

A Novel Method for Forest Fire Detection Based on Convolutional Neural Network

Jialiang Feng¹, Dingju Zhu¹⁺, and Lihua Liao²

¹School of Computer, South China Normal University, Guangzhou 510631, China

²School of Information Technology in Education, South China Normal University, Guangzhou 510631, China

Abstract. In a forest fire, smoke and fire always appear together. The study of smoke usually uses traditional methods, while deep learning focuses mostly on the features of flame. According to the fact that the smoke is always observed earlier than the flame in the forest fire, this paper presents a deep convolutional neural network model for forest fire detection based on the study of forest fire detecting images. Experimental results show that the deep convolutional neural network for forest fire detection has a higher accuracy than the traditional method in recognition of forest fires by detecting smoke and flame together. In addition, the deep convolutional neural network for forest fire detection combines with preprocessing of ZCA whitening and padding of same size output, which improves the experimental speed and prediction accuracy.

Keywords: smoke and flame recognition, CNN, forest fire, deep learning

1. Introduction

Forest fire is a phenomenon that causes harm and loss to forest ecosystems and human beings. Forest fires have a significant impact on the life and development of humankind. Therefore, it is very important to detect forest fires.

Detection technology has been widely used in many applications such as detecting video content [1], surgery fatigue[2] and parking space[3], and is also useful and important for prevention and control of forest fire. Fire detection is divided into flame detection and smoke detection. Angayarkkani et al. [4] deduced a set of fuzzy rules based on diffusion image segmentation and achieved some results in forest fire detection. In addition, Laplace Operator, Fourier Transform and Hidden Markov Model are also used in the extraction of features such as shape, color and texture of forest flame.

In recent years, deep learning has also been gradually applied to machine learning, which based on traditional methods such as support vector machines, naive Bayes, conditional random fields and maximum entropy. In the areas of recommendation system [5], text classification and face recognition, deep learning has a good effect.

In the current related work, the detection of forest fire smoke usually uses the traditional machine learning algorithm. However, deep learning, which has achieved good results in the field of image recognition, mostly focuses on the research of flame characteristics and lacks the research on the combination of smoke and flame in the actual forest fire scene. Therefore, this paper presents a deep convolutional neural network model for forest fire detection based on the images of smoke and flame in actual forest fire. The model combines with preprocessing of ZCA whitening and padding of same size output, which can carry out comparative experiments.

2. Smoke and Flame Detection of Forest Fire Based on CNN

+ Corresponding author.
E-mail address: zhudingju@m.scnu.edu.cn

2.1. Smoke and Flame Detection of Forest Fire Based on Deep CNN

There are phases of smoke, flame and coexistence between smoke and flames during the forest fire. This paper designs a deep learning network framework based on convolutional neural network image processing. Network shown in Figure 1:

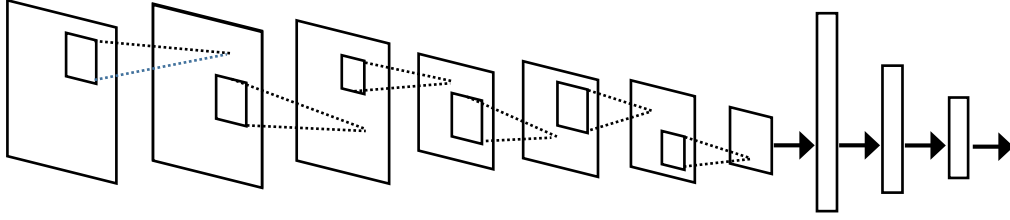


Fig. 1: Deep neural network architecture.

In the convolution layer, the convolution kernel convolutes the image matrix through a sliding window to extract features. After completing the convolution calculation of the local feature, it extracts the feature of the next local receptive field by moving the distance of the stride step. The size of the extracted output feature image is:

$$N^l = \frac{N^{l-1} - K^l + 2P^l}{S^l} + 1 \quad (1)$$

l is the convolution layer. K is the size of the sliding window. P is the pixel that is padded on each side of the input image. S is the stride of the sliding window. Because the input image is equal in width and height, this formula is used in both dimensions.

After the kernel extracted the feature, it needs to add bias and go through the RELU activation function to get output of the convolution layer. Because the image is multichannel, the convolution output of the j th convolution kernel corresponding to each step is:

$$x_j^l = RELU\left(\sum_{t \in \mathbb{R}^D} x_t^{l-1} w_{ij}^l + b_j^l\right) \quad (2)$$

t means the input matrix is the t th channel or t th feature graph. l is the current network layer number. \mathbb{R}^D indicates the number of channels or feature maps in $l-1$ layer. w_{ij}^l is the corresponding weight of convolution of the j th convolution kernel with the t th channel of the input feature. b_j^l is the bias. The network settings are shown in Table 1:

Table 1: Network parameter settings

Layer	Kernel	Stride	Padding	Neuron	Output
C1	3×3	1	same or not	32	$32 \times 32 \times 32$
C2	3×3	1	not	32	$30 \times 30 \times 32$
S2	2×2	2	not	32	$15 \times 15 \times 32$
C3	3×3	1	same or not	64	$15 \times 15 \times 64$
C4	3×3	1	not	64	$13 \times 13 \times 64$
S4	2×2	2	not	64	$6 \times 6 \times 64$
FC5	-	-	-	512	512×1
FC6	-	-	-	4	4×1

2.2. ZCA Whitening for Preprocessing Data

An associated preprocessing step called ZCA whitening is required in the algorithm. Because of the strong correlation between adjacent pixels in the image, the input is redundant for training. The purpose of ZCA whitening is to reduce the redundancy of input. The input has the following properties through the ZCA whitening process: the correlation between features is low; all features have the same variance.

In the process of ZCA whitening, the characteristic dimension of sample data, also known as the number of neurons in the input layer, should be zero-averaged. The result \bar{x}_i is calculated by subtracting the average of the dimensions from each dimension of the sample data. Then normalize the result to the [0,1] range.

$$\Sigma = \frac{1}{m} \sum_{i=1}^m x^i \cdot (x^i)^T \quad (3)$$

Σ is the covariance matrix of training sample data. After using Σ as singular value decomposition, the eigenvector U and eigenvalue diagonal matrix S can be obtained, $U = [u_1, u_2, \dots, u_n]$. u_1 is the most dominant eigenvector of covariance matrix Σ . u_2 is the secondary eigenvector of covariance matrix Σ . Then u_n is the least important eigenvector of covariance matrix Σ . The above feature vector becomes a set of vector bases under the new dimensional coordinates. Therefore, the new sample data after dimension conversion is $x_{rot} = U^T \cdot x$. The dimensions of x_{rot} are independent of each other. Dividing the standard deviation \sqrt{S} by x_{rot} to calculate the variance of each dimension is 1. So, it satisfies the two conditions that the mean of ZCA whitening is close to 0 and the variance is equal. The formula of ZCA whitening is:

$$X_{ZCA,Whitening} = U \cdot \frac{X_{rot}}{\sqrt{S + \varepsilon}} \cdot U^T \quad (4)$$

ε is ZCA whitening dynamic parameters, generally take between [0.01,0.1]. We can adjust the data preprocessing by changing the value of ε until finally obtaining a good feature.

2.3. Padding Method in Convolution Layer of CNN

C1 layer is connected to C2 layer and C3 layer is connected to C4 layer. Convolution layers can extract important features from the input matrix. When the kernel convolutes with the input matrix, the number of convolution operations performed by the edge pixels is often smaller than the number of convolution operations performed by the non-edge pixels. After the feature extraction of two consecutive convolution layers, the boundary feature information tends to weaken or disappear. Therefore, when C1 layer and C3 layer are convoluting, the padding of same size output can ensure that the boundary information can still be maintained when the C2 layer and C4 layer perform convolution extraction.

If the stride S is 1, according to formula (1) above, the size of padding each side is:

$$P = \frac{K-1}{2} \quad (5)$$

K is the size of sliding window. In FIG. 2, when the stride is 1, the size of padding is zero, the size of input image is 6×6 , the size of convolution kernel is 3×3 , and the size of feature after convolution is 4×4 . In FIG. 3, when the stride is 1, the size of padding is 1 each side, the size of input image is from 6×6 to 8×8 after padding, the size of convolution kernel is 3×3 , and the size of feature after convolution is 6×6 .

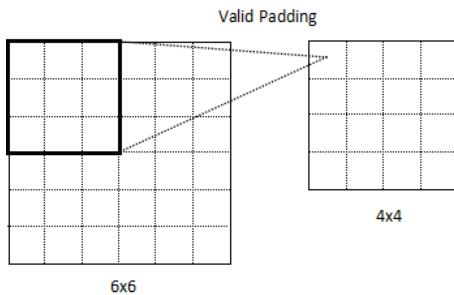


Fig. 2: Valid padding, not filled.

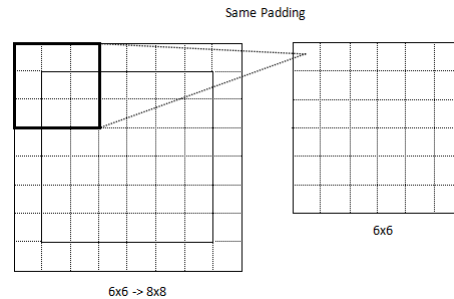


Fig. 3: Same padding, same size output.

3. Experimental Results and Analysis

3.1. Data and Preprocessing

There are ImageNet, CIFAR and other databases that have high usage rate of image recognition. But these databases do not have enough types of image samples for smoke and flame of forest fire that satisfy the

model. In order to train a model that can identify the fire process of forest fire, this paper selects the normal forest images, pre-fire smoke forest images, the fire images of forest smoke and flame coexistence, and a wide range of flame forest images as representative dataset categories. Choosing 125 different images for each category in the Internet as a training set, and 25 different images for each category as a test set.

3.2. Experimental Parameters and Comparative Methods

Experiments uses TensorFlow, Keras as the experimental platform, to build a deep convolutional neural network. In the experiment, we randomly select the validation set from the training set according to the ratio of 10%, and validate in the training of batch updating. Each set of experiments use ten fold cross-validation method to calculate the average.

Experiments are added a control group used SVM on the same dataset. The input feature vector selects the GiST feature. Based on the comparison with traditional methods, according to whether the input data is preprocessed with ZCA whitening, analyzing the effect of dimensionality reduction in smoke and flame images of forest fire. In addition, the features extracted by multiple convolution layers may result in a reduction in boundary information. A control group is set up to add padding of same size output to analyze the role of boundary information in identifying smoke and flame images of forest fire.

3.3. Experimental Results Analysis

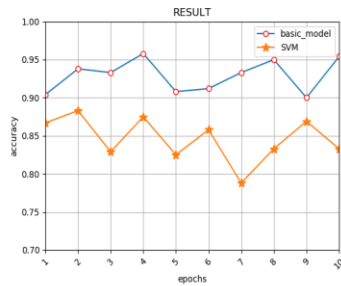


Fig. 4: Benchmark and SVM.

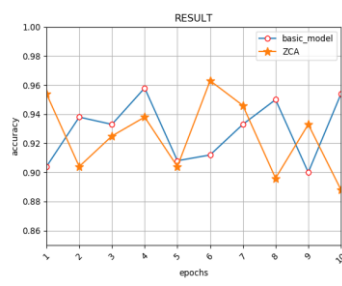


Fig. 5: Benchmark and ZCA.

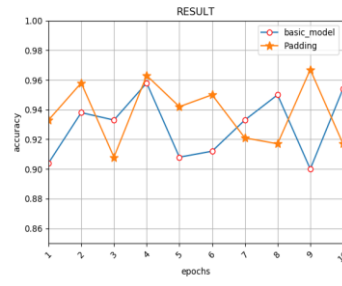


Fig. 6: Benchmark and padding.

Non-ZCA whitening and non-padding are combined with the basic deep convolutional neural network as a benchmark experimental model. Comparison between the SVM and benchmark experimental model is shown in Figure 4. Taking the benchmark experimental model as the standard, the input image data were preprocessed with ZCA whitening and then compared with the data without ZCA whitening, and the result is shown in Figure 5. Based on the benchmark model, padding of same size output is added to convolution layers of C1 and C3 as a control group, which is performed on the same data. The result is shown in tableure 6.

From Table 2, when using the benchmark experimental model, the recognition accuracy has significantly improved compared to the traditional method. The speed of experiment with using data after ZCA whitening preprocessing has improved significantly, and the correct rate is not much difference. The correct rate of padding with the same size output in the convolution layer has a little increase, but the experimental speed is slightly slower.

Table 2: Experimental results

Method	Time	Accuracy
benchmark model	215s/epoch	92.9%
SVM	-	84.6%
ZCA+benchmark model	197s/epoch	92.6%
padding+benchmark model	223s/epoch	93.8%

4. Conclusion

This paper presents a deep convolutional neural network model, for identifying the smoke and flame of real forest fires. Compared with the traditional SVM method, this model has a significant improvement in the accuracy of classifying normal forest images, pre-fire smoke forest images, smoke and flame coexistence of

forest images and a wide range of flame forest images. In addition, based on the benchmark experimental model for forest smoke and flame images in this paper, ZCA whitening preprocessing of smoke and flame images can accelerate the training speed slightly. In the convolution layer, padding the input data with same size output can get a better accuracy than the benchmark experimental model.

In the future, complexity and number of the sample will be increased, and based on the image recognition of smoke and flame in forest fire, we will analyze the development trend of the whole forest fire process according to the weather characteristic data of the fire image.

5. Acknowledgments

This work was supported by Major Project of National Social Science Fund [grant number 14ZDB101].

6. References

- [1] Mustafa R, Min Y, Zhu D. Obscenity detection using Haar-like features and gentle adaboost classifier[J]. The Scientific World Journal, 2014, 2014.
- [2] Zhu D J. Surgery fatigue detection method and system based on internet of things and big data [J]. Basic & Clinical Pharmacology & Toxicology, 2015, 117: 31.
- [3] Zhu D J. Hospital parking spaces detection method based on distributed collaborative awareness and reluctance wireless sensor network. Basic & Clinical Pharmacology & Toxicology, 2016.
- [4] Angayarkkani K,Radhakrishnan N. Efficient forest fire de-tection system: A spatial data mining and image processingbased approach[J]. International Journal of Computer Science and Network Security,2009.9(3):100-107.
- [5] Jian Wei,Jianhua He,Kai Chen,Yi Zhou,Zuoyin Tang. Collaborative filtering and deep learning based recommendation system for cold start items[J]. Expert Systems With Applications,2017,69.