

Forest fire image recognition based on convolutional neural network

Wang Yuanbin , Dang Langfei and Ren Jieying

Journal of Algorithms & Computational

Technology

Volume 13: 1–11

© The Author(s) 2019

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1748302619887689

journals.sagepub.com/home/act



Abstract

In order to detect fire automatically, a forest fire image recognition method based on convolutional neural networks is proposed in this paper. There are two main types of fire recognition algorithms. One is based on traditional image processing technology and the other is based on convolutional neural network technology. The former is easy to lead in false detection because of blindness and randomness in the stage of feature selection, while for the latter the unprocessed convolutional neural network is applied directly, so that the characteristics learned by the network are not accurate enough, and recognition rate may be affected. In view of these problems, conventional image processing techniques and convolutional neural networks are combined, and an adaptive pooling approach is introduced. The fire flame area can be segmented and the characteristics can be learned by this algorithm ahead. At the same time, the blindness in the traditional feature extraction process is avoided, and the learning of invalid features in the convolutional neural network is also avoided. Experiments show that the convolutional neural network method based on adaptive pooling method has better performance and has higher recognition rate.

Keywords

Forest fire recognition, deep learning, convolutional neural network, pooling model

Received 23 May 2019; Revised 25 July 2019; accepted 19 September 2019

Introduction

With the rapid development of digital camera technology and image processing technology, the flame detection method based on computer vision system has gradually replaced the traditional method and has become an important trend.^{1,2} Due to the complex background and large space of the forest fire image, certain difficulties are brought to the forest fire identification process, especially in the feature selection process, there are often some blind operations. Applying the convolutional neural network (CNN) technology to image recognition can avoid the blindness and randomness to a large extent in the feature extraction process, and theoretically extract deeper features, which can greatly improve the accuracy of flame image recognition. CNN technology has been applied to fire image recognition by many researchers.^{3,4}

Wang⁵ used the logistic regression and support vector machine to classify the images in the output classification layer of the CNN for forest fire detection. The experiments show that the accuracy of the forest fire detection algorithm based on deep learning has

exceeded than that of traditional algorithm. Chen et al.⁶ used the block detection method to preprocess the video fast on forest fire image, which greatly reduced the running time of the whole system. This method uses graphics processing unit (GPU) to accelerate the texture analysis and improves the real-time performance of the system. Fu et al.⁷ designed different network models to identify the forest fire image in the case of different forest fire backgrounds at night and during the day, and mainly analyze the performance of the network under different parameters. The experimental results show that the CNN-based forest fire flame recognition method has prominent advantages compared with the traditional image processing-based

School of Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an, China

Corresponding author:

Wang Yuanbin, School of Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an, China.

Email: wangyb998@163.com



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<http://www.creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

flame detection method. Frizzi et al.⁸ proposed a CNN method for detecting fires in video. Tested on real video sequences, the proposed method achieves better classification performance than other related traditional video forest fire detection methods. Hohberg⁹ applied a convolutional 3D (C3D) network that can extract temporal features to forest fire detection. The experimental results show that trained networks distinguish wildfires from other objects clearly.

These scholars have made certain contributions to the identification of forest fires and achieved good recognition results. However, in these studies, most of the original images were directly input to the CNN to accomplish the training. Due to the complex background and various forms of the original forest fire image, these non-flame features will be fully learned during the training process, which makes the network generalization performance very poor. Due to the above factors, conventional image processing techniques and CNN methods are combined in this paper. Firstly, the candidate flame area of fire is segmented, and then the CNN is used to complete the recognition of fire flames. In the pooling process of CNN, an adaptive pooling method is introduced.

Forest fire recognition based on CNN

Flow of the algorithm

At present, most of the applications in forest fire identification are directly applied to CNN on the original image set. Due to the complex background and a number of interference in the original image, the result of the training is not so good. Therefore, in this paper, a method is proposed to segment the candidate flame region based on the color feature, and then part of the image is sent to the CNN network for training, which can extract features more specifically and improve the recognition rate of forest fire image effectively. The algorithm flow is shown in Figure 1.

In the training phase, firstly, the binary image of the suspected flame region is segmented, and the result obtained by performing AND operation between the binary image and the original image is used as a training set, and a label is set for each image. A network model is obtained after training the CNN according to the training set. In the testing phase, similarly, the binary image of the suspected flame region is firstly segmented, and the result obtained by performing AND operation with the original image is used as a testing set. The testing set image is sent to the trained network model to obtain the recognition result.

Segmentation of suspected flame area

Color is a very important feature of fire image. In this paper, the flame pixel is segmented by the feature in RGB color space and YCbCr color space¹⁰ (the third problem of review 1). In order to analyze the characteristics of the flame pixel in RGB color space, the number of pixels is plotted on the abscissa, and the coordinate system is established with the pixel value as the ordinate, and the flame pixels on the *R* channel and the *G* channel are depicted, respectively, as shown in Figure 2.

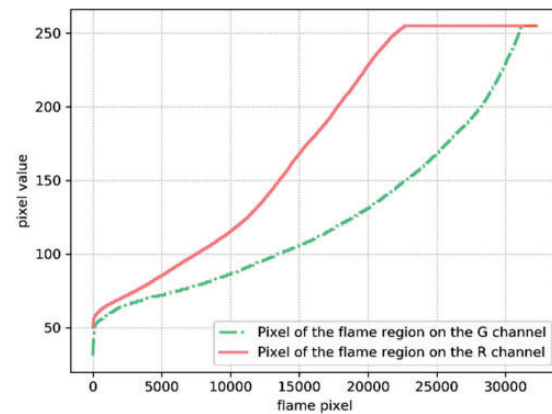


Figure 2. Distribution of flame pixel values on channels *R* and *G*.

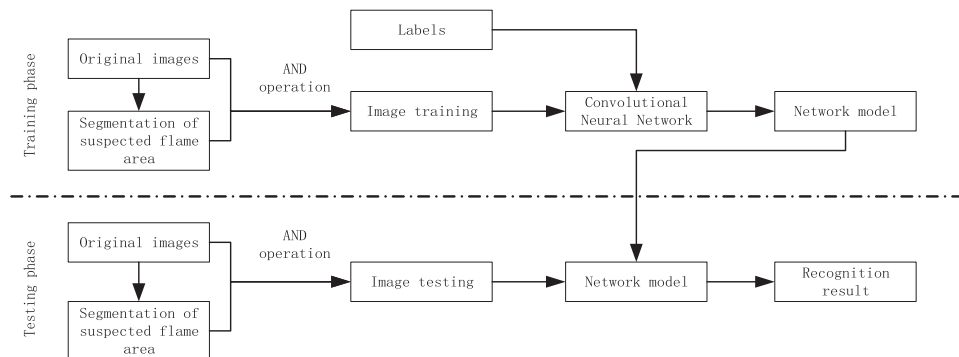


Figure 1. Flow chart of forest fire image recognition algorithm based on CNN.

From Figure 2, the rule of the flame pixel in RGB color space as shown in equation (1) can be found

$$R(i,j) > G(i,j) \quad (1)$$

where $R(i,j)$ and $G(i,j)$ represent the pixel values on channels R and G at the coordinates (i,j) , respectively.

Similarly, the characteristics of the flame pixel in YCbCr color space can be analyzed, as shown in Figure 3.

From Figure 3(a), the rule for the flame pixel in YCbCr color space as shown in equation (2) can be found

$$\begin{cases} Y(i,j) > Y_{mean} \\ Cb(i,j) < Cb_{mean} \\ Cr(i,j) > Cr_{mean} \end{cases} \quad (2)$$

where $Y(i,j), Cb(i,j), Cr(i,j)$ represent flame pixel values at the position (i,j) on Y channel, Cb channel, and Cr channel, respectively; $Y_{mean}, Cb_{mean}, Cr_{mean}$ represent the average pixel value on Y channel, Cb channel, and Cr channel, respectively.

In this paper, the flame image is analyzed pixel by pixel according to the feature information in formulas (1) and (2) to implement the segmentation of candidate flame region. The flowchart is shown in Figure 4 (the first problem of reviewer 2).

Selection and structure of the model

In this paper, a CNN model is designed based on the AlexNet to recognize flame area. Specifically, it consists of three convolution layers, three pooling layers, two fully connected layers, plus one input layer and output layer. The network is shown in Figure 5, and the network structure is shown in Table 1.

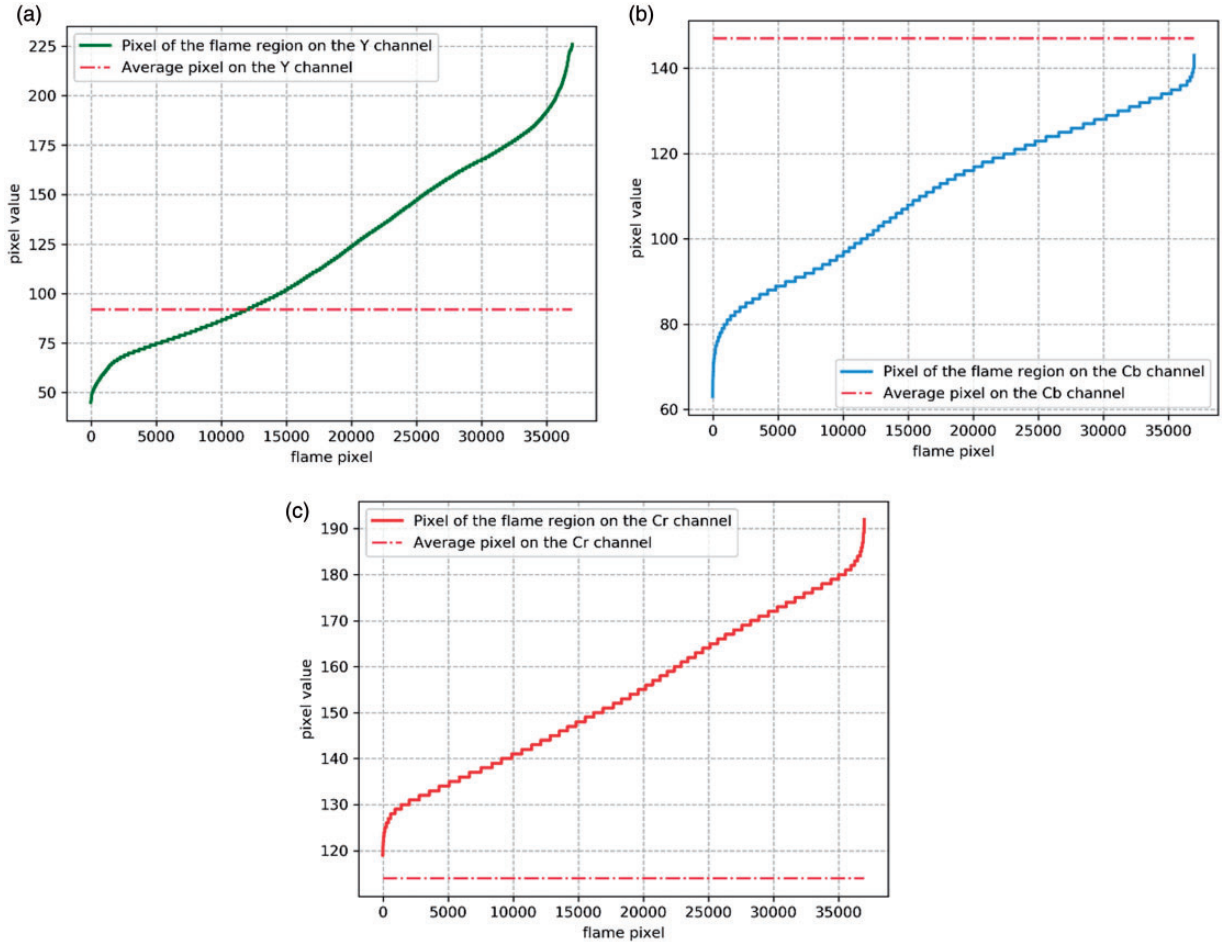


Figure 3. Characteristic distribution of flame pixels in YCbCr color space. (a) Distribution of flame pixel values on Y channels. (b) Distribution of flame pixel values on Cb channels. (c) Distribution of flame pixel values on Cr channel.

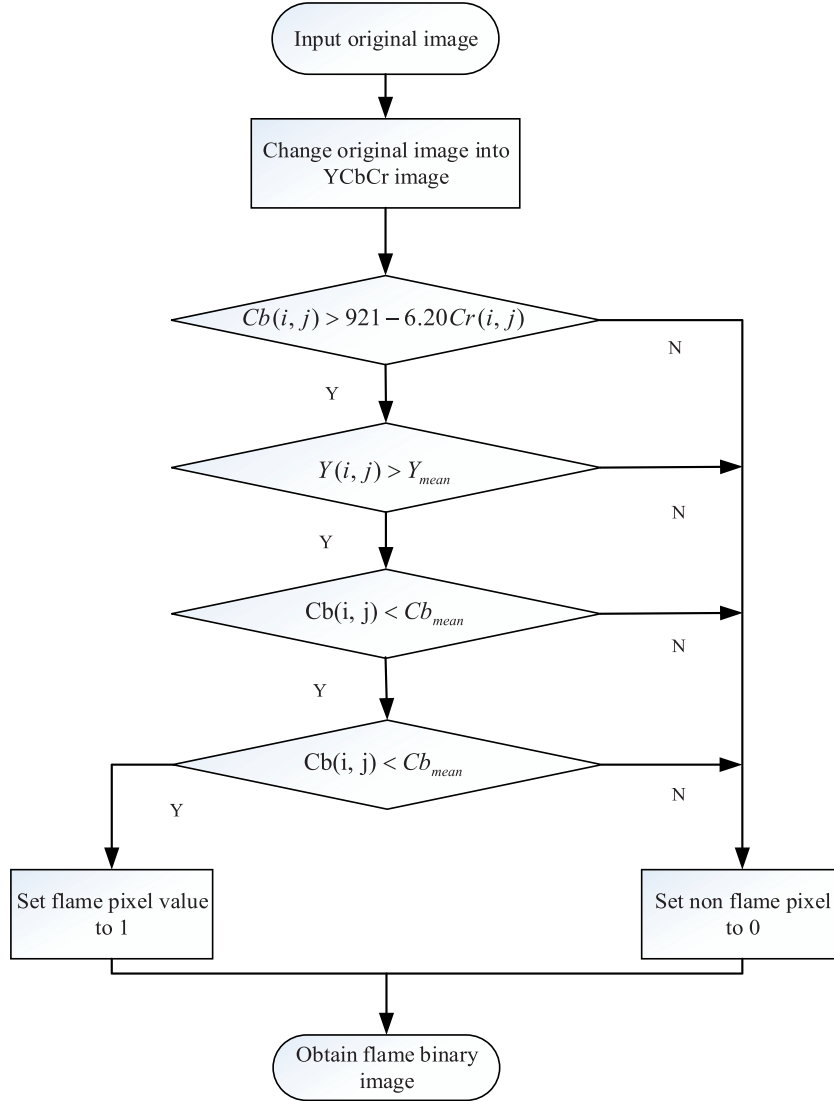


Figure 4. Flow chart of fire image segmentation based on color feature.

CNN based on adaptive pooling (the first problem of reviewer 1)

The traditional pooling methods are maximum pooling¹¹ and mean pooling,¹² and the formula is given as in equations (3) and (4)

$$P_{ij} = \max_{i=1, j=1}^s (F_{ij}) \quad (3)$$

$$P_{ij} = \frac{1}{s^2} \left(\sum_{i=1}^s \sum_{j=1}^s F_{ij} \right) \quad (4)$$

where F is the feature map, F_{ij} is the pixel value of the feature map at position (i, j) , the size of the pooled domain is $s \times s$, and the resulting pooled result is represented by P . As shown in Figures 6 and 7, the process of maximum pooling and averaging pooling is performed.

It is undeniable that these two pooling methods do have great advantage in terms of calculation. But for some extreme case, it is not very optimistic. The results of a pooled domain after maximum pooling and averaging pooling are shown in Figures 8 and 9.

It can be seen that the results based on traditional pooling cannot represent all the information of the original feature map accurately. Therefore, Liu et al.¹³ proposed an optimization algorithm based on the average pooling and maximum pooling, in which both of the two traditional pooling models are considered, and its expression is shown in equation (5)

$$P_{ij} = \frac{1}{2} \left[\frac{1}{c^2} \left(\sum_{i=1}^s \sum_{j=1}^s F_{ij} \right) + \max_{i=1, j=1}^s (F_{ij}) \right] \quad (5)$$

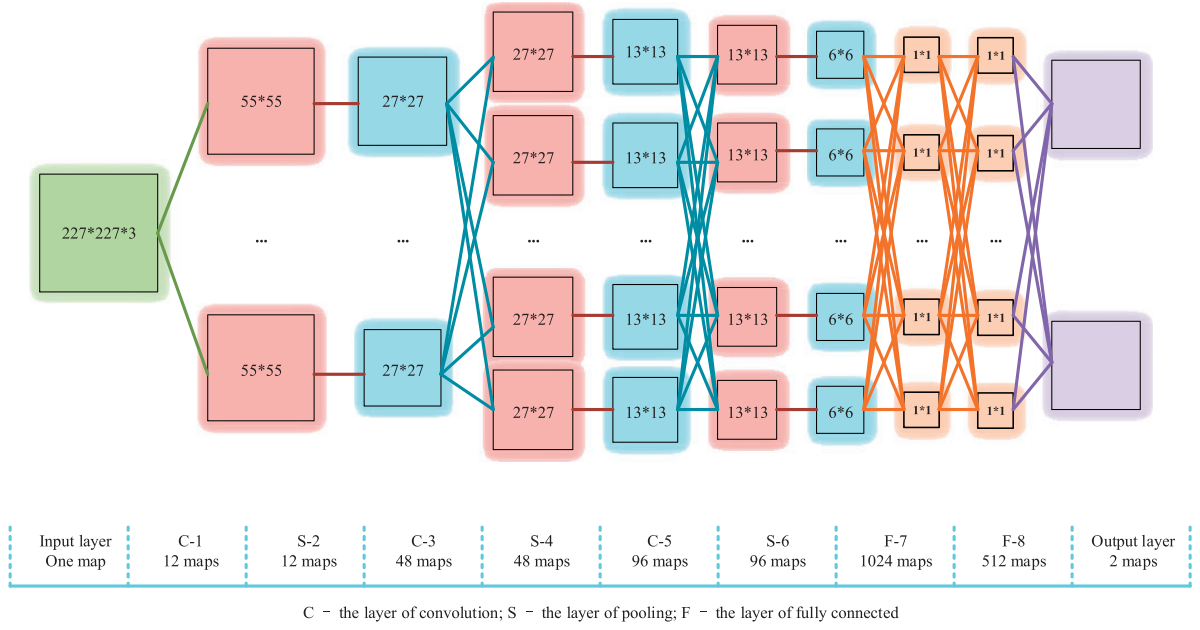


Figure 5. Convolutional neural network structure of fire recognition.

Table 1. Structure of the network.

Network number	Network type	Number of feature maps	Kernel size	Step size
Input	Input layer	3	—	—
C-1	Convolution	12	11×11	4
S-2	Pooling	12	3×3	2
C-3	Convolution	48	5×5	2
S-4	Pooling	48	3×3	2
C-5	Convolution	96	3×3	1
S-6	Pooling	96	3×3	1
F-7	Fully connected	Number of neurons: 1024		
F-8	Fully connected	Number of neurons: 512		
Output	Output layer	Number of categories: 2		

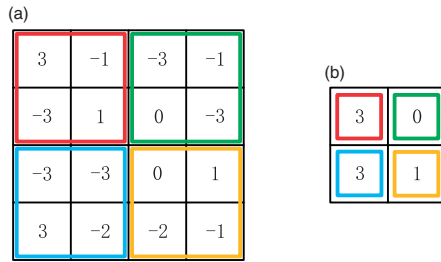


Figure 6. Max pooling. (a) 4×4 feature map. (b) Result of max pooling.

local information of the image undoubtedly. In view of this problem, an adaptive pooling based on the median pooling is proposed in this paper. Feature extraction based on the information of pooled domain can be performed dynamically. Its expression is as follows

$$P_{ij} = \frac{1}{2} \left[\frac{1}{c^2} \left(\sum_{i=1}^{s-1} \sum_{j=1}^{s-1} \omega_1 F_{ij} \right) + \max_{i=1, j=1}^s (\omega_2 F_{ij}) \right] \quad (6)$$

In this model, the feature extraction process can be optimized and the performance of the network can be improved. However, in the process of network training, various pooling domains will appear. For traditional CNN network, a fixed pooling model is adopted to extract features, which will cause loss of

In equation (6), two pooling factors ω_1 and ω_2 are introduced

$$\omega_1 = \frac{1}{1 + e^{-\frac{F_{ij}}{\text{sum} + \theta}}} \quad (7)$$

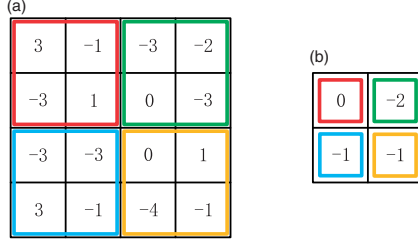


Figure 7. Average pooling. (a) 4×4 feature map. (b) Result of average pooling.

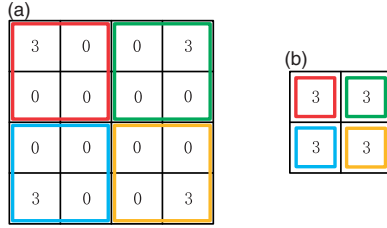


Figure 8. Extreme max pooling effect. (a) 4×4 feature map. (b) Result of max pooling.

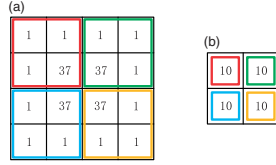


Figure 9. Extreme average pooling effect. (a) 4×4 feature map. (b) Result of max pooling.

Table 2. Impact of initial learning rate on network performance.

Initial learning rate	Epoch at convergence	Training accuracy (%)	Final loss	Testing accuracy (%)
1	14	31.5	1.9982	0
0.1	11	26	2.0328	0
0.01	75	100	0.0087	88.75
0.001	367	100	0.0071	72.08
0.0001	—	81	0.6415	54.17

$$\omega_2 = e^{-\frac{\max - \text{mean}}{\max + \theta}} \quad (8)$$

where *sum* represents the sum of all pixels in the pooled domain; *max* represents the maximum value, and *mean* represents the mean of the remaining pixels after removing the largest pixel; θ exists as a correction error, and it mainly deals with the case where the pixels in the pooled domain are all 0.

Both the mean pooling and the maximum pooling is considered in the design of the adaptive pooling model, and the pooling method can be selected dynamically. Furthermore, the loss of image information caused by the traditional pooling method in the feature extraction process is avoided.

Experiment results and analysis

The performance of CNN is affected by some hyperparameters, including initial learning rate, batch size, pooling, and so on. The performance of CNN from the above aspects is analyzed.

Impact of learning rate on network performance

The cost function $Loss(\omega)$ is a function of the weight ω , and the update of the weight is implemented by back-propagation, as in equations (9) and (10)

$$\Delta\omega = -\alpha \frac{\partial loss}{\partial \omega} \quad (9)$$

$$\omega_{i+1} = \Delta\omega + \omega_i \quad (10)$$

where α is the learning rate. The network structure is the same as in Table 1. In order to ensure the accuracy, the average value was taken from three experiments at each initial learning rate. As shown in Table 2, training period epoch during network convergence, final training accuracy, and accuracy of the verification set were analyzed.

When the initial learning rate is 1 and 0.1, the loss value explodes directly. Although the network converges quickly, the final training result is poor, and the accuracy in the verification data set is 0, and all the verification images are classified into the same category. When the initial learning rate is set too small, such as 0.0001, the network still does not converge at 2000 epoch, which shows that the convergence speed is slow. When the initial learning rate is 0.01 or 0.001, the network can converge more quickly, and the accuracy can reach 100% during training process, and has good generalization ability. Both time and performance are taken into consideration, then the initial learning rate is preferably set between 0.001 and 0.01.

Impact of batch size on network performance

In training process, batch processing technology is often used, and the size of the batch determines the optimization degree and speed of the model. Different batch sizes are used to train 500 flame and non-flame data sets on the network structure, as shown in Table 1. The data analysis of the training results is shown in Table 3.

Table 3. Training results under different batch size.

Batch size	512	256	128	64	32	8	1
Total epochs	500	500	500	500	500	500	500
Total iteration	1	2	4	7	13		
Time of one epochs	42.93	36.06	32.73	29.82	25.22		
Achieve 0.99 Accuracy at epoch	327	149	69	35	27	Cannot converge	
Time of achieve 0.99 accuracy	10,266.84	4810.26	1720.10	753.14	352.15		
Final training error (500 epochs)	0.0281	0.0084	0.0030	0.0015	0.0028		
Test score (%)	56.25	45.83	68.75	82.42	64.58		

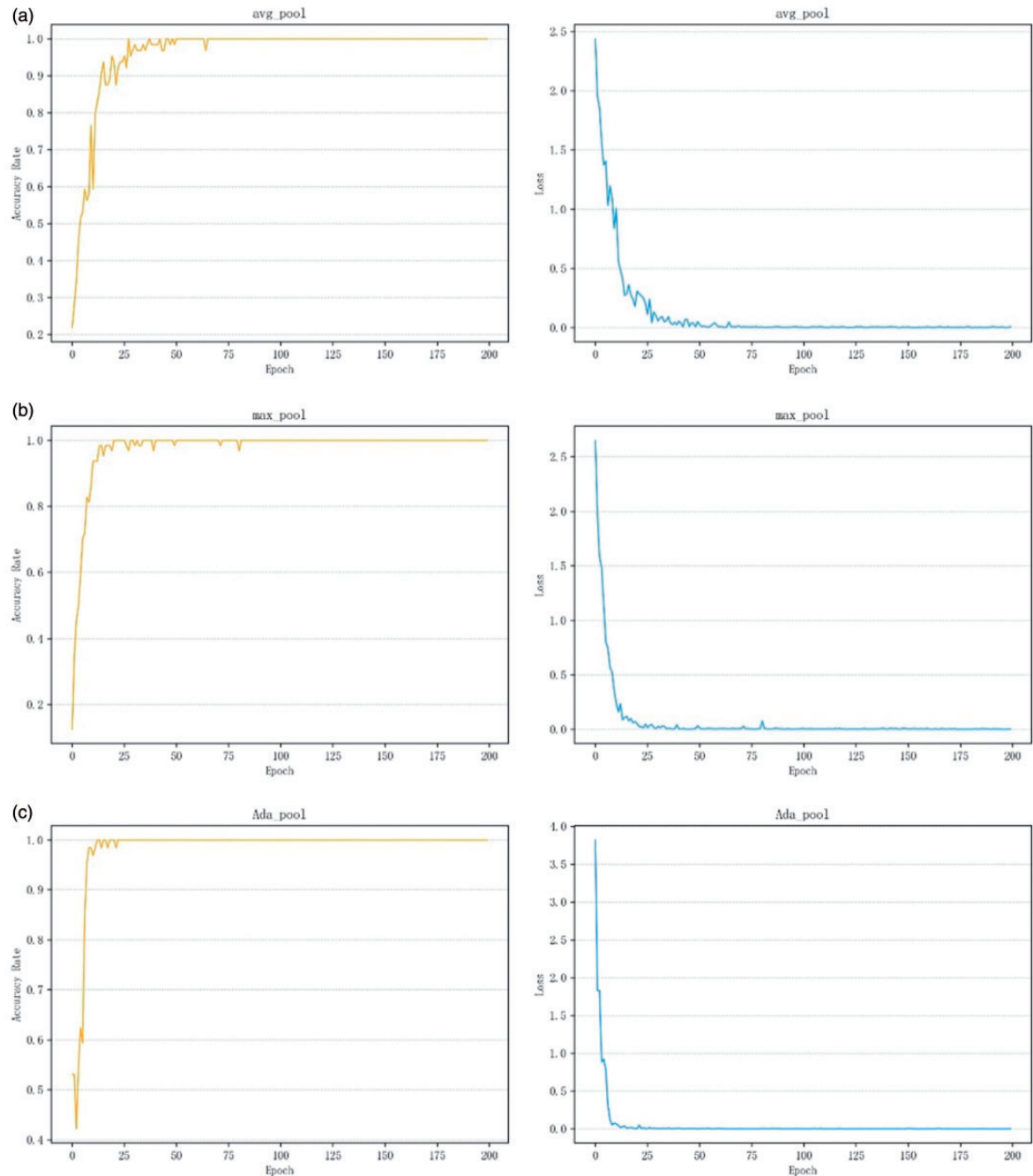
**Figure 10.** Training results in different pooling modes. (a) Average pooling. (b) Max pooling. (c) Adaptive pooling.

Table 4. Test accuracy under different pooling modes.

Pooling mode	Test accuracy (%)
Average	87.5
Max	78.13
Adaptive	93.75

From Figure 3, combining the final loss of training and the score of testing set, the batch size is selected as 64 for subsequent experiment.

Impact of pooling model on network performance

In order to verify the impact of adaptive pooling on the performance of CNN, the structure in Table 1 is adopted. The batch size is set to 64, the learning

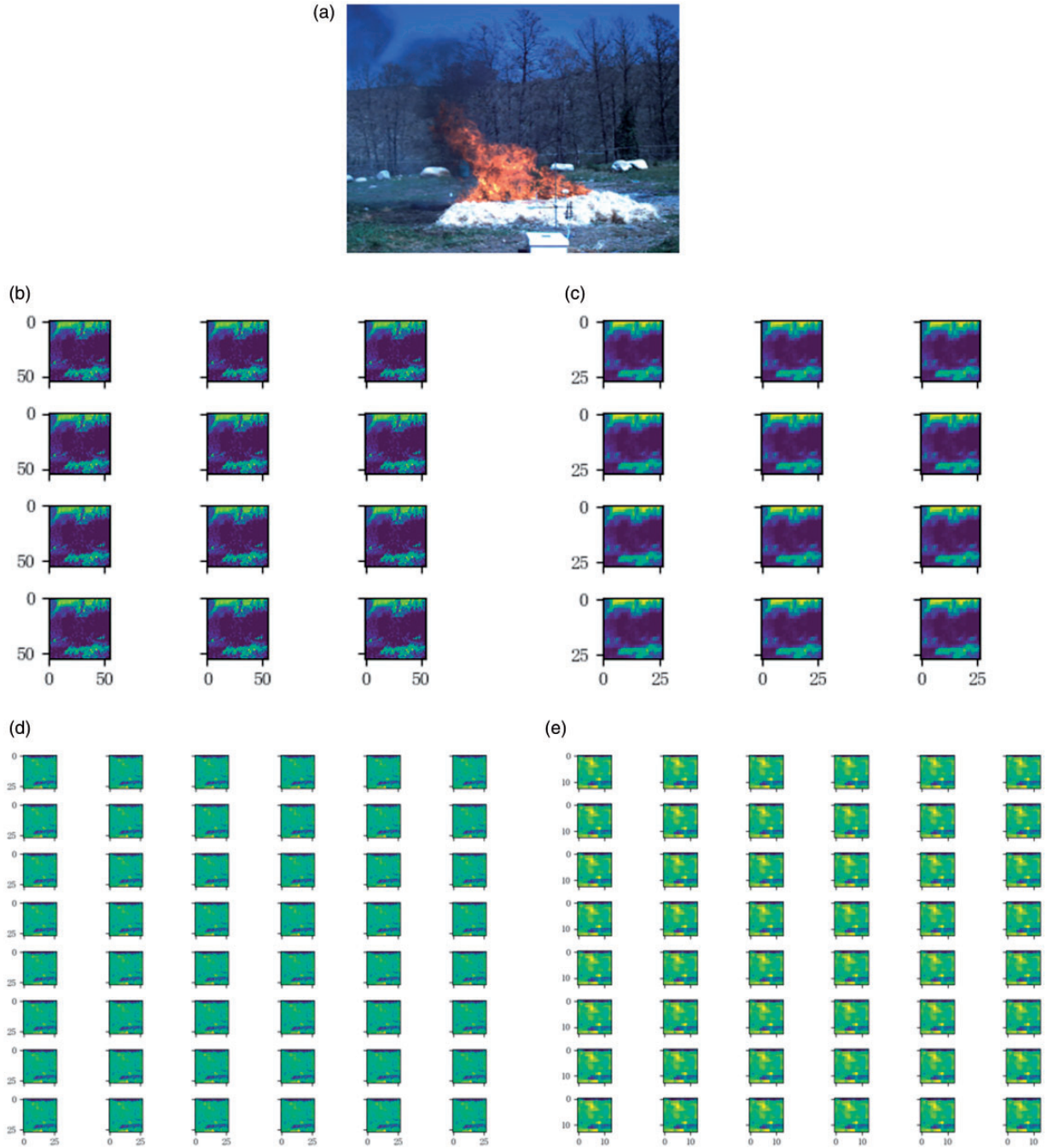


Figure 11. CNN hidden layer features of original fire image. (a) Original image. (b) Feature map with C-1. (c) Feature map with S-2. (d) Feature map with C-3. (e) Feature map with S-4.

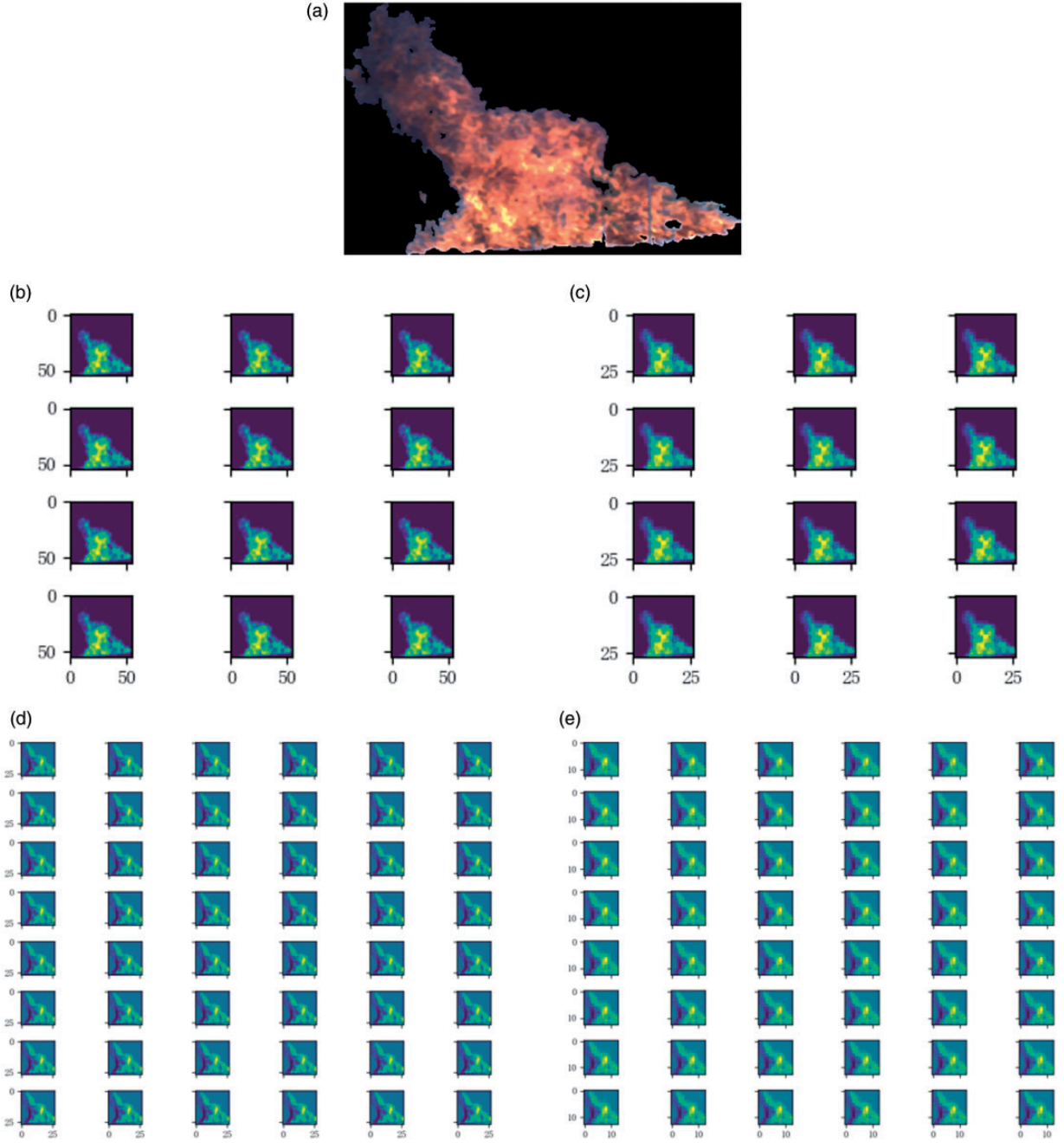


Figure 12. Fire image features of CNN hidden layer combined with color feature. (a) Original image. (b) Feature map with C-1. (c) Feature map with S-2. (d) Feature map with C-3. (e) Feature map with S-4.

rate is set to 0.005, and iteration time is set to 200. The average pooling, maximum pooling, and adaptive pooling modes are shown in Figure 10, respectively.

It can be observed from Figure 10 that after the 25th training, the accuracy of the adaptive pooling mode will remain unchanged at 1.0, and the training effect is slightly better than that of average pooling and maximum pooling. Testing accuracy in each

pooling mode is compared to illustrate the superiority. Sixteen frames of non-fire image and 16 frames of fire image are adopted to form a testing set and average value is taken from three tests. It is shown in Table 4.

It can be seen from Table 4 that the testing accuracy of adaptive pooling is higher than that of mean and maximum pooling; thus, the feasibility of adaptive pooling is verified.

Table 5. Comparison of fire recognition rate under different training samples.

Training samples	Testing samples	Accuracy (%)
Original image	Original image	84.2
Original image	Segmented fire area	81.6
Segmented fire area	Original image	80.4
Segmented fire area	Segmented fire area	90.7

Feasibility verification of the proposed algorithm

In order to verify the feasibility, the features in the middle layer of CNN are extracted. Figure 11 shows the feature map in hidden layer during training based on original image.

The feature map in CNN hidden layer where flame region is segmented before training is shown in Figure 12.

In Figure 11, since there are a lot of trees in the background, this part is extracted as a main feature in the feature map. In Figure 12, only the fire area exists, so that features such as shape and texture can be observed clearly in the feature map of hidden layer (the first problem of reviewer 1).

In order to verify the algorithm, based on the structure in Table 1, the training samples are only involved in the flame region. The results are shown in Table 5.

It is obvious from Table 5 that the recognition rate is only 84.2% and 81.6% based on the training of original image. If the segmented fire area image is adopted as training sample, and the original image is taken as testing sample, the accuracy is as low as 80.4%. This is because there is only the features of fire without background information. The image after segmentation is applied to training sample and testing sample, and the accuracy can be as high as 90.7%. Therefore, it is further proved that the CNN fire recognition algorithm combined with color features has higher accuracy.

Conclusion

In this paper, the process of the forest fire image recognition algorithm based on CNN is presented. Its main feature is that the flame image is employed for training and testing. Then, AlexNet model is introduced, and an adaptive pooling method combined with color features is proposed for the problem that the traditional pooling method in CNN may weaken the image features in some cases. The effects of learning rate, batch size, and other parameters on the performance of CNN are analyzed based on experiments, and the optimal parameters are determined. Candidate flame area is extracted based on color feature; thus, the image feature of non-flame area in the hidden layer is reduced, and the feature, such as shape and

texture, is enhanced. The information loss of image are avoided as adaptive pooling is adopted, and the rate of flame recognition in which fire area is segmentation than that of original image is adopted without segmentation. It is shown that the proposed algorithm has high recognition rate and is feasible. In this paper, the pooling of CNN is modified and applied on forest image recognition, recognition rate and consuming time will be developed deeply and compared with other algorithms in future.

Acknowledgements

This work (grant no. 61703329) was supported by the National Natural Science Foundation of China and the National Key Research and Development Program of Shaanxi Province, China (2019KW-046). The pictures were from Corsica fire Database which was made by the University of Corsica.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work (grant no. 61703329) was supported by the National Natural Science Foundation of China and the National Key Research and Development Program of Shaanxi Province, China (2019KW-046). The pictures were from Corsica fire Database which was made by the University of Corsica.

ORCID iD

Wang Yuanbin  <https://orcid.org/0000-0002-8848-7057>

References

1. Kanwal K, Liaquat A, Mughal M, et al. Towards development of a low cost early detection system using wireless sensor and machine vision. *Wireless Pers Commun* 2017; 95: 475–489.
2. Yuan C, Liu Z and Zhang Y. Aerial images-based forest fire detection for firefighting using optical remote sensing techniques and unmanned aerial vehicles. *J Intell Robot Syst* 2017; 88: 2–4.
3. Sharma J, Granmo OC, Goodwin M, et al. Deep CNN for fire detection in images. In: Boracchi G, Iliadis L, Jayne C, et al. (eds) *Engineering Applications of Neural Networks. EANN 2017. Communications in Computer and Information Science*. Cham: Springer.
4. Shi L, Long F, Lin CH, et al. Video-based fire detection with saliency detection and convolutional neural network. In: *International Symposium on Neural Networks*

- (eds Cong F, Leung A and Wei Q), 2017, pp. 299–309. Cham: Springer.
5. Wang G. *Research on fire detection methods based on machine learning*. China: Dalian University of Technology.
 6. Chen J, Wang Z, Chen H, et al. Dynamic smoke detection using cascaded convolutional neural network for surveillance videos. *J Univ Electron Sci Technol China* 2016; 45: 992–996.
 7. Fu T, Zheng C, Tian Y, et al. Forest fire recognition based on deep convolutional neural network under complex background. *Comput Modernization* 2016; 3: 52–57.
 8. Frizzi S, Kaabi R, Bouchouicha M, et al. Convolutional neural network for video fire and smoke detection. In: *IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society*, 23 October 2016, pp. 877–882. IEEE.
 9. Hohberg SP. *Wildfire smoke detection using convolutional neural networks*. Berlin: Freie Universität Berlin, 2015.
 10. Wang Y and Ren J. Image segmentation for forest fire in low illumination environment based on color feature. *Fire Sci Technol* 2017; 10: 75–78.
 11. Xia W and Xia Z. An improved algorithm for cervical cancer cell image recognition based on convolution neural network. *J China Univ Metrol* 2018; 29: 439–444.
 12. He X and Chen X. Figure vein recognition based on improved convolutional neural network. *Comput Eng Des* 2019; 40: 562–566.
 13. Liu WJ, Liang XJ and Qu HC. Learning performance of convolutional neural networks with different pooling models. *J Image Graph* 2016; 21: 1178–1190.