UAV Image-based Forest Fire Detection Approach Using Convolutional Neural Network

Yanhong Chen[1], Youmin Zhang[2], Jing Xin[1], Guangyi Wang[3], Lingxia Mu[1], Yingmin Yi[1], Han Liu[1], Ding Liu[1]

1. Xi'an University of Technology, Xi'an 710048, P.R.China
E-mail: 15891423152@sina.cn

2. Concordia University, Montreal, H3G 1M8, Canada,
E-mail: youmin.zhang@concordia.ca

3. Northwestern Polytechnical University, Xi'an 710048, P.R.China
E-mail: horton1314@foxmail.com

Abstract: Forest fires are very dangerous. Once they become disasters, it is very difficult to extinguish. In this paper, an unmanned aerial vehicle (UAV) image-based forest fire detection approach is proposed. Firstly, the local binary pattern (LBP) feature extraction and support vector machine (SVM) classifier are used for smoke detection, so as to make a preliminary discrimination of forest fire. In order to accurately identify it in the early stage of the fire, according to the convolutional neural network (CNN), it has the characteristics of reducing the number of parameters and improving the training performance through local receptive domain, weight sharing and pooling. This paper proposes another method for detecting forest fires in convolutional neural networks. Image preprocessing operations such as histogram equalization and smooth low-pass filtering are performed prior to inserting the image into the CNN network. The effectiveness of the proposed method is verified by detecting real forest fire images.

Key Words: Forest Fire Detection (FFD), Convolutional Neural Network (CNN),Unmanned Aerial Vehicles (UAVs), Local Binary Pattern (LBP), Image Preprocessing

1 Introduction

Forest fire refers to fire that are freely controlled by humans and are free to spread and expand in forests, causing certain damage and loss to forest ecosystems and human beings. Because of their rapid spread, firefighting and rescue are difficult, and they pose a serious threat to people's lives and the natural environment. Therefore, it is very important to extinguish the fire immediately when the fire is still in its infancy. In order to reduce the risk of forest fires, prevention is far more important than firefighting. If it can be identified at the beginning of the fire, it can reduce various losses.

UAVs are a new type of aviation platform. In recent years, with the continuous maturity of technology, it has been applied to many fields such as meteorological detection, disaster monitoring, power line inspection, and post-disaster rescue. In particular, the lightweight and small size UAVs have the characteristics of low cost, simple operation and flexible maneuvering. They can adjust the work plan and onboard remote sensing equipment according to the real-time situations on site, which is very suitable for the detection of forest fires [25]. At present, relevant theoretical research on forest fire detection based on computational vision has been carried out worldwide. Compared with thermal imaging systems such as hyperspectral sensors, infrared sensors, and multi-spectral sensors, UAVs equipped with ordinary cameras for forest fire identification research have the advantages of low price, common application, and simple operation and the like.

Forest fire detection mainly includes flame detection and smoke detection. There are many different methods of target detection, and we must find a suitable way to detect forest fires. First of all, the smoke characteristics of forest fires are easier to observe than flames. That is, more significant. Therefore, research on smoke detection is first considered. After comparison, the LBP is selected to extract the smoke texture features. Finally, the SVM classifier is used to detect the forest smoke. In addition, early forest fire detection make a little more sense, so we consider simultaneous detection smoke and flame.

Due to its multi-level network structure and thousands of network nodes, the neural network can implement the approximation of complex functions and represent the distribution characteristics of the input data. Among them, the convolutional neural network reduces the number of parameters by core ideas such as local receptive domain weight sharing and pooling to improve training performance. The characteristics of CNN make it have many advantages in speech recognition and image processing, especially for multi-dimensional input vector images that can be directly input into the model, avoiding the complexity of data reconstruction in feature extraction and classification.

This work was supported in part by National Natural Science Foundation of China under Grants 61573282, 61833013 and 61873200 and Natural Sciences and Engineering Research Council of Canada. Corresponding Author: youmin.zhang@concordia.ca.

2 Related Works

There are various methods for detecting forest fires, which can be divided into traditional methods based on manual monitoring and new fire detection systems based on computer vision. Among them, the traditional detection methods can be divided into four categories according to the development order: ground patrol, watchtower detection, satellite remote sensing and aircraft patrol [25]. However, the best prospects for the current situation are video surveillance systems based on computer vision, and this paper has also conducted more in-depth study.

2.1 Local Binary Pattern (LBP)

In the near-field monitoring of forest fire identification technology, even if a fire has already occurred, the flame cannot be directly observed in the field of view of the observation deck or the monitoring station. It is very likely that the smoke is the most obvious visual phenomenon. In the early stage of a forest fire, it is easier to observe scattered smoke. Therefore, smoke identification technology has more research value. After comparing various feature extraction methods, this paper chooses LBP to extract the texture features of smoke to realize smoke recognition in images.

2.1.1 Basic LBP

The basic idea of LBP is to sum the result of comparing the pixels of the image with the pixels around it: taking this pixel as the center, if the surrounding pixel value is greater than the central pixel value, the position of the pixel is marked as 1, otherwise 0, so each pixel can be represented by a binary number.

2.1.2 Round LBP

In order to adapt to the texture features of different scale images and achieve the requirements of gray scale and rotation invariance, Ojala et al. optimized the basic LBP operator: The improved LBP operator allows for any number of sample points within a circular neighborhood of radius R. Thus, an LBP operator containing P sample points in a circular region of radius R is obtained.

2.1.3 LBP Uniform Pattern

For a LBP operator with p samples in a circular region of radius R, 2^p modes will be generated. In order to solve the problem of excessive binary mode and improve statistics, Ojala proposed to use an "equivalent mode" to reduce the dimension of the LBP operator: when the cyclic binary number corresponding to an LBP has two transitions from 0 to 1 or from 1 to 0, the binary corresponding to the LBP is called an equivalent model class. Patterns other than the equivalent mode class are assorted as another class, called the mixed mode class.

2.1.4 LBP rotation invariant mode (RI-LBP)

As can be seen from the definition of LBP, the LBP operator is gray-invariant, but it is not rotation-invariant. Because the rotation of the image will get different LBP values. Maenpaa et al. proposed a LBP operator with rotation invariance: that is, rotating the circular neighborhood continuously to obtain a series of initially defined LBP values, taking the minimum value as the LBP value of the neighborhood.

The feature extraction experiment of the smoke image was carried out by using the above various LBP operators. The results are shown in Fig. 1.

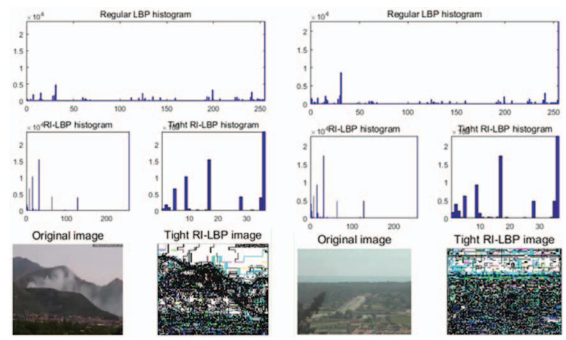

Figure 1: LBP feature extraction of sample images. The picture on the left is a smoke sample; the picture on the right is a negative sample.

Based on the experimental results, we decided to use the "circular equivalent rotation invariant LBP mode" (Tight RI-LBP) in the detection of smoke. Feature extraction is performed on all samples of the data set, and then the calculated Tight RI-LBP statistical histogram is connected into a feature matrix, that is, the LBP texture feature matrix of the entire data set, and finally classified by SVM.

2.2 Image Preprocessing

The main purpose of image preprocessing is to eliminate irrelevant information in the image, restore useful real information, enhance the detectability of relevant information and minimize data, thereby improving the reliability of feature extraction, image segmentation, matching and recognition. Therefore, before the acquired fire image is input to the CNN, the necessary preprocessing of the image may be helpful to improve the accuracy of the fire detection.

2.2.1 Histogram matching (hm)

The gray histogram counts the number of pixels in each gray level in the image. The basic idea of histogram matching is to transform the histogram of the original image into a uniformly distributed form. The gray level with a large number of pixels in the image is broadened, and the gray level with a small number of pixels is compressed, thereby expanding the dynamic range of the pixel value, improving the contrast and the change of the gray tone, and making the image clearer.

Let the original image have a gray level of f at (x, y) and the changed image be g, then the method of image enhancement can be expressed as mapping the gray level f at (x, y) to g. In the gray histogram matching process, the mapping function for the image can be defined as: $g = EQ(f)$, this mapping function $EQ(f)$ must satisfy two conditions (where L is the gray level of the image):

(a) $EQ(f)$ is a single-valued single-increasing function in the range of $0 \leq f \leq L - 1$. This is to ensure that the enhancement process does not disturb the gray-scale order of the original image, and the gray levels of the original image remain in an arrangement from black to white (or from white to black) after the transformation.

(b) $0 \leq g \leq L - 1$ for $0 \leq f \leq L - 1$. This condition guarantees the consistency of the dynamic range of the gray value before and after the transformation.

This method is very useful for images that are too bright or too dark in the background and foreground. A major advantage of this approach is that it is a fairly straightforward technique and is reversible. If the equalization function is known, the original histogram can be restored and the amount of computation is small.

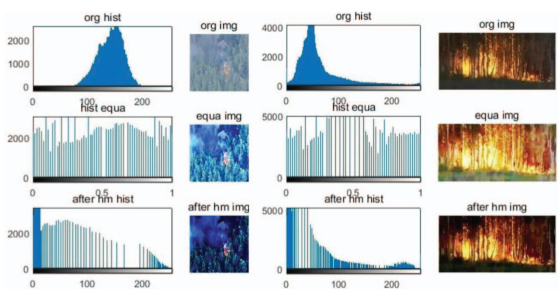

Figure 2: Results of histogram matching of two sample images.

2.2.2 Image smoothing

Image smoothing refers to an image processing method for highlighting a wide area of an image, a low-frequency component, a trunk portion, or suppressing image noise and interfering with high-frequency components, which is essentially low-pass filtering, which aims to smooth the brightness of the image, reduce the abrupt gradient and improve the image quality. Among them, the commonly used neighborhood averaging method (na) belongs to the spatial domain processing method. The neighborhood average is based on the background of the image or the change in gray of the target portion is continuous and slow, while the particle noise causes a sudden change in the gray level of some pixels on the image. The idea is to use the average value of the gray level of the point (x, y) and its neighbors in the image instead of the gray value of the point (x, y), which can produce a "smooth" effect on the point where the brightness is abrupt. The gradation of the mutation can be smoothed by the neighborhood averaging. Therefore, the neighborhood averaging method is used to smooth the image.

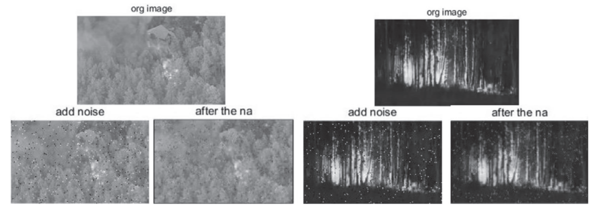

Figure 3: Results of neighborhood averaging smoothing.

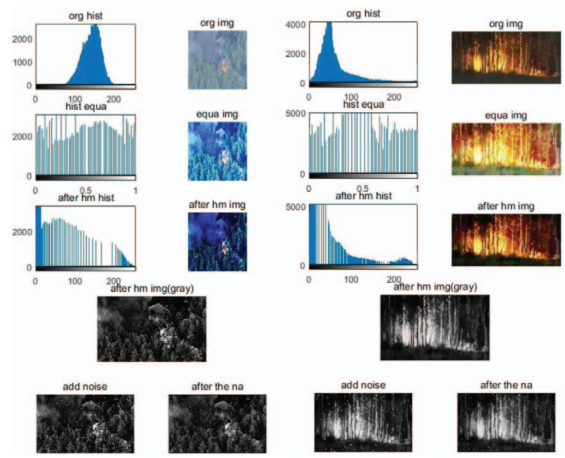

Figure 4: Results of histogram matching + neighborhood averaging smoothing.

It can be seen from Fig. 4 that the image becomes smoother after preprocessing operations such as histogram matching and field averaging. The experiment in chapter 4 also confirmed that image preprocessing has improved network performance to some extent.

3 Forest Fire Detection Using CNN

The flowchart of traditional fire detection method and the proposed CNN-based detection method is shown in Fig. 5 [30].

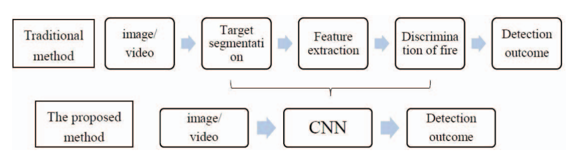

Figure 5: Comparison of two fire detection methods.

Video images are usually affected by the detection environment and uneven illumination during the acquisition process, and pre-processing operations such as image enhancement and filtering were required before detection. In this paper, the histogram matching is used for image enhancement, and the neighborhood averaging method is used for noise filtering. In addition, in order to simultaneously perform smoke and flame detection, the network structure proposed in [25] is improved. The system flow adopted in this paper is shown in Fig. 6.

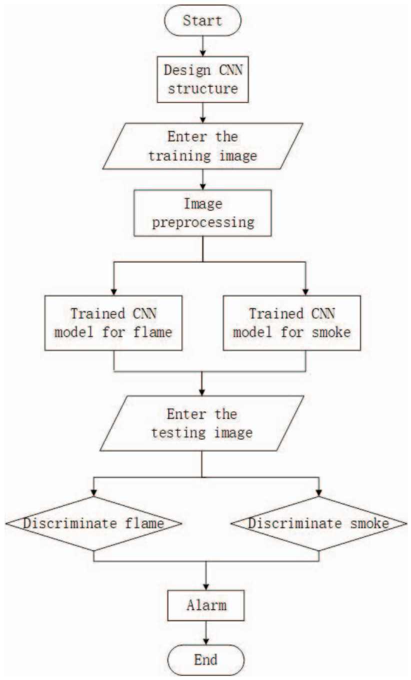

Figure 6: System flow chart.

3.1 Characteristics of CNN

Convolutional neural network is a new neural network method that combines artificial neural network and deep learning technology. CNN uses the core ideas of local receptive domain and weight sharing (pooling) to reduce the number of parameters to improve training performance, and to ensure the robustness of the image to displacement, scaling and deformation.

The first major feature of CNN is the local receptive domain. For each neuron, it is not necessary to perceive the global image, only need to perceive the local part, and then combine the local information at a higher level to get the global information.

The second major advantage of CNN, weight sharing, can further reduce the parameters. To describe a large image, the average (or maximum) of a particular feature over an area of the image can be calculated. The implicit principle is that the statistical properties of a part of the image are the same as the other parts. The operation of this aggregation is called pooling. According to the method of computing pooling, there are commonly used mean pooling and maximum pooling.

3.2 Structure of CNN

A convolutional neural network is a multi-layered neural network, each layer comprising a plurality of two-dimensional feature plans, each of which includes a plurality of independent neurons. In most of the current research work, CNN is used in a variety of machine learning problems including face recognition, document analysis and language detection. Here is an example of the CNN model proposed in [25]. Its structure is shown in Fig. 7.

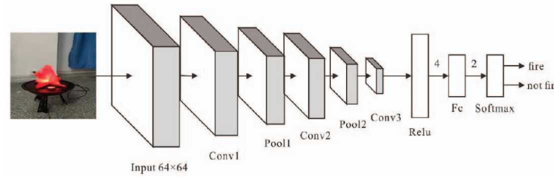

Figure 7: CNN-9 model of forest fire detection experiment.

After conducting the forest fire detection experiment, we found that the 9-layer CNN model is still difficult to achieve high accuracy, so it was improved to $CNN-17$ and the samples of the data set were tested.

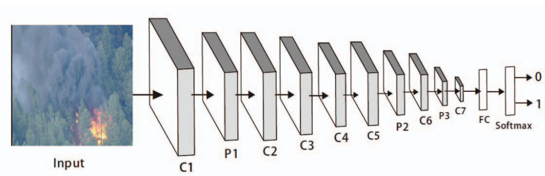

Figure 8: CNN-17 model for forest fire detection experiments (some layers are hidden).

The network parameters of the $CNN-17$ model are shown in Table 1. Compared with the 9-layer network, the $CNN-17$ model adds three layers of convolutional layers, a set of convolution-pooling layers, and multiple regularization layers, which theoretically increases the reliability and adaptability of the model. It is worth mentioning that the model design of $C2-C3-C4-C5$ refers to the "bottleneck structure" proposed in [18], in which the $C2$ and $C5$ layers are responsible for reducing and then increasing (restoring) the size of the feature map, making the $C3-C4$ layer become a "bottleneck" with smaller input/output sizes.

4 Experiment

At present, there is no universal open fire image dataset. In order to verify the feasibility of the proposed LBP-based smoke detection and CNN-based forest fire detection methods, this paper uses the six-rotor UAV $(DJI900)$ equipped with a $SONY A7$ camera to obtain forest fire images. The training data sets for the two groups of experiments used images taken from real forest fire. The test images are forest fire related images obtained from the website. The two data sets are shown in Table 2 and Table 3, respectively.

4.1 Experiment 1: LBP+SVM for smoke detection

The test was performed on the validation sample set at a correct rate of 100%, which may be due to the high similarity between the validation set and the training set sample. In addition, the test data in data set 1 was tested with an accuracy of 99.81%. The accuracy rate is calculated as follows:

$$accuracy = 1 - \frac{1}{300}\sum_{j=1}^{300}(result - label) \qquad (1)$$

Table 1: Model parameters of CNN-17.

Input	C1	P1	C2	C3	C4	C5	P2	C6	P3	C7	Fc
128*128*3	20*9*9	2*2	50*1*1	100*7*7	200*9*9	200*1*1	2*2	500*7*7	2*2	500*7*7	4*1*1

Table 2: Experimental data set 1 for LBP extraction.

Smoke images	Positive samples	Negative samples	Testing samples
Number of samples	200	200	300
Pixel size	400*240	320*240	Uncertain

Table 3: Data sets used in CNN-based forest fire detection experiments. The dataset contains a total of 2100 images, including training sets and testing sets, divided into two categories, the first type is smoke images; the other is flame images, both of which include positive and negative samples.

Categories		Training sets	Testing sets	Total
Smoke images	Positive	500	230	730
	Negative	400	70	470
Flame images	Positive	500	174	674
	Negative	400	126	526
Total		1800	300	2100

Table 4: Comparison of detection performance before and after model improvement.

Results	CNN-9	CNN-9(hm+na)	CNN-17(hm+na)
α	0.53	0.61	0.86
β	0.61	0.80	0.98
γ	0.68	0.72	0.34

4.2 Experiment 2: Forest fire detection through CNN

The effects on the performance of forest fire detection experiments before and after the model's improvement are shown in Table 4.

Among them, the meaning of each parameter is as follows:

(i) Accuracy is:

$$\alpha = \frac{Number of correct predictions}{Total number of testing datas} \tag{2}$$

(ii) The True Positive Rate (which corresponds to the proportion of positive data points that are correctly considered positive) is expressed as:

$$\beta = \frac{TruePositive}{FalseNegative + TruePositive} \tag{3}$$

(iii) False Positive Rate (corresponding to the proportion of all negative data points that are incorrectly considered positive):

$$\gamma = \frac{FalsePositive}{FalsePositive + TrueNegative} \tag{4}$$

From the results shown in Table 4, it is clear that the improvement of the model from $CNN-9$ to $CNN-17$ improves the accuracy and true positive rate, and the false positive rate is also greatly reduced, which verifies the accuracy of the model. In addition, image preprocessing also increases the performance of the network to some extent. For the samples of the test set, the test results of a single image without preprocessing are shown in Fig. 9; the smoothed test results are shown in Fig. 10. It can be seen

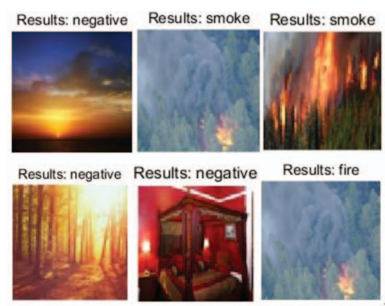

Figure 9: Results of forest fire detection(without pre-processing).

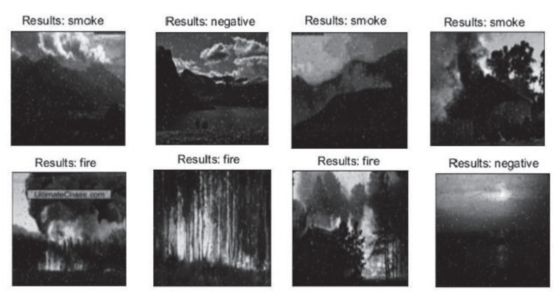

Figure 10: Results of forest fire detection(hm+na).

from Fig. 9 and Fig. 10 that the accuracy of the CNN-based forest fire detection method is still very high, which verifies the effectiveness of the proposed method.

5 Conclusion

This paper proposes: (1) a smoke detection method combining LBP feature extraction and SVM classifier, using texture features for smoke detection; (2) a forest fire detection method based on convolutional neural network: smoke and flame detection are performed by two CNN models. Based on the result of feature extraction (Fig. 1), the "circular equivalent rotation-invariant LBP mode (tight RI-LBP)" is used in the detection of smoke. Compared with the CNN-9 model, the CNN-17 model reduces the complexity of the algorithm and improves the detection accuracy. In addition, the method proposed in this paper detects both smoke and flame. Finally, the effectiveness of the proposed method is verified by two sets of experiments.

REFERENCES

[1] A. Krizhevsky, I. Sutskever, G. Hinton. "Imagenet classification with deep convolutional neural networks[J]". *Advances in Neural Information Processing Systems*, 25, 2012: 1106-1114.

[2] B. C. Arrue, A. Ollero, J. R. Martinez de Dios. "An intelligent system for false alarm reduction in infrared forest-fire detection". *Intelligent Systems & Their Applications*, 2000, 15(3): 64-73.

[3] B. U. Toreyin, Y. Dedeoglu, A. E. Cetin. "Wavelet based real-time smoke detection in video". *European Signal Processing Conference*, 2005: 1-4.

[4] C. Emmy Premal, S. S. Vinsley. "Image processing based forest fire detection using YCbCr colour model". *The International Conference on Circuits, Power and Computing Technologies*, 2014

[5] C. Feichtenhofer, A. Pinz, R. P. Wildes. "Spatiotemporal multiplier networks for video action recognition". *IEEE Conference on Computer Vision & Pattern Recognition(CVPR)*, 2017: 7445-7454.

[6] C. Gulcehre, M. Moczulski, M. Denil, and Y. Bengio. "Noisy activation functions". *International Conference on machine learning*, 2016: 3059-3068.

[7] L. Chang, X. M. Deng, M. Q. Zhou, Z. K. Wu, Y. Yuan, S. Yang, H. G. Wang. "Convolutional neural networks in image understanding". *Journal of Automation*, 2016, 42(9): 1300-1312.

[8] T.-H. Chen, L. K. Cheng, S.-M. Chang. "An intelligent real-time fire-detection method based on video[A]". *In Proceedings of IEEE 37th Annual* 2003 *International Carnahan Conference on Security Technology*, 2003: 104-111.

[9] C. Yuan, Z. X. Liu, Y. M. Zhang. "Aerial images-based forest fire detection for firefighting using optical remote sensing techniques and unmanned aerial vehicles". *Journal of Intelligent and Robotic Systems*, (2017) 88: 635-654.

[10] C. Yuan, Y. M. Zhang, and Z. X. Liu. "A survey on technologies for automatic forest fire monitoring, detection, and fighting using unmanned aerial vehicles and remote sensing techniques", *Canadian Journal of Forest Research*, 2015, 45(7): 783-792.

[11] C. Yuan, K. A. Ghamry, Z. M. Liu and Y. M. Zhang. "Unmanned aerial vehicle based forest fire monitoring and detection using image processing technique". *IEEE Chinese Guidance, Navigation & Control Conference*, 2017: 1870-1875.

[12] C. Yuan, Z. X. Liu, Y. M. Zhang. "Fire detection using infrared images for UAV-based forest fire surveillance". *International Conference on Unmanned Aircraft Systems*, 2017:567-572

[13] E. Prema1, S. S. Vinsley. "Image processing based forest fire detection using YCbCr colour model". *The* 2014 *International Conference on Circuit, Power and Computing Technologies (ICCPCT)*, doi:10.1109/ICCPCT.2014.7054883.

[14] H. Cruz, M. Eckert, J. Meneses and J. Mart nez. "Efficient forest fire detection index for application in unmanned aerial systems (UASs)". *Sensors*, 2016. Vol. 16, Pages 893.

[15] H. Yamagishi, J. Yamaguchi. "Fire flame detection algorithm using a color camera". *Proceedings of* 1999 *International Symposium on Micro Electro – Mechanical System and Human Science*, 1999, 255-260.

[16] K. He, X. Zhang, S. Ren, J. Sun. "Deep residual learning for image recognition". *The* 2016 *IEEE Conference on Computer Vision & Pattern Recognition(CVPR)*, 2016, arXiv: 1512.03385 [cs.CV]

[17] M. Bisquert, E. Caselles, J. M. Snchez, and V. Caselles. "Application of artificial neural networks and logistic regression to the prediction of forest fire danger in Galicia using MODIS data". *International Journal of Wildland Fire*, 2012, 21(8): 1025-1029.

[18] R. K. Srivastava, K. Greff, and J. Schmidhuber. "Highway networks". *Machine Learning (cs.LG); Neural and Evolutionary Computing (cs.NE)*, 2015, arXiv: 1505.00387.

[19] Q. Zhang, J. Xu, L. Xu and H. Guo. "Deep convolutional neural networks for forest fire detection". *International Forum on Management, Education and Information Technology Application*, 47: 568-575, 2016.

[20] S. A. Saleh, S. A. Suandin, and H. Ibrahim. "Recent survey on crowd density estimation and counting for visual surveillance". *Engineering Applications of Artificial Intelligence*, 2015.

[21] V. Lempitsky, and A. Zisserman. "Learning to count objects in images". *The 24th Annual Conference on Neural Information Processing Systems (NIPS)*, 2010.

[22] W. Ye, J. Zhao, S. Wang, Y. Wang, D. Y. Zhang, Z. Y. Yuan. "Dynamic texture based smoke detection using Surfacelet transform and HMT model". *Fire Safety Journal*, 2015, 73: 91-101.

[23] W. Horng, J. Peng, and C. Chen. "A new image based real-time flame detection method using color analysis [A]". In *Proceeding of Networking, Sensing and Control*, 2005, Tucson, Arizona, USA, 2005: 100-105.

[24] W. Phillips III, M. Shah, and N. Vitoria Lobo, "Flame recognition in video". *Pattern Recognition Letters*, 2002, 23(13): 319-327.

[25] Y. H. Chen, Y. M. Zhang, J. Xin, Y. Yi, D. Liu, and H. Liu. "A UAV-based forest fire detection algorithm using convolutional neural network". *The 37th Chinese Control Conference (CCC2018)*, 58-63, 2018.

[26] Y. Sun, X. Wang, and X. Tang. "Deep learning face representation from predicting 10,000 classes". *The* 2014 *IEEE Conference on Computer Vision and Pattern Recognition*, 1891-1898.

[27] Y. Zhang, D. Zhou, S. Chen, S. Gao, and Y. Ma. "Single-image crowd counting via multi-column convolutional neural network". *The IEEE Conference on Computer Vision & Pattern Recognition*, 2016.