Smoke Detection Algorithm based on Negative Sample Mining

Pei Ma
Wuhan Textile University of Hubei Province, People's Republic of China
mp_668@163.com

Feng Yu
Wuhan Textile University of Hubei Province, People's Republic of China
yufeng@wtu.edu.cn

Changlong Zhou
Wuhan Textile University of Hubei Province, People's Republic of China
zcl@wtu.edu.cn

Minghua Jiang
Wuhan Textile University of Hubei Province, People's Republic of China
minghuajiang@wtu.edu.cn

Abstract: *Forest fire is one of the most dangerous disasters that threaten the safety of human life and property. In order to detect fire in time, we detect the smoke when the fire breaks out. However, it is still a challenging task due to the variations of smoke in color, texture, shape and the disturbances of smoke-like objects. Therefore, the accuracy of smoke detection is not high, and it is accompanied by a high false positive rate, especially in the real environment. To tackle this problem, this paper proposes a novel model based on Faster Region-based Convolutional Network (R-CNN) which utilizes negative sample mining method. The proposed method allows the model to learn more negative sample features, thereby reducing false positives in smoke detection. The experiments are performed on self-created dataset containing 11958 images which are collected from cameras placed in villages or towns and existing datasets. Compared to other smoke datasets, the self-created dataset is larger and contains complex scenes. The proposed method achieves 94.59% accuracy, 94.35% precision and 5.76% false positive rate on self-created dataset. The results show that the proposed network is better and more robust than previous works.*

Keywords: *Smoke detection, negative sample mining, false positives, convolutional neural network.*

Received August 15, 2020; accepted October 10, 2021
https://doi.org/10.34028/iajit/19/4/15

1. Introduction

Forest fire accidents cause a great impairment to human health and nature. To solve fire hazards in time, we usually use smoke detection methods to detect fire phenomena [3, 4, 17, 22, 44]. In the early days, sensors are used to determine whether fire occurs [7, 19]. However, this method exists some inherent limitations. Concretely, they must be installed close to the location of fire. Therefore, sensors are suitable for indoor scenes and narrow places but not for outdoor scenes. In fact, forest fire accidents usually occur in the wild, so the practical value of sensors is relatively low.

Nowadays, various methods based on computer vision have been proposed for smoke detection [5, 39, 40, 42]. In the past few years, smoke features are mainly extracted by the traditional manual feature extraction method. In these methods, hand-crafted features are input into classifiers, such as Support Vector Machines (SVM) [38], Adaboost [10], for features learning and classification. Zhao [47] utilize kalman filtering to segment smoke regions in images, then, calculate the flutter direction angle of segmented regions and adopt Local Binary Motion Pattern (LBMP) to extract smoke dynamic texture features. Finally, Adaboost is used to classify smoke. Pundir and Raman [29] adopt the background subtraction method to obtain region of high motion. Then, RGB values and Luminance are used to extract smoke pixel intensities and Local Extrema Co-occurrence Pattern is utilized to extract smoke texture. Finally, these features are fed into Deep Belief Network (DBN) for smoke classification. The methods mentioned above have the following disadvantages:

1. The procedures of extracting smoke features are complicated.
2. Smoke features cannot be comprehensively extracted. In realistic scenes, it is difficult to achieve high accuracy and low false positive rate, due to the diversity of smoke characteristics and the interference from smoke-like objects.

Recently, Convolutional Neural Networks (CNN) have achieved outstanding performance in the fields of image classification [18, 24, 33] and object detection [6, 16, 20, 25, 30, 31, 32]. In contrast to hand-crafted feature methods, deep learning can automatically learn features of images for object detection and classification. Khan *et al.* [20] propose an energy efficient system based on Visual Geometry Group Network (VGG-16) [33] for early smoke detection in both normal and foggy Internet of Things (IoT) environments. Muhammad *et al.* [25] apply AlexNet [23] to fire detection. This fire detection scheme achieves higher accuracy compared to earlier methods. Although the above CNN smoke detection methods avoid the complex process of hand-crafted feature extraction, there are still many false positives in realistic applications. Therefore, they still cannot meet the needs of real scenes. The complexity of real scenes is reflected in the following aspects:

a) The varieties of smoke in color, texture, and shape.
b) Objects with similar smoke features, such as fog, trail, and water surface reflection.

In this paper, a new model based on Faster R-CNN [31] is proposed, which uses improved negative sample mining to reduce false positives in smoke detection. Concretely, first, Faster R-CNN is used to obtain the first round results in smoke detection. Then, through analysis of detection results, images of frequent false positive areas could be obtained. According to the rules of selecting candidate regions in Region Proposal Network (RPN) [31], the prepared smoke templates are integrated into these images. Finally, the dataset is fed back to Faster R-CNN to mine more negative samples and learn features. In addition, existing smoke datasets such as Ko *et al*. [21], Toreyin *et al*. [35], and Vezzani *et al*. [37] are relatively small. These datasets cannot apply to complex scenes. To meet actual application scenes, dataset used for smoke detection should be varied and abundant. Therefore, a large number of real-world images are collected by cameras located in villages and towns. The proposed dataset contains smoke images and non-smoke images from different cities, complex scenes, and diverse weather. Different from previous methods, the proposed method is divided into two steps:

1. Convolutional neural network is trained on self-created dataset to obtain the model for detecting smoke;
2. Smoke detection results of the first step are analyzed to get frequently false positive areas and negative sample mining method is integrated into the backbone network to reduce smoke false positives. Experimental results show that the proposed method achieves 94.59% accuracy, 94.35% precision, and 5.76% false positive rate on self-created dataset. The main contributions of this paper are summarized as follows:
    a) A dataset based on Pattern Analysis, Statical Modeling and Computational Learning (PASCAL) VOC2007 [9] benchmark was created. This dataset contains 11958 images in total, 7,773 images (65%) for training, 1,794 images (15%) for validation and 2,391 images (20%) for testing.
    b) This paper proposes a negative sample mining method that is applied to smoke detection framework to reduce false positives.
    c) Experimental results on several datasets show that the proposed method outperforms previous methods.

The rest of this paper is organized as follows: related work is given in section 2; the proposed method and the overall framework are detailed in section 3; experimental environments and results are summarized in section 4; finally, the summary of this paper is presented in section 5.

2. Related Work

In recent years, many smoke detection methods have been proposed [5, 16, 35, 39, 42, 45, 46]. These methods can be divided into hand-crafted features extraction and deep learning methods. In this section, the corresponding methods are briefly reviewed.

2.1. Hand-Crafted Feature Extraction

In the early stage, researchers studied smoke detection by manually extracting smoke characteristics. Toreyin *et al.* [35] propose an algorithm that mainly determines edge region based on its wavelet sub-band energy decreasing with time. Subsequently, RGB and chrominance values are analyzed to get smoke region. Yuan [42] propose an accumulative motion model based on integral image by fast estimating motion orientation of smoke. Chunyu *et al.* [5] adopt Lucas Kanade optical flow algorithm to calculate optical flow of candidate regions. Then, motion characteristics are calculated according to the optical flow results. Finally, back-propagation neural network is adopted to get classification results. Zhao [46] present a candidate smoke region segmentation method based on rough set theory. Kalman filtering [12] is used to exclude interference of static smoke-color objects and adopt color space to get smoke regions.

The process of smoke feature extraction by hand-crafted feature extraction method is very complex, and it cannot extract smoke feature sufficiently. Therefore, it is difficult to achieve high accuracy, high precision, and low false positive rate in complex scenes.

2.2. Deep Learning

In recent years, deep learning methods have shown superior performance in many fields such as face recognition [8], pedestrian detection [43], object segmentation [11, 13], and image classification [34]. Yin *et al.* [41] propose a novel video-based smoke detection method that captures space and motion context information by using deep convolutional motion-space networks. Hu and Lu [16] propose a spatio-temporal convolutional neural network for video smoke detection. This network uses a multi-task learning strategy to identify smoke and estimate optical flow while capturing intra-frame appearance features and inter-frame motion features. Xu *et al.* [39] propose deep saliency network to detect video smoke.

Concretely, the pixel-level and object-level salient convolutional neural networks are combined to extract informative smoke saliency map. Zhang *et al.* [45] propose joined Deep Convolutional Neural Networks (DCNN), in which global image-based CNN is used for smoke classification and local patch-based CNN is used for smoke localization.

Compared with hand-crafted feature extraction methods, the above methods have achieved significant improvements. However, real-world scenarios are quite complex. Due to the diversity of smoke characteristics and the interference of smoke-like objects, these above methods still have difficult to reduce false positives. In order to tackle this problem, we use an improved negative sample mining method during smoke detection.

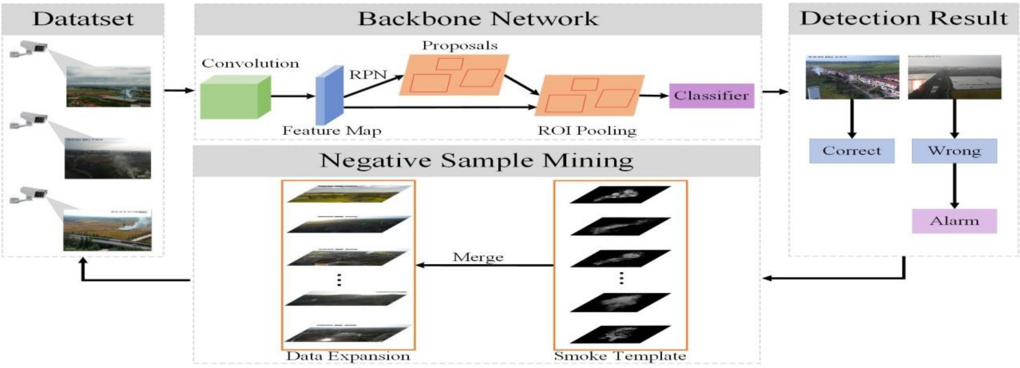


Figure 1. The architecture of proposed method consists of four modules. In the experiment of this paper, backbone network uses ResNet-50 for feature map extraction.

3. Proposed Approach

In this section, we give detailed description of the proposed network which applies improved negative sample mining method to the backbone network, aiming at decreasing false positives in smoke detection.

3.1. Overall Framework

As shown in Figure 1, the architecture of the proposed method consists of four modules:

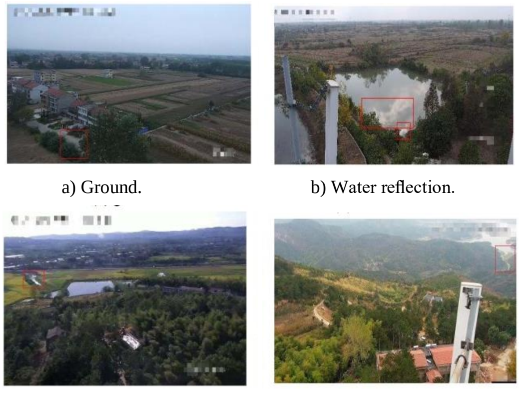
a) Dataset: the experimental dataset is collected from cameras that are set in villages and towns. The areas monitored by cameras are prone to fire.
b) Backbone Network: we use Faster R-CNN as our backbone network, which is mainly composed of Convolution Layer, Region Proposal Network, Region of Interest (ROI) Pooling Layer, and Classification. This network is trained on self-created smoke dataset and gets the first-round model.
c) Detection Result: we contain frequently false positive regions by analyzing the marked smoke location obtained from the first-round results.
d) Negative Sample Mining: through statistical analysis of smoke shapes in our dataset, the network can mine more negative samples, so as to reduce false positives in smoke detection.

3.2. Backbone Network

Faster R-CNN is used as backbone network which is an object detection framework that generates proposals by using RPN. The backbone network consists of four main components:

a) Convolution layer: extracts feature maps for input pictures. The feature maps are shared for subsequent RPN layer and ROI pooling layer.
b) Region Proposal Network: generate proposals that may contain smoke.
c) ROI Pooling layer: map the proposals to the corresponding position of the feature map.
d) Classification: Classify and regress the obtained proposal box.

Although the backbone network can detect the area where the smoke is located very well, however, false positives rate is relatively high in realistic applications. As shown in Figure 2, through detailed analysis of the first-round results, false positive areas in images frequently concentrate in ground, trail, water reflection, rivulet and etc., To tackle this problem, negative sample mining method is used to reduce false positives.



a) Ground. b) Water reflection.

c) Rivulet. d) Trail.

Figure 2. The red border is the area where false positives frequently occur.

3.3. Negative Sample Mining

In the RPN, proposals are selected on the feature map using a 3*3 sliding window. The center point of each sliding window in the feature map can correspond to a certain point of the original image. At each center point, RPN selected nine boxes with different proportions and aspect ratios. RPN has two branches: classification and box regression. In the classification branch, foreground or background are distinguished by setting Intersection Over Union (IOU) [28] threshold between the proposal and the ground-truth box, and each proposal is assigned a binary label (that is, object class or background). In the box regression branch, the foreground coordinates are adjusted according to the ground-truth box. Finally, the final foregrounds are obtained by combining results of two branches.

The proposed negative sample mining method aims to enable the network to learn more comprehensive feature information. First, according to the first-round detection results, we get images of regions where false positives often occur. Second, we use a clustering algorithm to statistically analyze the smoke shapes in the dataset, which allowed us to obtain nearly twenty typical smoke templates. As shown in Figure 3, these are typical smoke templates selected by statistical analysis. Third, according to the rules for RPN to select candidate regions, these prepared smoke templates are merged into appropriate positions in these images which frequently have false positives. As shown in Figure 4, these smoke templates are merged into images. Finally, these processed images are input into the network for retraining. The negative sample mining method can mine the frequently false positive proposals as negative samples, improve the feature extraction ability of the network, and reduce the false positives in smoke detection.

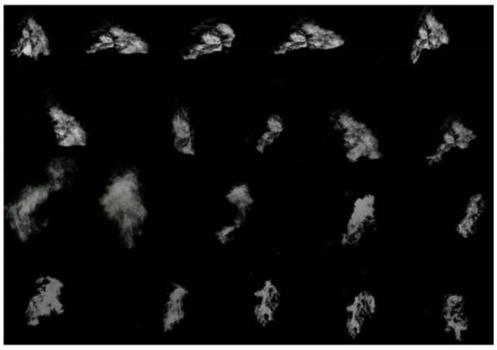


Figure 3. According to smoke shapes in original dataset, some typical smoke templates are clustered.

In our experiment training phase, when a proposal's IOU is higher than 0.7, this proposal is foreground and is assigned a positive label. A negative label is assigned to a non-positive proposal if its IOU is lower than 0.3 for all ground-truth boxes. Proposals that are neither positive nor negative do not contribute to the network training. Subsequently, the Non-Maximum Suppression (NMS) algorithm [26] is used to remove duplicate proposals. The NMS algorithm is a local maximum search, non-local maximums values will be excluded that can reduce computational cost in the network.

In the experimental testing phase, many proposals will be generated when test images pass through the RPN. Subsequently, the trained model predicts these proposals categories and generates corresponding scores. We set the prediction threshold to 0.8. If the proposal's score is more than or equal to 0.8, this image is considered to exist smoke.

a) Village. b) Field.

c) Lake. d) House.

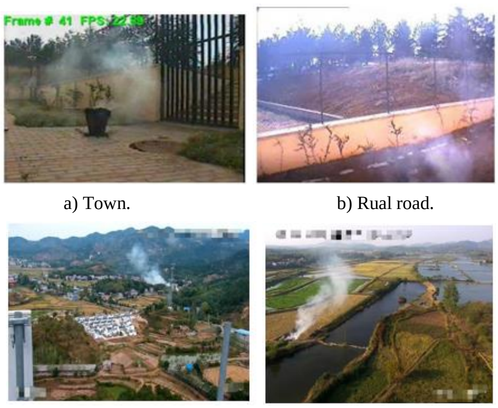
Figure 4. The above pictures are part of the dataset after incorporating smoke templates.

4. Experimental Results and Analysis

Tensorflow framework [1] is used to evaluate the proposed method. Experiments are performed on a workstation with Intel Core I7, 64GB RAM, NVIDIA GeForce RTX 2080 Ti GPU, and the operating system is Ubuntu18.04. For the experiment strategies, the parameters of each layer are updated using Stochastic Gradient Decent (SGD) with the batch-size of 256, momentum of 0.9, weight decay of 0.0001, and initial learning rate of 0.01. In RPN, there are two initialized thresholds $\alpha 0$ and $\alpha 1$. If the IOU between the target box and the ground truth is more than $\alpha 0=0.7$, then it is a positive sample; if the IOU between the target box and the ground truth is less than $\alpha 1=0.3$, then it is a negative sample. In the implementation process, ResNet-50 is adopted to generate feature map, and Faster R-CNN is applied as the backbone network. The training procedure was set to maintain the verification accuracy for four consecutive epochs.

4.1. Dataset and Evaluation Metrics

Existing smoke detection datasets generally cannot meet requirements in complex scenes. For example, as shown in Figure 5-a) and b), made by Ko *et al*. [21], Toreyin *et al*. [35], and Vezzani *et al*. [37] are relatively small and the scenes of dataset are simple. For this reason, we propose a self-created smoke dataset containing 11,958 images, collected by cameras installed in various villages and towns. As shown in Figure 5-c), and d), the scenes contained in the self-created dataset are more complex than previous datasets, which involve complicated environments (e.g., town, village) and diverse objects (e.g., lake, forest, house, trail). These images are labeled according to PASCAL VOC2007 benchmark. The self-created smoke dataset consists of 9,567 train-validation images and 2,391 test images. In the test set, there are 1,196 images with smoke and 1,195 images without smoke.



a) Town. b) Rual road.

c) Village. d) Field.

Figure 5. The results of smoke detection in different scenarios.

All methods are evaluated by the following widely-used criterion. Concretely, True Positive (TP) is the number of images with smoke detected as smoke images. True Negative (TN) is the number of images that without smoke and detected as non-smoke images. False Positive (FP) is the number of images without smoke but detected as smoke images. False Negative (FN) is the number of images with smoke detected as non-smoke images. As follows, accuracy, precision, and false positive rate are used as evaluation metrics.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$False\ Positive\ Rate = \frac{FP}{TN+FP} \quad (3)$$

4.2. Experiments on Self-Created Dataset

As shown in Table 1, the results of the proposed method and previous method are presented. In [44], due to smoke dataset being insufficient, author uses Faster R-CNN to extract features of synthetic smoke dataset. This method is replicated on the self-created dataset which achieves accuracy, precision, and false positive rates are 90.39%, 86.35%, and 15.38% respectively. For fairness, the proposed method also uses Faster R-CNN for feature extraction. However, different from the above method, the negative sample mining method is incorporated into the proposed method. The experimental results showed that the false positive rate of the proposed method on the test set is 5.76%, which is 9.62% lower than the method reported in [24]. Meanwhile, accuracy and precision are improved by 4.20% and 8.00% respectively. On the whole, the proposed method significantly improves the performance of smoke detection. Figure 6 shows several images in the test set and the detection results in the proposed method.

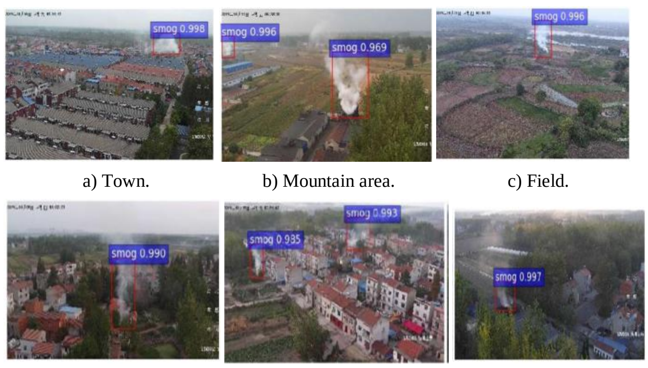
Table 1. Comparison with different smoke detection methods.

Methods	Accuracy	Precision	False Positive Rate
[44]	90.39%	86.35%	15.38%
Our method	**94.59%**	**94.35%**	**5.76%**

To further confirm that the proposed method can reduce false positives, we conducted comparative experiments to compare these methods. As shown in Table 2, the comparison methods are all based on RPN to select proposals. It is obvious that the performance is improved after incorporating negative sample mining methods into these networks. To ensure the authenticity and reliability of the data, both Libra R-CNN and Cascade R-CNN were trained and tested on self-created datasets. In Libra R-CNN, accuracy, precision, and false positive rate are 72.25%, 81.84%, and 12.95% respectively. After incorporating the proposed method, accuracy and precision are improved by 2.15% and 2.94% respectively. Meanwhile, false positive rate is dropped by 2.07%. Before incorporating the proposed method, accuracy, precision, and false positive rate are 84.58%, 96.00%, and 3.06% respectively in cascade R-CNN. After incorporating the proposed method, accuracy and precision are improved by 0.59% and 2.18% respectively. At the same time, false positive rate is decreased by 1.71%. In general, the proposed method can be applied to any networks that select proposals based on RPN. This method can reduce false positive and improve accuracy and precision.

Table 2. Detection results of Libra R-CNN, Cascade R-CNN and our method. NSM means negative sample mining.

Methods	With NSM			Without NSM		
Libra R-CNN [27]	74.40%	84.78%	10.88%	72.25%	81.84%	12.95%
Cascade R-CNN [2]	85.17%	**98.18%**	**1.35%**	84.58%	**96.00%**	**3.06%**
Our method	**94.59%**	94.35%	5.76%	**90.39%**	86.35%	15.38%



a) Town. b) Mountain area. c) Field.

d) Village. e) Courtyard. f) House chimney.

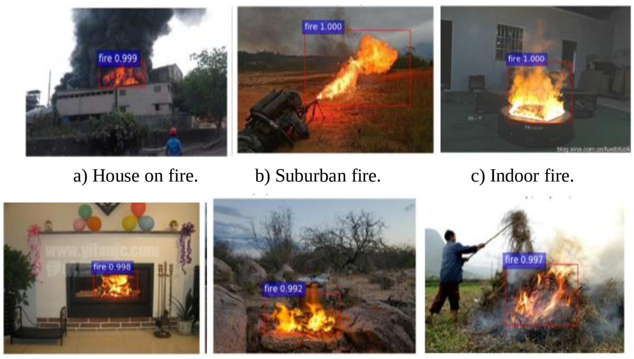
Figure 6. Probability scores and location prediction produced by the proposed method on test set.

4.3. Performance on other Datasets

In order to further verify that the proposed method achieves good performance in other datasets, experiments are conducted on the dataset used in [36]. This dataset is about flame detection and processed according to PASCAL VOC2007 benchmark. Among them, training set has 2059 images and test set has 400 images. Table 3 shows the experimental results, F1-score in [14] achieves 91%, but it is still 2.02% lower than F1-score of the proposed method. Compared with [37], the proposed method is also higher in accuracy and precision, and false positive rate is much lower than it. Figure 7 shows [36] flame dataset and the corresponding detection results.

Table 3. Experiments on the dataset in [35].

Methods	Accuracy	Precision	False Positive Rate	F1-score
[4]	-	-	-	88%
[15]	-	-	-	78%
[43]	-	-	-	91%
[14]	87.75%	81.3%	22.5%	88.87%
Our method	**92.75%**	**89.77%**	**11%**	**93.02%**



a) House on fire. b) Suburban fire. c) Indoor fire.

d) Fireplace. e) Picnic fire. f) Crop burning.

Figure 7. Probability scores and location prediction produced by the proposed method on [36] dataset.

4.4. Ablation Studies

The backbone network adopts RPN strategy to extract proposals. α_0 and α_1 are two key parameters in the RPN which are used to determine thresholds for selecting positive samples and negative samples, respectively. In this paper, the thresholds α0 and α1 are set to 0.7 and 0.3, respectively.

In order to prove that the threshold we set in RPN is optimal, ablation experiment will be used to evaluate the selection of threshold. From previous experiences [31], we set the range of values for α_0 to (0.6, 0.9) and α_1 to (0.1, 0.4) and adopt control variates to search in steps of 0.1. As shown in Figures 8 and 9, when α_0 is 0.7 and α_1 is 0.3, accuracy and precision of smoke detection is the highest. Similarly, the same search strategy is adapted to find optimal parameter values for reducing false positive rate. As shown in Figure 10, the false positive rate is also lowest when α0 is 0.7 and α1 is 0.3.

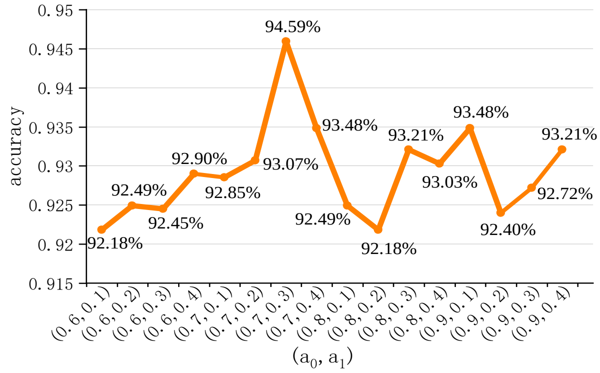


Figure 8. Accuracy of the backbone network.

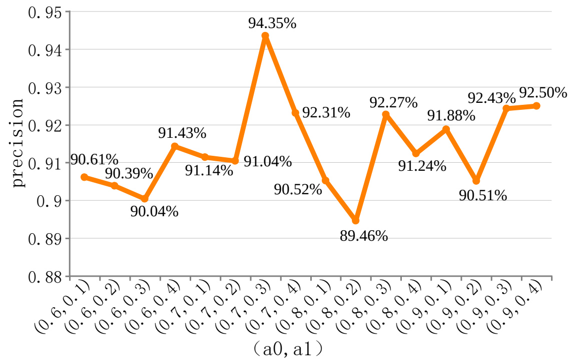


Figure 9. Precision of the backbone network.

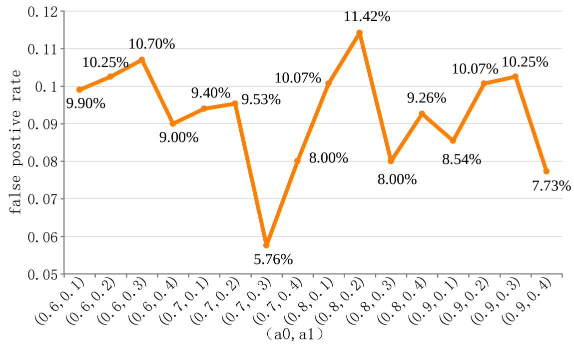


Figure 10. False positive rate of the backbone network.

5. Conclusions

In this paper, we propose a negative sample mining method to reduce the false positive rate in smoke detection. The method is based on the training sample selection rule of RPN, which selects more complex samples as negative samples to adjust the input of the detection network, thus improving the ability of the network to learn positive samples. In addition, a new challenging smoke detection dataset is created which is richer and more authentic than previous datasets. In this dataset, the proposed method outperforms than preceding smoke detection methods, achieving 94.59% accuracy, 94.35% precision and 5.76% false positive rate. Thus, it can be found that the improved network can substantially reduce the false positive rate caused by the interference of smoke-like objects in the real environment.

For now, false positive pictures are selected manually, which is inefficient. In the future, we will use deep learning or machine learning to select false positive pictures automatically.

Acknowledgement

This work was supported by young talents programme of scientific research program of Hubei education depa rtment, No.Q20201709, research on the key technolog y of flexible intelligent manufacturing of garments bas ed on digital twin of Hubei key research and developm ent program, open topic of engineering research center of Hubei province for clothing information, No.90020.

References

[1] Abadi M., Agarwal A., Barham P., and et al., "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," *arXiv preprint arXiv:1603.04467*, 2016.

[2] Cai Z. and Vasconcelos N., "Cascade R-CNN: Delving into High Quality Object Detection," *arxiv.org/abs/1712.00726v1*, 2017.

[3] Cao Y., Yang F., Tang Q., and Lu X., "An Attention Enhanced Bidirectional LSTM for Early Forest Fire Smoke Recognition[J]," *IEEE Access*, vol. 7, pp. 154732-154742, 2019.

[4] Chen T., Wu P., and Chiou Y., "An Early Fire-Detection Method Based on Image Processing," *in Proceedings of International Conference on Image Processing*, Singapore, pp. 1707-1710, 2004.

[5] Chunyu Y., Jun F., Jinjun W., and Yongming Z., "Video Fire Smoke Detection Using Motion and Color Features," *Fire Technology*, vol. 46, no. 3, pp. 651-663, 2010.

[6] Dai J., Li Y., and Sun J., "R-Fcn: Object Detection Via Region Based Fully Convolutional Networks," *Advances in Neural Information Processing Systems*, vol. 29, 2016.

[7] Da-Penha O. and Nakamura E., "Fusing Light and Temperature Data for Fire Detection," *in Proceedings of IEEE Symposium on Computers and Communications*, Riccione, pp. 107-112, 2010.

[8] Deng J., Guo J., Xue N., and Zafeiriou S., "Arcface: Additive Angular Margin Loss for Deep Face Recognition," *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Long Beach, pp. 4690-4699, 2019.

[9] Everingham M., Van Gool L., Williams C., Winn J., and Zisserman A., "The Pascal Visual Object Classes (VOC) Challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303-338, 2010.

[10] Freund Y. and Schapire R., "A Decision-Theoretic Generalization of On-Line Learning and an Application Tto Boosting," *in Proceedings of European Conference on Computational Learning Theory*, Barcelona, pp. 23-37, 1995.

[11] Ge Y., Zhang R., Wu L., Wang X., Tang X., and Luo P., "Deep Fashion2: A Versatile Benchmark for Detection, Pose Estimation, Segmentation and Re-Identification of Clothing Images," *arxiv.org/pdf/1901.07973.pdf*, 2019.

[12] Haykin S., *Kalman Filtering and Neural Networks*, Wiley Online Library, 2001.

[13] He K., Gkioxari G., Dollár P., and Girshick R., "Mask r-cnn," *in Proceedings of the IEEE International Conference on Computer Vision*, Honolulu, pp. 2961-2969, 2017.

[14] He K., Zhang X., Ren S., and Sun J., "Deep Residual Learning For Image Recognition," *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770-778, 2016.

[15] Horng W., Peng J., and Chen C., "A New Image-Based Real-Time Flame Detection Method Using Color Analysis," *in Proceedings of IEEE Networking, Sensing and Control*, Tucson, pp. 100-105, 2005.

[16] Hu Y. and Lu X., "Real-Time Video fire Smoke Detection by Utilizing Spatial-Temporal Convnet Features," *Multimedia Tools and Applications*, vol. 77, no. 8, pp. 29283-29301, 2018.

[17] Jiang B., Luo R., Mao J., Xiao T., and Jiang Y., "Acquisition of Localization Confidence for Accurate Object Detection," *in Proceedings of the European Conference on Computer Vision*, pp. 784-799, 2018.

[18] Johnston F., Henderson S., Chen Y., Randerson J., Marlier M., DeFries R., Kinney P., Bowman DM., and Brauer M., "Estimated Global Mortality Attributable to Smoke From Landscape Fires," *Environmental Health Perspectives*," vol. 120, no. 5, pp. 695-701, 2012.

[19] Kaiser T., "Fire Detection with Temperature Sensor Arrays," *in Proceedings of IEEE 34th Annual International Carnahan Conference on Security Technology*, Ottawa, pp. 262-268, 2000.

[20] Khan S., Muhammad K., Mumtaz S., Baik S., and De-Albuquerque V., "Instituto de Telecomunicações, Aveiro, Portugal Sejong University, Seoul, South Korea "Energy-Efficient Deep CNN for Smoke Detection in Foggy IoT Environment," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 9237-9245, 2019.

[21] Ko B., Kwak J., and Nam J., "Wildfire Smoke Detection Using Temporo-Spatial Features and Random Forest Classifiers," *Optical Engineering*, vol. 51, no. 1, pp. 017208, 2012.

[22] Koman P., Billmire M., Baker K., De Majo R., Anderson F., Hoshiko S., Thelen B., and French

N., "Mapping Modeled Exposure of Wildland Fire Smoke for Human Health Studies in California," *Atmosphere*, vol. 10, no. 6, pp. 308, 2019.

[23] Krizhevsky A., Sutskever I., and Hinton G., "ImageNet Classification with Deep Convolutional Neural Networks," *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097-1105, 2012.

[24] Liu W., Anguelov D., Erhan D., Szegedy C., Reed S., Fu C., and Berg A., "Ssd: Single Shot Multibox Detector," *in Proceedings of European Conference on Computer Vision*, Amsterdam, pp. 21-37, 2016.

[25] Muhammad K., Ahmad J., and Baik S., "Early fire Detection Using Convolutional Neural Networks During Surveillance for Effective Disaster Management," *Neurocomputing*, vol. 288, pp. 30-42, 2018.

[26] Neubeck A. and Van-Gool L., "Efficient Non-Maximum Suppression," *in Proceedings of 18th International Conference on Pattern Recognition*, Hong Kong, pp. 850-855, 2006.

[27] Pang J., Chen K., Shi K., Feng H., Ouyang W., and Lin D., "Libra R-CNN: Towards Balanced Learning for Object Detectionl," *in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, pp. 821-830, 2019.

[28] Phillips W., Shah M., and Lobo N., "Flame Recognition in Video," *in Proceedings 5th IEEE Workshop on Applications of Computer Vision*, Palm Springs, pp. 224-229, 2000.

[29] Pundir A. and Raman B., "Deep Belief Network for Smoke Detection," *Fire Technology*, vol. 53, pp. 1943-1960, 2017.

[30] Redmon J., Divvala S., Girshick R., and Farhadi A., "You Only Look Once: Unified, Real-Time Object Detection," *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779-788, 2016.

[31] Ren S., He K., Girshick R., and Sun J., "Faster R-Cnn: Towards Real-Time Object Detection With Region Proposal Networks," *Advances in Neural Information Processing Systems*, vol. 28, 2015.

[32] Renwei T., Zhongjie Z., Yongqiang B., Ming G., and Zhifeng G., "Key Parts Of Transmission Line Detection Using Improved Yolo V3," *The International Arab Journal Of Information Technology*, vol. 18, no. 6, pp. 747-754, 2021.

[33] Simonyan K. and Zisserman A., "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[34] Sudharshan P., Petitjean C., Spanhol F., Oliveira L., Heutte L., and Honeine P., "Multiple Instances Learning for Histopathological Breast Cancer Image Classification," *Expert Systems with Applications*, vol. 117, pp. 103-111, 2019.

[35] Toreyin B., Dedeoglu Y., and Cetin A., "Wavelet Based Real-Time Smoke Detection in Video," *in Processing of European Signal Processing Conference*, Antalya, pp. 1-4, 2005.

[36] Toulouse T., Rossi L., Campana A., Celik T., and Akhloufi M., "Computer Vision for Wildfire Research: an Evolving Image Dataset for Processing and Analysis," *Fire Safety Journal*, vol. 92, pp. 188-194, 2017.

[37] Vezzani R., Calderara S., Piccinini P., and Cucchiara R., "Smoke detection in video surveillance: The use of ViSOR (Video Surveillance On-line Repository," *in Proceedings of the International Conference on Content-based Image and Video Retrieval*, Niagara Falls, pp. 289-298, 2008.

[38] Wen T., Edelman A., and Gorsich D., "A Fast Projected Conjugate Gradient Algorithm for Training Support Vector Machines," *Contemporary Mathematics*, vol. 323, pp. 245-263, 2001.

[39] Xu G., Zhang Y., Zhang Q., Lin G., Wang Z., Jia Y., and Wang J., "Video Smoke Detection Based On Deep Saliency Network, "*Fire Safety Journal*, vol. 105, pp. 277-285, 2019.

[40] Ye S., Bai Z., Chen H., Bohush R., and Ablameyko S., "An Effective Algorithm to Detect Both Smoke and Flame Using Color and Wavelet Analysis," *Pattern Recognition and Image Analysis*, vol. 27, no. 1, pp. 131-138, 2017.

[41] Yin M., Lang C., Li Z., Feng S., and Wang T., "Recurrent Convolutional Network for Video-Based Smoke Detection," *Multimedia Tools Applications*, vol. 78, no. 8, pp. 237-256, 2019.

[42] Yuan F., "A Fast Accumulative Motion Orientation Model Based on Integral Image for Video Smoke Detection," *Pattern Recognition Letters*, vol. 29, no. 7, pp. 925-932, 2008.

[43] Zhang L., Zhu X., Chen X., Yang X., Lei Z., and Liu Z., "Weakly Aligned Cross-Modal Learning for Multispectral Pedestrian Detection," *in Processing of the IEEE International Conference on Computer Vision*, Seoul, pp. 5127-5137, 2019.

[44] Zhang Q., Lin G., Zhang Y., and Xu G., and Wang J., "Wildland Forest Fire Smoke Detection Based on Faster R-CNN using Synthetic Smoke Images," *Procedia Engineering*, vol. 211, pp. 441-446, 2018.

[45] Zhang Q., Xu J., Xu L., and Guo H., "Deep Convolutional Neural Networks for Forest fire Detection," *in Proceedings of International Forum on Management, Education and Information Technology Application*, pp. 568-575, 2016.

[46] Zhao Y., Zhou Z., and Xu M., "Forest fire Smoke Video Detection Using Spatio-Temporal and Dynamic Texture Features," *Journal of Electrical and Computer Engineering*, 2015.

[47] Zhao Y., "Candidate Smoke Region Segmentation of Fire Video Based on Rough Set Theory," *Journal of Electrical and Computer Engineering*, 2015.

Pei Ma comes from the School of Computer and Artifical Intelligence, Wuhan Textile University. Her major is computer science and technology, and her main research direction is computer vision and object detection.

Fen g Yu is currently a lecturer with the school of computer science and intelligence, Wuhan Textile University. He received the Ph.D. degree with the school of computer science and technology, Huazhong University of Science and Technology. His research interests include machine vision algorithm, artificial intelligence application, and clothing intelligent manufacturing.

Changlong Zhou is currently a experimentalist of Wuhan Textile University. His research interests include computer system architecture, artificial intelligence application, and computer application technology.

Minghua Jiang is currently the vice-chancellor of Wuhan Textile University, and also is a professor with school of computer science and intelligence, Wuhan Textile University. He received the Ph. D. degree from the school of computer science and technology, Huazhong University of Science and Technology. His research interests include computer system architecture, artificial intelligence application, and clothing intelligent manufacturing