus, the gray co-occurrence matrix of two pixels are defined by

$$P(i,j) = \{[(x,y),(x+\Delta x, y+\Delta y)]\,|f(x,y) = i, f(x+\Delta x, y+\Delta y) = j\} \tag{29}$$

Based on the gray co-occurrence matrix defined in (29), 11 texture features of flame image are calculated as follows:

① Angular second moment (ASM)

$$ASM = \sum_{i=1}^{L}\sum_{i=1}^{L}\{P(i,j)\}^2 \tag{30}$$

② Contrast

$$Contrast = \sum_{\substack{n=0 \\ |i-j|=n}}^{L-1} n^2\{\sum_{i=1}^{L}\sum_{j=1}^{L} P(i,j)\} \tag{31}$$

③ Correlation

$$Correlation = \frac{\sum_{i=1}^{L}\sum_{j=1}^{L}(ij)P(i,j) - \mu_x\mu_y}{\sigma_x\sigma_y} \tag{32}$$

④ Sum of squares (SS)

$$SS = \sum_{i=1}^{L}\sum_{j=1}^{L}(i-\mu)^2 P(i,j) \tag{33}$$

⑤ Sum average (SA)

$$SA = \sum_{i=2}^{2L} i \cdot P_{x+y}(i) \tag{34}$$

⑥ Inverse difference moment (IDM)

$$IDM = \sum_{i=1}^{L}\sum_{j=1}^{L}\frac{1}{1+(i-j)^2}P(i,j) \tag{35}$$

⑦ Sum entropy (SE)

$$SE = -\sum_{i=2}^{2L} P_{x+y}(i)\log\{P_{x+y}(i)\} \tag{36}$$

⑧ Sum variance (SV)

$$SV = \sum_{i=2}^{2L}(i-SE)^2 P_{x+y}(i) \tag{37}$$

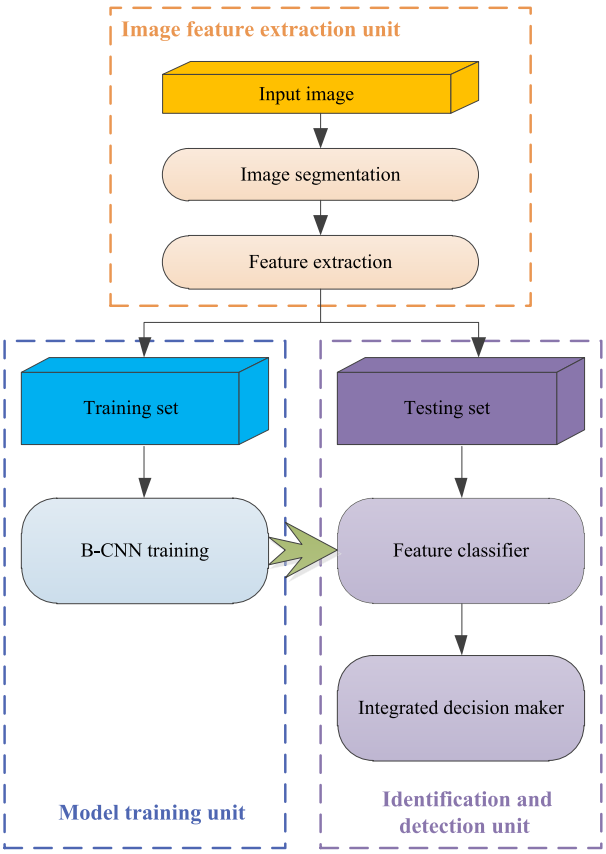


FIGURE 5. Flame image recognition model based on MISBCNN.

⑨ Entropy

$$Entropy = -\sum_{i=1}^{L}\sum_{j=1}^{L} P(i,j)\log\{P(i,j)\} \tag{38}$$

⑩ Difference variance (DV)

$$DV = \text{Variance of} P_{x-y} \tag{39}$$

Difference entropy (DE)

$$DE = \sum_{i=0}^{L-1} P_{x-y}(i)\log\{P_{x-y}(i)\} \tag{40}$$

V. FLAME IMAGE RECOGNITION MODEL

The proposed flame image recognition model based on multi-information source bilinear convolutional neural network (MISBCNN) consists of image feature extraction unit, model training unit and, recognition and detection unit, as shown in Fig. 5.

The image feature extraction unit is responsible for the preprocessing and feature extraction of the input image, including image segmentation and feature extraction. Image segmentation is to segment the target region after image preprocessing; feature extraction is to extract the feature vector of the image of the target region. The model training

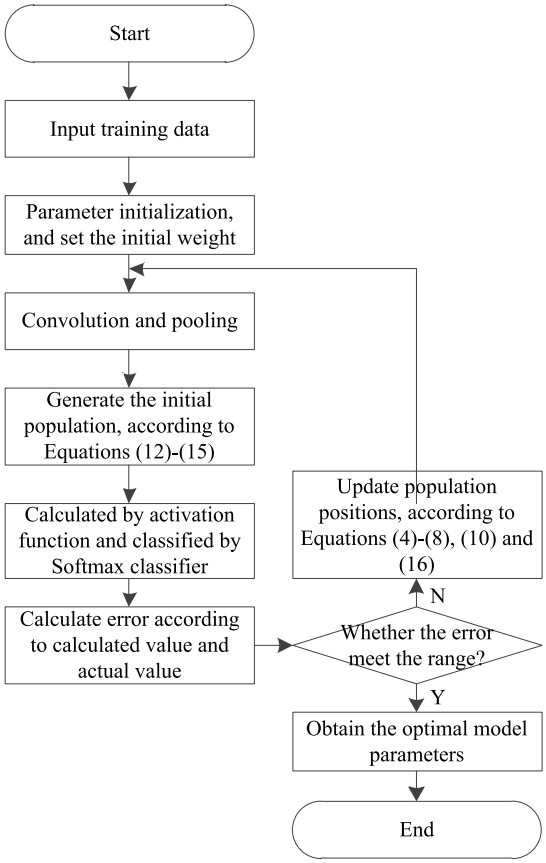


FIGURE 6. Main steps of the B-CNN model training based on IGOA optimization.

unit is to develop the optimal B-CNN network by the training set. The recognition and detection unit takes the trained B-CNN network as the feature classifier, in which, the test set is input into the multi-information source B-CNN, and makes the integrated decision to obtain the final recognition result. The model training unit and recognition and detection unit are described as follows.

A. MODELTRAINING UNIT

After the training set data is determined, B-CNN network is trained. For the training data, the extracted 19 image feature vectors are used as input vectors, and the corresponding classification labels are used as output vectors. Then convolution and pooling operations are carried out. After activation function calculation, the Softmax classifier is used for classification. The calculated output is compared with the actual output to get the prediction error and adjust the weight reversely. When the prediction error meets the constraint conditions, the training ends and the optimal model parameters are determined. The training process of the model training unit is shown in Fig. 6.

According to Fig. 6, the whole training of B-CNN model involves convolution and pooling operations, calculation of activation function, classification calculation of Softmax classifier, etc. Then, the network structure of B-CNN model is shown in Fig. 7. Among them, network *A* and network *B* are based on VGG-NET [36], each of them containing 13 convolutional layers and 4 maximum pooling layers.

In Fig. 7, *Conv* represents the convolution layer, and *Max-pool* represents the maximum pooling layer. In convolution operation, the convolution function is to realize a feature mapping; the pooling function in the pooling operation is to combine the features of all locations into a general feature. After convolution and pooling, all feature graphs are connected and mapped to a one-dimensional vector, and then the image features are classified by Softmax classifier.

B. RECOGNITION AND DETECTION UNIT

The trained B-CNN model is used for flame image recognition at different positions and angles. At the same time, a group of flame images are collected by image acquisition devices in different locations, and they are input into the trained B-CNN models respectively. The integrated decision is made on the recognition results of different models, and the final conclusion is obtained. The structure diagram of the recognition and detection unit based on multi-information source B-CNN model is shown in Fig. 8.

According to Fig. 8, the workflow of recognition and detection unit is shown in Fig. 9.

In the stage of multi-information source integration identification, this paper adopts the "voting" strategy, and adopts odd number of information sources to produce odd number of classification results. If the classification results obtained are inconsistent, the result with more votes is the final result.

VI. SIMULATION EXPERIMENTS

A. PERFORMANCE VERIFICATION OF IGOA

Firstly, the advanced optimization performance of IGOA algorithm proposed in this paper is verified. The following four Benchmark functions are used for simulation experiments.

(1) Ackley

$$f_1(x) = -20\exp(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2}) - \exp(\frac{1}{n}\sum_{i=1}^{n}\cos 2\pi x_i) + 20 + e \quad (41)$$

where $x \in$[-5,5].

(2) Sphere

$$f_2(x) = \sum_{i=1}^{n} x_i^2 \quad (42)$$

where $x \in$[-2.5,2.5].

(3) Griewank

$$f_3(x) = \sum_{i=1}^{n}\frac{x_i^2}{4000} - \prod_{i=1}^{n}\cos(x_i/\sqrt{i}) + 1 \quad (43)$$

where $x \in$[-250,250].

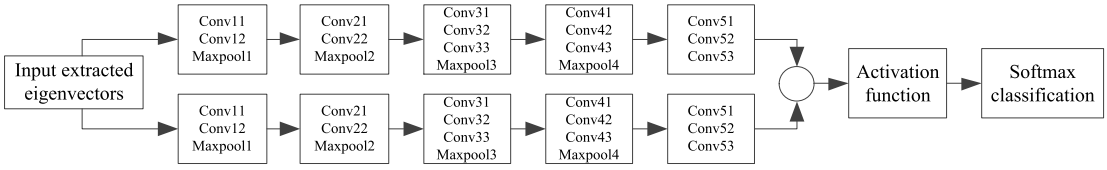


FIGURE 7. Network structure of the B-CNN model.

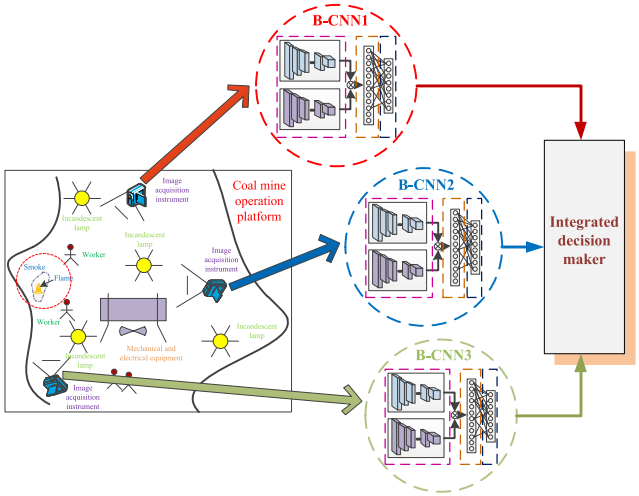


FIGURE 8. Structure diagram of recognition and detection unit based on multi-information source B-CNN model.

TABLE 1. Adjustable parameters in each algorithm.

Method	Parameters setting
IFS	Small search steps is random given in [15,25] search radius is random given in [0.1,0.5]
IHS	The harmony number is random given in [20,40], the retention probability is random given in [0.5,0.95], the memory disturbance probability is random given in [0.05,0.5], the minimum bandwidth is 0.0005, the maximum bandwidth is 1.
GOA	f *is* random given in [0.1,0.9], l *is* random given in [1,2], c_{max}=1, c_{min}=0.00004.
IGOA	a is random given in [1,5], f is random given in [0.1,0.9], l *is* random given in [1,2], c_{max}=1, c_{min}=0.00004, ps is random given in [0,0.3].

(4) Bohachevsky

$$f_4(x) = x_1^2 + 2x_2^2 - 0.3\cos(3\pi x_1)\cos(4\pi x_2) + 0.3 \quad (44)$$

where $x_1, x_2 \in$[-120,120].

Six algorithms of FBH [37], IFS [38], IHS [39], LVCMFOA [40], GOA and IGOA are respectively used for optimization of f_1-f_4, in which n is 4, the population size is 30, the number of iteration is 1000. The values of adjustable parameters in each algorithm are shown in Table 1.

Each algorithm is repeatedly run by 20 times, and Best value, Worst value, Average value and Std. value of them are recorded in Table 2.

As can be seen from Table 2, for Ackley function (f_1), IHS and LVCMFOA have poor optimization performance,

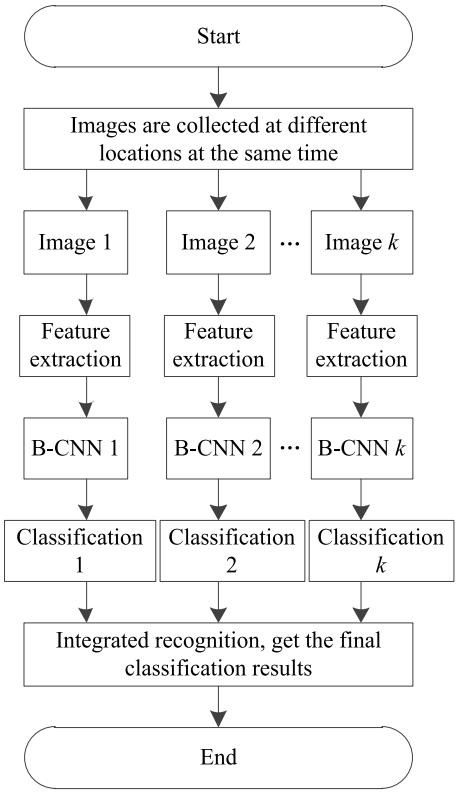


FIGURE 9. Workflow of the recognition and detection unit.

while FBH and IGOA have better optimization performance among FBH, IFS, GOA and IGOA, among which IGOA has superior performance; in addition, as standard deviation std. reflects the stability of optimization results of the optimization algorithm in multiple runs, it can be seen that IFS and GOA have poor stability, while IGOA has significantly better stability than other methods. For Sphere function (f_2), compared with FBH, IFS, IHS and LVCMFOA, both GOA and IGOA have obvious optimization accuracy; however, of the two methods, GOA has poor stability performance, and IGOA has advantages in both optimization accuracy and stability. For Griewank function (f_3), among the four methods of FBH, IFS, IHS and LVCMF- OA, although IFS and IHS have better optimal values, they have unsatisfactory worst values, resulting in poor overall optimization performance; although the optimal values for GOA and IGOA are not the best in these methods, the worst values for multiple runs are combined with Std.

TABLE 2. Comparison of several algorithms for different Benchmark functions.

	Algorithm	FBH	IFS	IHS	LVCMFOA	GOA	IGOA
f_1	**Best**	4.66e-05	2.81e-05	0.02	3.60	3.17e-07	3.17e-07
	Worst	1.20e-03	1.87	0.54	3.60	1.84	2.85e-06
	Average	4.13e-04	0.30	0.24	3.60	0.18	1.31e-06
	Std.	4.64e-04	0.63	0.16	9.74e-05	0.58	8.35e-07
f_2	**Best**	5.72e-05	1.74e-04	0.07	0.62	4.21e-14	4.87e-15
	Worst	1.26e-01	0.18	0.28	3.60	1.11e-12	9.50e-14
	Average	1.49e-02	0.04	0.20	2.01	2.10e-13	4.60e-14
	Std.	3.93e-02	0.06	0.07	1.06	3.20e-13	3.14e-14
f_3	**Best**	2.81	0.37	0.93	2.70	1.00	1.00
	Worst	7.63	9.25	6.94	3.60	1.00	1.00
	Average	4.74	4.42	4.15	3.42	1.00	1.00
	Std.	1.70	2.95	1.97	0.37	5.81e-13	4.88e-13
f_4	**Best**	2.81e-05	0.21	0.38	3.60	7.24e-12	1.56e-12
	Worst	1.87	66.67	131.70	3.60	1.43e-10	8.49e-11
	Average	0.30	9.84	39.84	3.60	4.63e-11	1.48e-11
	Std.	0.63	20.63	38.95	1.09e-04	4.00e-11	2.55e-11

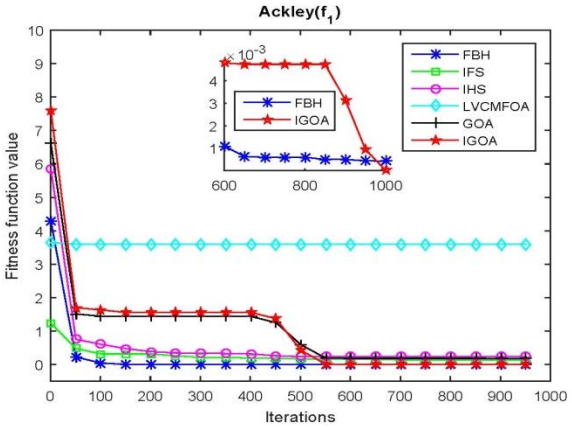


FIGURE 10. Comparison of the average fitness curves for Ackley function (f_1).

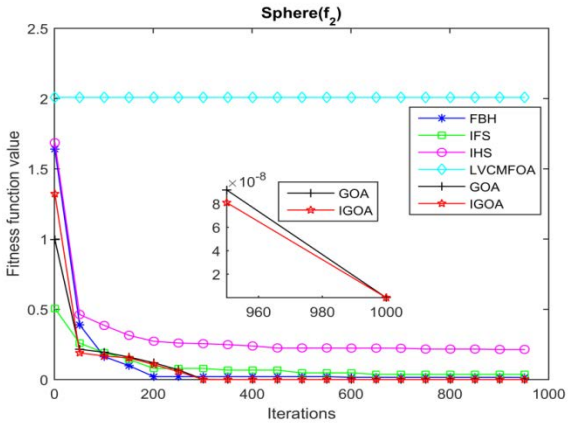


FIGURE 11. Comparison of the average fitness curves for Sphere function (f_2).

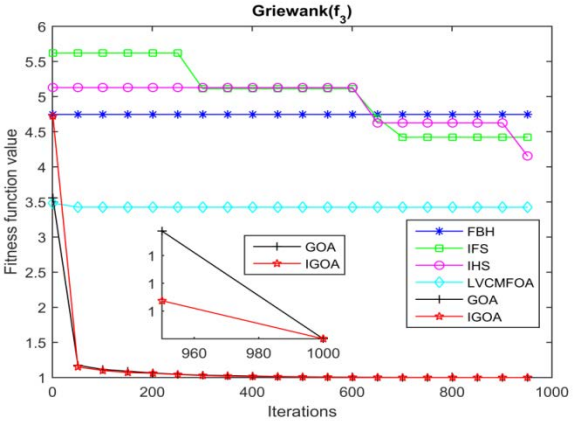


FIGURE 12. Comparison of the average fitness curves for Griewank function (f_3).

Value, their comprehensive optimization ability is the best. For Bohachevsky function (f_4), compared with FBH, IFS, IHS, and LVCMFOA, both GOA and IGOA have obvious advantages in the optimal value, mean value, and Std. value, indicating that their performance of multiple optimization is the best; by comparing GOA and IGOA, it can be seen that IGOA has more advantages in optimization accuracy and stability.

In order to better illustrate the advantages of IGOA, the average fitness curves of multiple optimizations for each algorithm are presented in Figs. 10-13.

According to Figs. 10-13, for f_1-f_4, compared with FBH, IFS, IHS, LVCMFOA and GOA, IGOA has significantly better convergence speed and accuracy in the process of repeated optimization. The calculation accuracy distribution of several algorithms under multiple optimization runs is shown in Figs. 14-19.

As can be seen from Figs. 14-19, compared with FBH, IFS, IHS, LVCMFOA and GOA, the IGOA algorithm proposed in this paper has obvious advantages in the calculation accuracy distribution of repeated runs. Therefore, the optimal selection of two parameters $\boldsymbol{W}$ and θ by the IGOA algorithm should achieve satisfactory results.

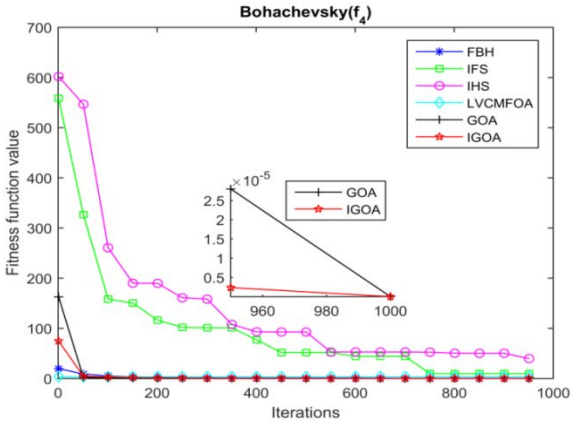


FIGURE 13. Comparison of the average fitness curves for Bohachevsky function (f_4).

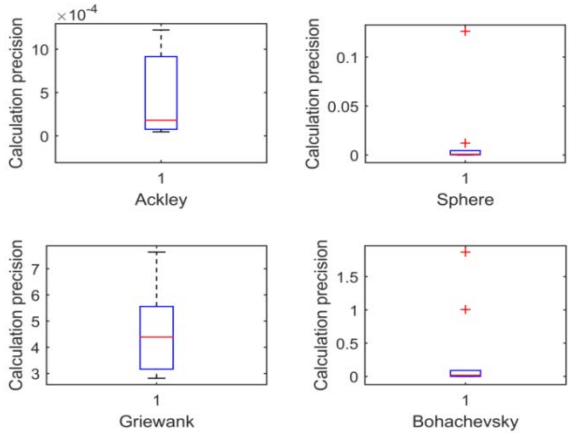


FIGURE 14. Calculation accuracy distribution of FBH.

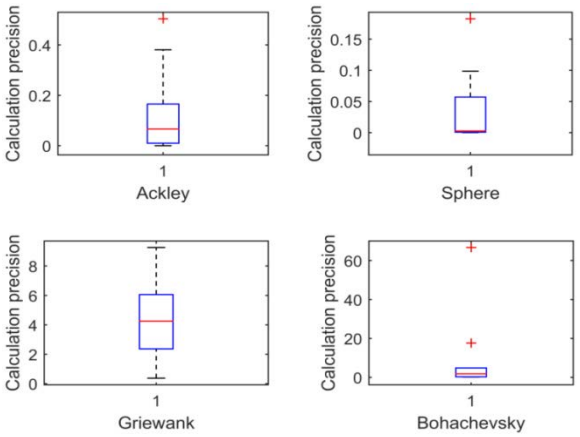


FIGURE 15. Calculation accuracy distribution of IFS.

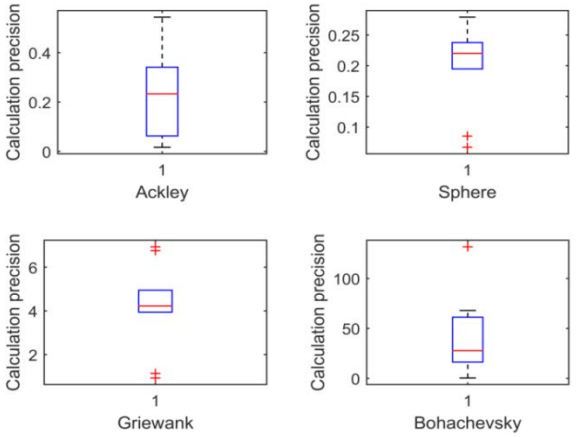


FIGURE 16. Calculation accuracy distribution of IHS.

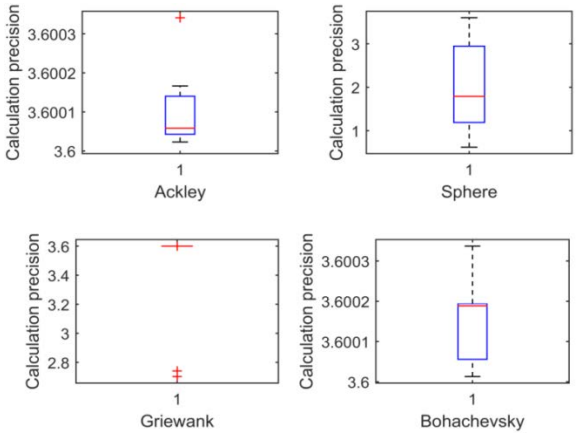


FIGURE 17. Calculation accuracy distribution of LVCMFOA.

B. PERFORMANCE VERIFICATION OF MULTI-SOURCE B-CNN MODEL OPTIMIZED BY IGOA

339 images are used for simulation experiments, which are from bowfire public data set and Bilkent university laboratory in http://signal.ee.bilkent.edu.tr/VisiFire/Demo/SmokeClips. 300 of them are used for training model and the rest 39 are used for model testing. In the training set, there are 200 flame pictures and 100 non-flame pictures; in the test set, there are 26 flame images and 13 non-flame images. It should be noted that non-flame images are suspected images of sunset clouds, sunset, lights, lanterns, red objects, etc.

Classification accuracy rate, classification precision rate and classification recall rate are used to evaluate the recognition performance of the system.

(1) Classification accuracy. The proportion of the number of results with correct model classification to the total number of results, which is defined as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \qquad (45)$$

where, TP represents the true example, which means that the model correctly classifies the positive category samples into positive categories; TN represents true negative example, which means that the model correctly classifies negative category samples into negative category; FP represents for false positive example, which means that the model incorrectly classifies negative category samples into positive categories; FN represents for false negative example, which

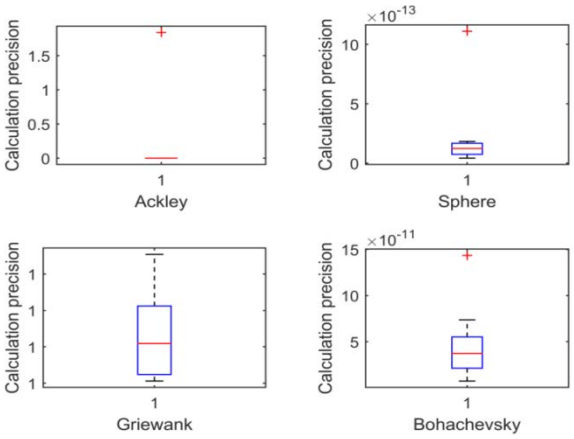


FIGURE 18. Calculation accuracy distribution of GOA.

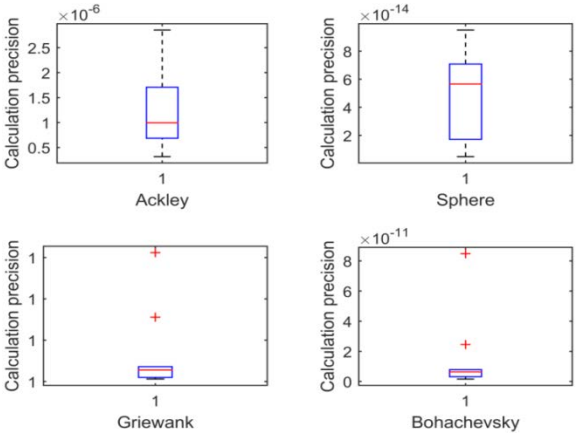


FIGURE 19. Calculation accuracy distribution of IGOA.

means that the model correctly classifies negative category samples into negative category.

(2) Classification precision rate.

The proportion of samples identified as positive categories that are indeed positive categories, which is defined as follows:

$$Precision = \frac{TP}{TP + FP} \times 100\% \tag{46}$$

(3) Classification recall rate. The proportion of positive category samples correctly identified as positive category, which is defined as follows:

$$Recall = \frac{TP}{TP + FN} \times 100\% \tag{47}$$

1) FEATURE EXTRACTION

Firstly, the target region of the image is marked, and then the target region is segmented. Finally, the feature extraction of the segmented target region is carried out as training set and test set. This paper adopts an RGB-HSV-YUV mixed color model for image segmentation, where RGB color rules [41]

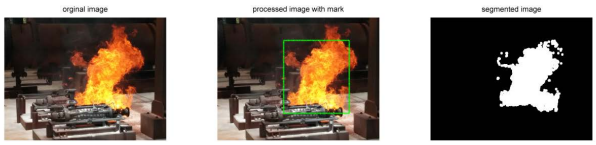


FIGURE 20. Flame region segmentation effect based on the RGB-HSV-YUV mixed color model.

are as follows:

$$\begin{cases} \text{Rule 1} : R > R_T \\ \text{Rule 2} : S > (225 - R) \times S_T / R_T \\ \text{Rule 3} : R \geq G > B \end{cases} \tag{48}$$

where R, G and B represent the red, green, and blue components of each pixel in the image; S represents color saturation of the image; R_T represents a threshold, there is $R_T \in[115,135]$; S_T represents a threshold, there is $S_T \in[45,60]$. Only pixels that meet these three rules are identified as flame pixels of RGB color models.

HSV color rules [42] are as follows:

$$\begin{cases} \text{Rule 1} : 0.02 < H < 0.3 \\ \text{Rule 2} : 0.2 < S < 1.0 \\ \text{Rule 3} : 0.98 < V < 1.0 \end{cases} \tag{49}$$

where H, S and V represent the hue, saturation and brightness of the image.

YUV color rules [42], [43] are as follows:

$$\begin{cases} \text{Rule 1} : \ Y \geq 80, Y > Y_{mean} \\ \text{Rule 2} : \ |U - 128| \leq 60, U < U_{mean} \\ \text{Rule 3} : \ |V - 128| \leq 40, V > V_{mean} \end{cases} \tag{50}$$

where Y represents brightness of a pixels, and Y_{mean} represents the mean value of pixels on the Y channel, in the YUV color space; U and V represents chroma of a pixel, and U_{mean} represents the mean value of pixels on the U channel, and V_{mean} represents the mean value of pixels on the V channel.

Thus, the RGB-HSV-YUV mixed color model established for flame area identification is as follows:

$$F_{color} = F_{RGB} \cup F_{HSV} \cup F_{YUV} \tag{51}$$

where F_{color} represents pixel in the mixed color model; F_{RGB} represents pixel meets RGB color rules; F_{HSV} represents pixel meets HSV color rules; F_{YUV} represents pixel meets YUV color rules.

The flame region segmentation effect of a pipeline fire image based on the RGB-HSV-YUV mixed color model [44] is shown in Fig. 20.

After the segmentation of the flame region, 19 image feature vectors can be extracted to train the B-CNN model, according to (24)-(40).

TABLE 3. Parameter settings of convolution layer and maximum pooling layer in network A and network B.

Layer name	Size of convolution kernel	Number of output channel
Conv11	3×3	64
Conv12	3×3	64
Conv21	3×3	128
Conv22	3×3	128
Conv31	3×3	256
Conv32	3×3	256
Conv33	3×3	256
Conv41	3×3	512
Conv42	3×3	512
Conv43	3×3	512
Conv51	3×3	512
Conv52	3×3	512
Conv53	3×3	512
Maxpool1	2×2	64
Maxpool2	2×2	128
Maxpool3	2×2	256
Maxpool4	2×2	512

TABLE 4. Recognition results of training samples by B-CNN model.

	Accuracy	Precision	Recall
Training samples	96.99%	94.76%	99.50%

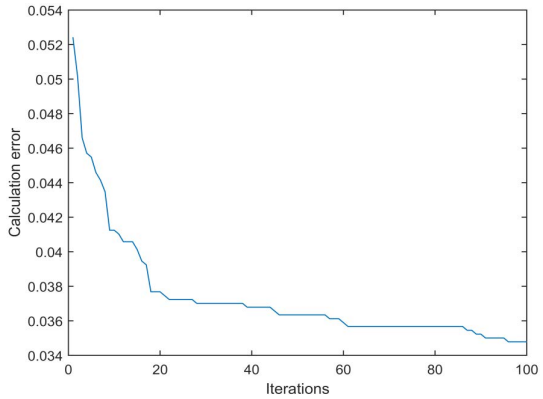


FIGURE 21. The curve of training error with the number of iterations.

2) TRAINING OF THE B-CNN MODEL

In the B-CNN model training, parameters of each layer are set in Table 3.

According to the training process of B-CNN model (as shown in Fig. 6), the model is trained by the feature vectors of 300 flame images, and then 300 groups of feature vectors are input into the trained B-CNN model, and the recognition results of training samples are obtained, as shown in Table 4.

According to Fig. 6, in the model training, weights need to be reversely updated according to the errors of actual values and calculated values, until the calculated errors meet the requirements. Set the number of iterations in the update process to 100. Then, the curve of calculation error changing with the number of iterations is shown in Fig. 21.

As can be seen from Fig. 21, with the increase of iterations, the training error gradually decreases. Combined with the recognition results of training samples by B-CNN model in Table 3, it can be shown that the designed B-CNN network structure has satisfactory recognition performance.

TABLE 5. Classification accuracy rate of different models (accuracy/%).

Model	Source 1	Source 2	Source 3	Integrated decision making
Multi-information source ESN	83.46	76.92	66.67	84.62
Multi-information source PNN	79.49	76.92	69.23	82.05
Multi-information source SVM	77.88	71.79	59.62	79.49
Multi-information source CNN	84.62	82.05	79.49	88.40
Multi-information source B-CNN	94.87	94.87	88.92	94.87

TABLE 6. Classification precision rate of different models (precision/%).

Model	Source 1	Source 2	Source 3	Integrated decision making
Multi-information source ESN	83.47	73.53	57.14	83.47
Multi-information source PNN	79.17	70.80	66.67	81.48
Multi-information source SVM	77.11	70.80	66.67	81.48
Multi-information source CNN	87.25	81.48	79.17	90.48
Multi-information source B-CNN	90.48	87.47	90.48	92.86

3) TESTING OF MULTI-INFORMATION SOURCE B-CNN MODEL

In order to simulate the results of flame image acquisition by device with multiple positions and angles, the test images are scaled, expanded and rotated respectively. In this paper, flame image processing of three information sources is adopted, and 39 test images are shrunk by 0.5 times, expanded by 1.5 times and rotated by 30° respectively. The effects are shown in Fig. 22.

After 39 groups of test images are shrunk by 0.5 times, expanded by 1.5 times and rotated by 30°, respectively, a total of 39×3 groups of test images are obtained. According to (17)-(33), 19 image feature vectors are extracted and input into three B-CNN information sources. In order to verify the effectiveness of the proposed method, four models including multi-information source ESN, multi-information source PNN, multi-information source SVM and multi-information source CNN are used for comparison. The results of the test samples are shown in Tables 4-6.

According to the results in Tables 4-6, compared with ESN, PNN, SVM and CNN, B-CNN has better performance in classification Accuracy rate, classification Precision rate and classification Recall rate, as it comprehensively considers information from different angles and positions of images. From the perspective of three information sources, information source 1 is to enlarge the image by 1.5 times, information source 2 is to reduce the image by 0.5 times,

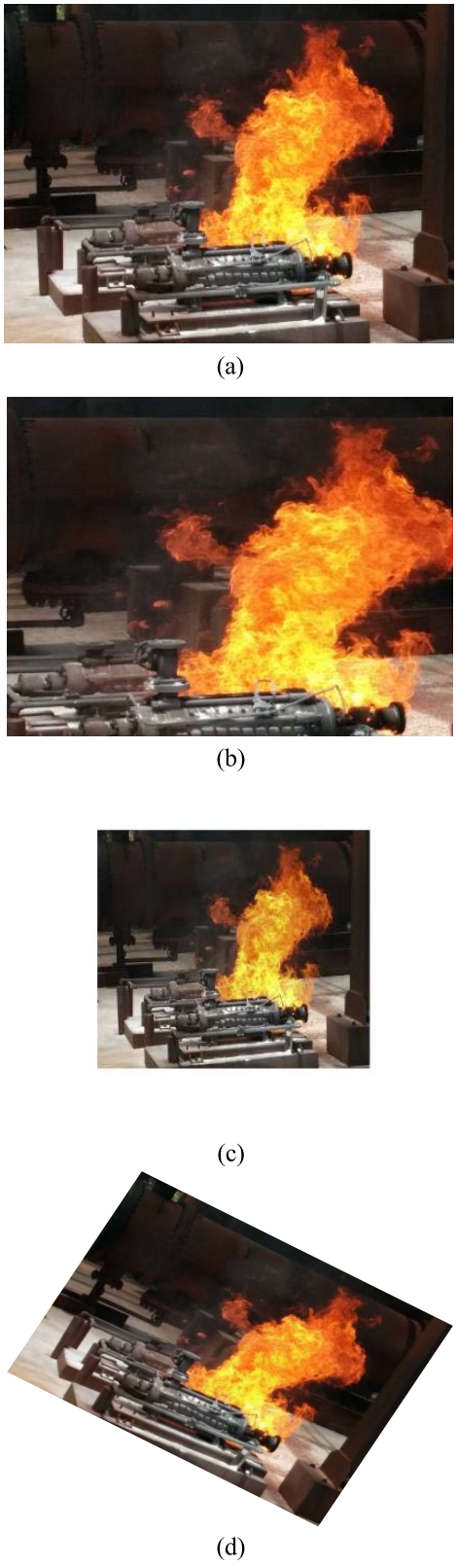


FIGURE 22. Schematic diagram of multi-source image information collection ((a) Original image; (b) Information source 1 (expanded by 1.5 times); (c) Information source 2 (shrunk by 0.5 times); (d) Information source 3 (rotated by 30°)).

TABLE 7. Classification Recall rate of different models (recall/%).

Model	Source 1	Source 2	Source 3	Integrated decision making
Multi-information source ESN	84.62	80.77	84.29	88.36
Multi-information source PNN	84.62	84.62	84.29	88.36
Multi-information source SVM	80.77	84.29	67.31	84.62
Multi-information source CNN	88.36	84.62	84.62	93.37
Multi-information source B-CNN	92.31	94.33	90.48	96.15

and information source 3 is to rotate the image by 30°; according to Tables 3-5, the result of the comparative analysis of five kinds of models, all classification effects on the information source 1 is good, and the worst classification effects on the information source 3, illustrating that when the image produces rotating, the negative influence on the classifier, on the contrary, when to expand the image processing, image characteristics of flame area are more apparent, which improves the performance of the classifier. From perspective of the integrated decision-making, although the classification effect of single information source is not ideal in some cases, the deficiency of it can be made up by information fusion processing and decision-making of multiple information sources. As can be seen from the results in Tables 3-5, although the classification accuracy rate of SVM by information source 3 is only 59.62%, the classification precision rate of ESN by information source 3 is only 57.14%, and the classification recall rate of SVM by information source 3 is only 59.62%, the final classification ability has been greatly improved through the integration decision of multiple information sources.

VII. CONCLUSION

In this paper, an underground coal mine fire identification method based on multi-information source B-CNN is proposed. By setting up image acquisition devices in different locations and using multi-source image information processing method to make integrated decision, the problem that single information source processing method is greatly affected by various external interference factors can be solved. In order to verify this conclusion, the test images were scaled, expanded and rotated in the simulation experiment to simulate the flame images collected at different positions and angles, and the flame recognition was carried out by using three B-CNN information sources. Accuracy, Precision and Recall were used to verify the effectiveness of the proposed method. In addition, An initial solution generation method based on sine mapping and a bad solution acceptance method with a certain probability were adopted to improve the performance of GOA; in the training process of B-CNN model, it can improve the classification performance of

B-CNN model by using it to optimize the two parameters W and θ in Softmax loss function. Moreover, in view of the B-CNN in training and integration recognition stage had the problem of high computational complexity, this article used the image characteristics in advance, that was, several image feature vectors under different types were extracted as input of B-CNN, which could reduce the computational complexity of the input image. In the stage of multi-information source integration recognition, the ''voting'' strategy was adopted, that was, an odd number of information sources produced an odd number of classification results, and the result with more votes was the final result. However, in many cases, there might be many channels of multiple information sources, or the number of information sources is even; so in the face of these problems, how to adopt a more rigorous integration decision method is a research focus of our future work.

REFERENCES

[1] J. Zhao, Z. Zhang, S. Han, C. Qu, Z. Yuan, and D. Zhang, ''SVM based forest fire detection using static and dynamic features,'' *Comput. Sci. Inf. Syst.*, vol. 8, no. 3, pp. 821–841, 2011.

[2] X. Huang and L. Du, ''Fire detection and recognition optimization based on virtual reality video image,'' *IEEE Access*, vol. 8, pp. 77951–77961, 2020.

[3] J. Baek, T. J. Alhindi, Y.-S. Jeong, M. K. Jeong, S. Seo, J. Kang, J. Choi, and H. Chung, ''Real-time fire detection algorithm based on support vector machine with dynamic time warping kernel function,'' *Fire Technol.*, vol. 57, no. 6, pp. 2929–2953, Nov. 2021, doi: 10.1007/s10694-020-01062-1.

[4] Y. Gao, L. Xie, Z. Zhang, and Q. Fan, ''Twin support vector machine based on improved artificial fish swarm algorithm with application to flame recognition,'' *Int. J. Speech Technol.*, vol. 50, no. 8, pp. 2312–2327, Aug. 2020.

[5] Y. Bian, M. Yang, X. Fan, and Y. Liu, ''A fire detection algorithm based on tchebichef moment invariants and PSO-SVM,'' *Algorithms*, vol. 11, no. 6, p. 79, May 2018.

[6] Y. T. Chen, W. H. Xu, J. W. Zuo, and Y. Kai, ''The fire recognition algorithm using dynamic feature fusion and IV-SVM classifier,'' *Cluster Comput.*, vol. 10, pp. 1–11, Mar. 2018.

[7] M. Kamran and N. M. Shahani, ''Decision support system for the prediction of mine fire levels in underground coal mining using machine learning approaches,'' *Mining, Metall. Explor.*, vol. 39, no. 2, pp. 591–601, Apr. 2022.

[8] M. J. Barros-Daza, K. D. Luxbacher, B. Y. Lattimer, and J. L. Hodges, ''Real time mine fire classification to support firefighter decision making,'' *Fire Technol.*, Jan. 2022, doi: 10.1007/s10694-022-01215-4.

[9] H. Zhao, ''Research on fire detection in coal mine based on GA-improved wavelet neural networks,'' *Adv. Mater. Res.*, vols. 490–495, pp. 1636–1639, Mar. 2012.

[10] Y. L. Jiang, ''Research of multi-sensor information fusion fire detection system,'' *Adv. Mater. Res.*, vols. 860–863, pp. 2745–2749, Dec. 2013.

[11] Z. Wen, L. Xie, H. Feng, and Y. Tan, ''Infrared flame detection based on a self-organizing TS-type fuzzy neural network,'' *Neurocomputing*, vol. 337, pp. 67–79, Apr. 2019.

[12] A. Andrew, A. Zakaria, S. M. Saad, and A. M. Shakaff, ''Multi-stage feature selection based intelligent classifier for classification of incipient stage fire in building,'' *Sensors*, vol. 16, no. 1, p. 31, Jan. 2016.

[13] Z. Wen, L. Xie, H. Feng, and Y. Tan, ''Robust fusion algorithm based on RBF neural network with TS fuzzy model and its application to infrared flame detection problem,'' *Appl. Soft Comput.*, vol. 76, pp. 251–264, Mar. 2019.

[14] L. Wu, L. Chen, and X. Hao, ''Multi-sensor data fusion algorithm for indoor fire early warning based on BP neural network,'' *Information*, vol. 12, no. 2, p. 59, Jan. 2021.

[15] J. Sharma, O. C. Granmo, and M. Goodwin, *Deep CNN-ELM Hybrid Models for Fire Detection in Images*. Cham, Switzerland: Springer, 2018.

[16] K. Muhammad, J. Ahmad, and S. W. Baik, ''Early fire detection using convolutional neural networks during surveillance for effective disaster management,'' *Neurocomputing*, vol. 288, pp. 30–42, May 2018.

[17] F. Saeed, A. Paul, P. Karthigaikumar, and A. Nayyar, ''Convolutional neural network based early fire detection,'' *Multimedia Tools Appl.*, vol. 79, nos. 13–14, pp. 9083–9099, Apr. 2020.

[18] G. H. de Almeida Pereira, A. M. Fusioka, B. T. Nassu, and R. Minetto, ''Active fire detection in Landsat-8 imagery: A large-scale dataset and a deep-learning study,'' 2021, *arXiv:2101.03409*.

[19] A. Ayala, B. Fernandes, F. Cruz, D. Macêdo, A. L. I. Oliveira, and C. Zanchettin, ''KutralNet: A portable deep learning model for fire recognition,'' 2020, *arXiv:2008.06866*.

[20] J. Zhu and H. Ren, ''Fire picture recognition based on deep learning and particle algorithm,'' *J. Ambient Intell. Humanized Comput.*, Apr. 2021, doi: 10.1007/s12652-021-03150-8.

[21] X. Sun, L. Sun, and Y. Huang, ''Forest fire smoke recognition based on convolutional neural network,'' *J. Forestry Res.*, vol. 2020, pp. 1–7, Oct. 2020.

[22] Y. Cao, F. Yang, Q. Tang, and X. Lu, ''An attention enhanced bidirectional LSTM for early forest fire smoke recognition,'' *IEEE Access*, vol. 7, pp. 154732–154742, 2019.

[23] A. S. Pundir and B. Raman, ''Dual deep learning model for image based smoke detection,'' *Fire Technol.*, vol. 55, no. 6, pp. 2419–2442, Nov. 2019.

[24] Z. Yang, L. Bu, T. Wang, P. Yuan, and O. Jineng, ''Indoor video flame detection based on lightweight convolutional neural network,'' *Pattern Recognit. Image Anal.*, vol. 30, no. 3, pp. 551–564, Jul. 2020.

[25] T.-Y. Lin, A. RoyChowdhury, and S. Maji, ''Bilinear CNNs for fine-grained visual recognition,'' 2015, *arXiv:1504.07889*.

[26] F. Zhou, S. Kong, C. C. Fowlkes, T. Chen, and B. Lei, ''Fine-grained facial expression analysis using dimensional emotion model,'' *Neurocomputing*, vol. 392, pp. 38–49, Jun. 2020.

[27] X. Dong, H. Zhou, and J. Dong, ''Texture classification using pair-wise difference pooling-based bilinear convolutional neural networks,'' *IEEE Trans. Image Process.*, vol. 29, pp. 8776–8790, 2020.

[28] Z. Tang, E. Tian, Y. Wang, L. Wang, and T. Yang, ''Nondestructive defect detection in castings by using spatial attention bilinear convolutional neural network,'' *IEEE Trans. Ind. Informat.*, vol. 17, no. 1, pp. 82–89, Jan. 2021.

[29] S. Saremi, S. Mirjalili, and A. Lewis, ''Grasshopper optimisation algorithm: Theory and application,'' *Adv. Eng. Softw.*, vol. 105, pp. 30–47, Mar. 2017.

[30] F. B. Demir, T. Tuncer, and A. F. Kocamaz, ''A chaotic optimization method based on logistic-sine map for numerical function optimization,'' *Neural Comput. Appl.*, vol. 32, no. 17, pp. 14227–14239, Sep. 2020.

[31] C. E. Prema, S. S. Vinsley, and S. Suresh, ''Efficient flame detection based on static and dynamic texture analysis in forest fire detection,'' *Fire Technol.*, vol. 54, no. 1, pp. 255–288, Jan. 2018.

[32] K. Li, S. Li, R. Liu, J. Wang, and D. Liu, ''Flame detection based on video,'' *Comput. Sci. Appl.*, vol. 6, no. 3, pp. 171–177, 2016.

[33] X. Wu, Y. Yan, J. Du, S. Gao, and Y. Liu, ''Flame detection based on fusion of multiple features,'' *CAAI Trans. Intell. Syst.*, vol. 10, no. 2, pp. 240–247, 2015.

[34] R. M. Haralick, K. Shanmugam, and I. Dinstein, ''Textural features for image classification,'' *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973.

[35] J. F. Vargas, M. A. Ferrer, C. M. Travieso, and J. B. Alonso, ''Off-line signature verification based on grey level information using texture features,'' *Pattern Recognit.*, vol. 44, no. 2, pp. 375–385, 2011.

[36] S. Liu and W. Deng, ''Very deep convolutional neural network based image classification using small training sample size,'' in *Proc. 3rd IAPR Asian Conf. Pattern Recognit. (ACPR)*, Kuala Lumpur, Malaysia, Nov. 2015, pp. 730–734.

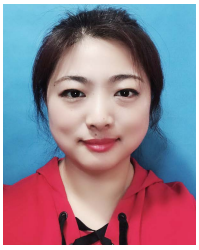
[37] K. Li, X.-W. Gao, H.-B. Zhou, and Y. Han, ''Fault diagnosis for down-hole conditions of sucker rod pumping systems based on the FBH–SC method,'' *Petroleum Sci.*, vol. 12, no. 1, pp. 135–147, Mar. 2015.

[38] Z. Tian and S. Li, ''A network traffic prediction method based on IFS algorithm optimised LSSVM,'' *J. Eng. Syst. Model. Simul.*, vol. 19, no. 4, pp. 200–213, 2017.

[39] Z. Tian, S. Li, Y. Wang, and X. Wang, ''A network traffic hybrid prediction model optimized by improved harmony search algorithm,'' *Neural Netw. world*, vol. 25, no. 6, pp. 669–686, 2015.

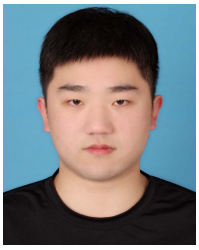
[40] Y. Han, Y. Jing, G. M. Dimirovski, and L. Zhang, ''Multi-step network traffic prediction using echo state network with a selective error compensation strategy,'' *Trans. Inst. Meas. Control*, vol. 44, no. 8, pp. 1656–1668, May 2022, doi: 10.1177/01423312211050296.

[41] T.-H. Chen, P.-H. Wu, and Y.-C. Chiou, "An early fire-detection method based on image processing," in *Proc. Int. Conf. Image Process. (ICIP)*, Singapore, Oct. 2004, pp. 1707–1710.
[42] J. Seebamrungsat, S. Praising, and P. Riyamongkol, "Fire detection in the buildings using image processing," in *Proc. 3rd ICT Int. Student Project Conf. (ICT-ISPC)*, Nakhonpathom, Thailand, Mar. 2014, pp. 95–98.
[43] A. E. Çetin, K. Dimitropoulos, B. Gouverneur, N. Grammalidis, O. Günay, Y. H. Habiboglu, B. U. Töreyin, and S. Verstockt, "Video fire detection—Review," *Digit. Signal Process.*, vol. 23, no. 6, pp. 1827–1843, 2013.
[44] H. Wu, A. Zhang, Y. Han, J. Nan, and K. Li, "Fast stochastic configuration network based on an improved sparrow search algorithm for fire flame recognition," *Knowl.-Based Syst.*, vol. 245, Jun. 2022, Art. no. 108626, doi: 10.1016/j.knosys.2022.108626.



LI ZHANG received the B.Sc. degree from the Shenyang University of Chemical Technology, in 2005, and the M.Sc. degree from Liaoning Technical University, in 2008. She is currently a Senior Engineer of the State Key Laboratory of Coal Mine Safety Technology, Shenyang Research Institute Company Ltd., China Coal Science and Industry Group. Her current research interests include design and development of mechanical and electrical safety equipment in coal mine.

YUQIN ZHU received the B.Sc. degree from the Qingdao University of Science and Technology, in 2011, and the M.Sc. degree from the Dalian University of Technology, in 2014. She is currently a Simulation Engineer of Shenyang Research Institute Company Ltd., China Coal Science and Industry Group. Her current research work focuses on numerical simulation analysis and scientific research work on coal mine robot.



HAO WU received the B.Sc. degree from the College of Mechanical & Electrical Engineering, Anhui Polytechnic University, in 2018. He is currently pursuing the M.S. degree in pattern recognition and intelligent system with Bohai University. His research interests include flame image processing, neural networks, and intelligent optimization.

KUN LI was born in Tai'an, Shandong, China, in 1983. He received the B.Sc. degree from the Shandong University of Science and Technology, in 2005, the M.Sc. degree from Liaoning Technical University, China, in 2008, and the Ph.D. degree from Northeastern University, China, in 2013.

From September 2013 to January 2021, he was an Assistant Professor with the College of Control Science and Engineering, Bohai University, China. Since March 2021, he has been a Professor at the Faculty of Electrical and Control Engineering, Liaoning Technical University. Now, he has collaboration with the State Key Laboratory of Coal Mine Safety Technology (China) on scientific research. He is the author of more than 30 articles and more than ten inventions. His research interests include complex system modeling, intelligent computing, machine learning, multi-objective optimization, swarm intelligence optimization algorithm, and their applications in petroleum, coal mine, power, communication networks, and UAV path planning.