International Journal of Innovative Computing, Information and Control
Volume 19, Number 2, April 2023
ICIC International ©2023 ISSN 1349-4198
pp. 355–364

DYNAMIC SMOKE DETECTION BY ELIMINATING STATIC TARGETS IN VIDEO

Hao Chen and Zhe-Ming Lu*

School of Aeronautics and Astronautics
Zhejiang University
No. 38, Zheda Road, Hangzhou 310027, P. R. China
chen_hao_zju@zju.edu.cn; *Corresponding author: zheminglu@zju.edu.cn

Received August 2022; revised December 2022

Abstract. *In this paper, we propose a method to improve the accuracy of traditional object detection algorithm for dynamic objects in video. This paper focuses on the flames and smoke of the real fire. Generally, when detecting fires, we only care about those burning flames, and do not need to detect static flames such as candles, which increase the false rate of our algorithm. This paper compares the performance of three main object detection algorithms in fire detection, and selects the algorithm that satisfies both real-time and accuracy requirements. This paper also proposes and compares two methods of eliminating static targets in video detection. In order to eliminate static flame, we apply the simple linear iterative clustering (SLIC) to the "you only look once" algorithms (YOLOs). We also believe, the algorithm proposed in this paper may be applied to other video detection fields.*
Keywords: Eliminating static targets, Smoke detection, YOLO, Video detection

1. **Introduction.** In recent years, frequent forest wildfires have increasingly aroused public concern about fire disaster. An uncontrolled fire can cause great destructive power and seriously threat human health and property. In 2014 alone, fire disaster losses in the United States amounted to nearly 2% of gross domestic product (GDP) [1]. The frequent occurrence, changeable location and serious damage make fire become one of the most worrying disasters in modern society. Therefore, an effective fire early warning method that can cope with complex environment is particularly important. Consider that smoke, not flames, comes first before a fire. Currently, the most commonly used technology is smoke sensors to detect smoke from fires. However, affected by the performance, only when the smoke concentration is high enough, can trigger the sensor, which greatly reduces the timeliness of fire detection. In addition, smoke sensors cannot detect the location or size of the fire [2]. At the same time, with thousands of cameras installed all over the city, Skynet allows city managers to easily call up massive amounts of video surveillance. Considering the existing large-scale hardware facilities, we can use video detection to detect urban fire, and only need to transplant a set of algorithms, without additional hardware costs.

Nowadays, deep learning plays an increasingly important role in the field of image and video detection. Pioneers in the field of computer vision have proposed a large number of target detection algorithms. In the field of video detection, the most classical algorithms [3-6] usually improve the performance of the algorithm from two perspectives of improving accuracy and reducing detection cost [7]. Even so, these video detection algorithms are still very slow and require very good hardware to work. This requires major improvements to existing monitoring equipments, and these operations still incur high costs. Considering

DOI: 10.24507/ijicic.19.02.355

the above factors, we adopted a more mature object detection algorithm, first divided the video into frames, and then calculated whether each frame contained smoke or flame. Existing target detection algorithms, such as SSD [8], Faster R-CNN [9] and YOLO [10,17] series, are very mature and can meet the requirements of real-time and accuracy at the same time.

However, the biggest problem that object detection algorithms usually face in application is false detection, which detects static flames and clouds as dynamic targets when fire occurs. Therefore, this paper hopes to eliminate the interference of static targets and improve the accuracy of video detection. In this paper, we compare the performance of three main object detection algorithms in smoke detection, and select the algorithm that satisfies both real time and accuracy. To eliminate static targets in video detection, this paper proposes and compares two methods: one is the basic method and the other is the improved method. In the improved one: we apply SLIC to YOLOs to get better result. The algorithm proposed in this paper can be applied to other video detection fields.

In Section 2, we compare the object detection algorithms in smoke detection. In Section 3, we compare two methods that we propose to eliminate the interference of static objects. In Section 4, we introduce the new algorithm combining SLIC. In Section 5, we summarize the whole work and put forward the next research direction.

2. Smoke Detection.

2.1. **Dataset.** For all deep learning projects, test set and train set must be different. However, for most projects, the test set and the train set are separated from the same data set. Therefore, although they are different data, they still belong to a same set. Does the model still work for data outside of the set? In this paper, we use a completely different video dataset to evaluate the recognition effect of our model.

In our project, it is difficult to find a suitable data set from public. Therefore, we organized our own smoke dataset according to the format of VOC. This dataset contains 10,000 smoke or fire images, which are divided into train set, validation set and test set by $8 : 1 : 1$. In addition, in order to research the static objects interference in smoke detection, our project also collected static fire, cloud and dynamic fire videos to test the proposed algorithm. The above video dataset is completely independent of the image dataset.

2.2. **Object detection algorithms.** Among object detection algorithms, there are two mainstream paths: one is the traditional object detection represented by R-CNN, and the other is the regression algorithm represented by YOLO and SSD. In classical object detection algorithms, the output of NN is the probability of each class. Find the region proposals by clustering or cutting and then classify each proposal. Regression algorithms get all the results including bounding box and probability by one neural network (R-CNN needs two NNs). Regression algorithms not only get the probability of each class, but also obtain the diagonal coordinates of bounding box. This is different from the traditional idea that the output result of neural network is the classification probability. The correlation between two paths is proposed in Faster R-CNN, and the two paths can be correlated by Anchor [9].

Table 1 shows the detection effect of mainstream target detection algorithms on large public data sets. By comparing the performance of several mainstream target detection algorithms on COCO data set, it can be seen that of YOLO algorithm, the latest YOLOv5 still has a good performance with the simpler structure compared to R-CNN.

Table 1. The distribution of our dataset. Our dataset is split to four types: non-target, static fire, static smoke and real fire.

All-data	Non-target	Static-fire	Static-smoke	Real-fire
10642	1692	1864	1119	5967

2.3. **Smoke detection algorithms.** The images collected by mainstream large data sets are almost images of fixed shape objects, so the neural network mainly learns the shape between different categories during training, because shape is easier to learn. Fire and smoke are both dynamic objects with no fixed shape. Therefore, smoke detection needs to learn multi-dimensional feature vectors from the input smoke image [14], which can be color, texture, shape, irregularity, flutter or frequency, and classify them into "smoke" or "non-fire" categories.

We use different models to train the image dataset, hoping to get the best recognition effect on the fire and smoke dataset. Using the transfer learning, the model is initialized with the excellent results verified on the large public dataset, such as COCO or VOC, and different training strategies of the neural network are adjusted to obtain the best fitting result of the dataset.

In Table 2, this paper compares several object detection algorithms in fire detection. Considering the requirements of model detection speed in practical application, we mainly focus on the YOLOs, while other deep learning algorithms are also considered for comparison. YOLOs are known for their simplicity and speed. Due to its low requirements on hardware, it is often used as an application-level object detection algorithm. Compared with several mainstream object detection models, although YOLOs only need to look once

Table 2. Comparison of the object detection algorithms in terms of mAP

Algorithms	Backbone	Input size	Training data	AP50 (VOC 2007 test)	AP50 (VOC 2012 test)	AP50 (COCO test-dev)
Faster R-CNN	VGG-16 [9]	512	COCO trainval	–	–	42.7
	VGG-16	512	COCO	76.1	73.0	–
	VGG-16	512	VOC2007	69.9	67.0	–
SSD [8]	VGG-16	512	VOC2007	71.6	–	–
	VGG-16	512	VOC2007 +VOC2012	76.8	74.9	–
	VGG-16	512	VOC2007 +VOC2012 +COCO	81.5	80.0	–
	VGG-16	512	Trainval35k [11]	–	–	46.4
YOLOv3 [12]	DarkNet53	608	COCO	–	–	57.9
YOLOv4 [13]	CSPdarknet53	608	MS COCO	89.0	–	65.7
YOLOv5s	Focus+CSP*5	640	COCO	–	–	56.0
YOLOv5m	Focus+CSP*5	640	COCO	–	–	63.9
YOLOv5l	Focus+CSP*5	640	COCO	–	–	67.2
YOLOv5s6	Focus+CSP*5	1280	COCO	–	–	63.0
YOLOv5m6	Focus+CSP*5	1280	COCO	–	–	69.0
YOLOv5l6	Focus+CSP*5	1280	COCO	–	–	71.6

to get all the results, it seems to be only a crude detection of the objects. However, in our dataset, the effect of smoke and flame is not much different from the Faster R-CNN. In YOLOs, different generations of algorithms show a great difference in detection accuracy. Considering what has mentioned above, our project is mainly based on YOLOv5.

3. **Two Methods of Eliminating Static Targets.** In practical applications, existing object detection algorithms can achieve the effect of real-time dynamic detection, especially the latest YOLO algorithm, which not only detects fast, but also has highly detection accuracy. The biggest problem with image-only training is that the continuity of video detection is not taken into account. In the applications with YOLO only, the problem of false alarm is particularly prominent. Not only does our algorithm recognize the real fire, but it also detects the static flames, such as candles. When a fire starts, the flames and smoke spread quickly, and the flames that burn intensely come in different shapes. A static flame does not spread so quickly. In order to solve the problem of false alarm, it is natural to introduce the continuity of video into fire detection and consider the difference in recognition results of two adjacent frames.

In Figure 1, the shape of a burning flame changes dramatically and the smoke spreads rapidly, while the flame of a quietly burning candle hardly changes shape. This gives us a good angle to eliminate static objects, such as candles.

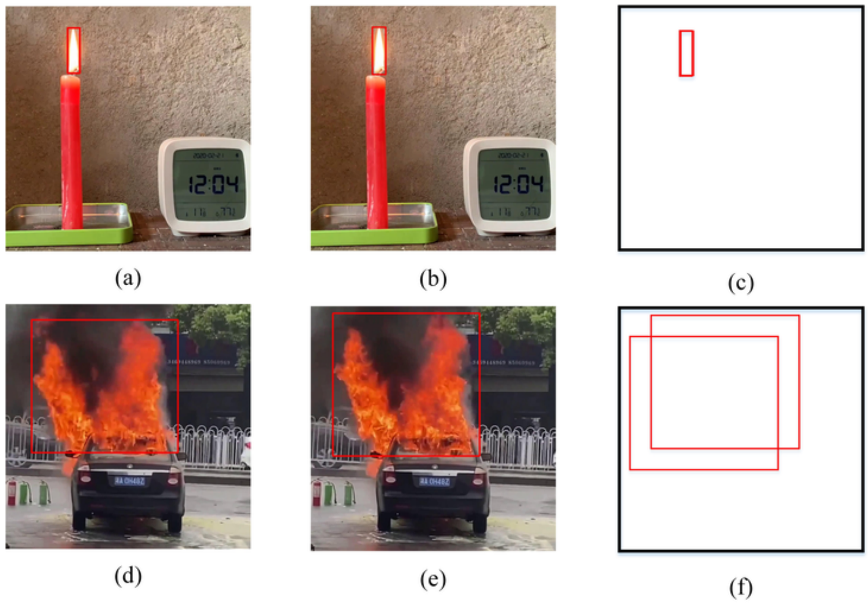

FIGURE 1. IoU comparison of static object and dynamic object: (a) A frame from a candle burning video, (b) the next frame from the same candle burning video, (c) IoU of two ground truths in the candle burning video, (d) a frame from a real fire video, (e) the next frame from the same real fire video, and (f) IoU of two ground truths in the real fire video

3.1. **Basic method.** Inspired by the above phenomenon, we try to judge whether the target is static or dynamic by calculating the IoU of two adjacent frames in the video. The specific idea is to draw the outer box of the object in each frame of the video according to the recognition result of the object detection algorithm, and then calculate the IoU of the target in the adjacent frame. In addition, we employ a threshold to determine whether the target is dynamic or static, referring to paragraph 4.4 for the threshold selection process.

3.2. **Advanced method.** However, the basic method is a rough detection of video, and there is obvious error in judging whether two frames detect the same object by using the bounding box of it. The bounding box does not represent the full range of a target, and a large part of the box is the background of the image. Considering the interference of background, we try to judge whether the target is static by objects' mask. In the first scheme, we calculate the IoU of the bounding box, and in the second scheme, we calculate the IoU of the objects' mask.

In the field of image processing, there are many excellent image segmentation algorithms, the most famous of which is Mask R-CNN. Mask R-CNN inherits many structures of Faster R-CNN, such as Anchor and RPN. However, the two-stage algorithm generally has the problem of complex algorithm and slow running speed. Therefore, it is obviously unwise to add Mask R-CNN to YOLOs for eliminating the interference of static objects on detection. We choose SLIC, which is more efficient. This machine-learning algorithm is added to YOLOs to improve the accuracy of YOLOs without greatly reducing the speed of the algorithm.

We compare the accuracy of different algorithms for three different types of video with different static elimination modules. As shown in Table 3, when there is almost no difference between the initial YOLOs in terms of accuracy, the more complex network has the stronger anti-interference ability. In addition, the two algorithms, YOLOv5s6 and YOLOv5l6, can resist the interference of static smoke naturally. In this algorithm, the threshold of IoU is 0.8, and only when the overlap of bounding box between two adjacent frames is more than 80% can be regarded as a static object. From Table 3, we can see that both two schemes can effectively eliminate the interference of static flames and smoke, especially combined with the algorithm that has more complex structure.

Table 3. Comparison of the mAP between different object algorithms in fire detection

Algorithms	**Backbone**	**Input size**	**Training data**	**AP50 (val set)**	**AP50 (test set)**
R-CNN	ResNet-50	640	Trainval set	72.49	61.76
SSD	VGG-16	512	Trainval set	–	70.69
YOLOv3	DarkNet53	416	Train set	64.19	44.15
YOLOv4	DarkNet53	416	Train set	–	65.4
YOLOv5s	Focus+CSP$*$5	640	Train set	76.38	71.5
YOLOv5m	Focus+CSP$*$5	640	Train set	76.35	71.5
YOLOv5l	Focus+CSP$*$5	640	Train set	75.23	70.9
YOLOv5s6	Focus+CSP$*$5	640	Train set	75.83	72.2
YOLOv5m6	Focus+CSP$*$5	640	Train set	76.22	72.7
YOLOv5l6	Focus+CSP$*$5	1280	Train set	76.26	73.5

4. **Algorithms for the Advanced Method.** This section will detail how to add the SLIC to deep learning. The flow of two methods is shown in Figure 2.

4.1. **SLIC.** The core of this algorithm is to segment each frame of video. However, adding a new image segmentation algorithm based on deep learning will inevitably increase the computational complexity. Therefore, it is important to find an efficient image segmentation algorithm. Considering that image segmentation algorithms based on deep learning, such as Mask R-CNN are too complex, the mainstream superpixel algorithms are divided into two paths, including graph-based and cluster-based. In the work of Wang et al., they

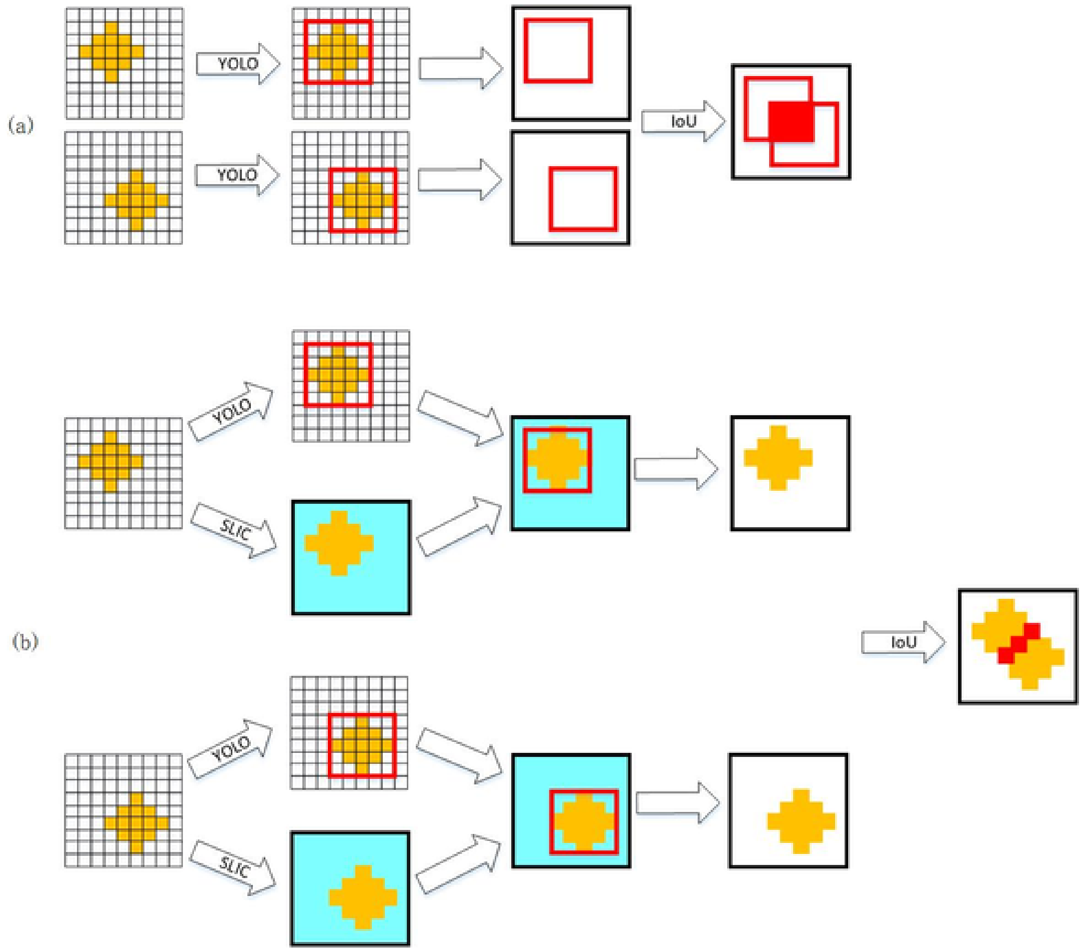

FIGURE 2. Comparison of two static object elimination algorithms: (a) Flow chart of the basic method; (b) flow chart of the advanced method

designed a standard to evaluate the image segmentation effect and compared the segmentation effect of some typical image segmentation methods [15]. From their research, the cluster-based superpixel segmentation algorithms are more efficient than the graph-based superpixel segmentation algorithms.

Nowadays, SLIC is the most widely used superpixel segmentation algorithm based on clustering. Compared with the traditional clustering algorithm, the biggest improvement of SLIC lies in two points: the reduction of the search range and the improvement of the distance measure function. In SLIC, the search range for the next anchor is narrowed down from the entire image to a rectangular area centred on the old anchor. The side of rectangle is twice of the size of segmentation. Reducing the research range increases the number of iterations, but also reduces the computational complexity of each iteration. After balancing the two algorithms, the computational complexity of the whole algorithm is greatly reduced compared with the original clustering algorithm. In addition, based on the traditional distance measurement, the algorithm adds the colour space measurement of $[l\ a\ b]^T$. The accuracy of clustering is greatly improved [16].

$$d_c = \sqrt{(l_i - l_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2} \tag{1}$$

$$d_s = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{2}$$

$$D = \sqrt{\left(\frac{d_c}{N_c}\right)^2 + \left(\frac{d_s}{N_s}\right)^2} \tag{3}$$

where d_c is the metric formula of colour space, $[l\ a\ b]$ represents a pixel point in the space; d_s is the distance formula between pixels; N_c, N_s are used to balance the two distances.

4.2. **YOLOs.** In our work, we adopt YOLO algorithm as the body algorithm of our method. The core of YOLO is detecting objects in multi-scale with multi branch, which contributes to detecting objects with different scales. Cross Stage Partial Network (CSP-Net) is widely used in YOLOv5 algorithm, which is also one of the newest algorithms in YOLO family.

Multiple multi-stage fusions guarantee the completely fusing between different scales, which is effective to multi-scale detection. The output of each branch contains all the results, including the coordiates of bounding boxes (BBoxes) and probability that objects belong to different classes. The comparison of different object detection algorithms is shown in Table 2.

Considering the effectiveness and efficiency of YOLOs in object detection, we choose YOLOv5 as the body of our static object elimination algorithm.

4.3. **Adding SILC to YOLO.** Obviously, the basic method in Section 3 is the simplest application of this idea. It directly calculates the IoU of the boxes obtained by YOLOs between two frames of the video to judge whether the detected object is a static object.

Based on the basic method in Section 3, SLIC is added into the detection of YOLOs by combining machine learning and deep learning. Every frame in the video is detected by YOLO to get the position information of bounding box. We segment each frame to get the objects' mask from the segmentation result. SLIC is used for image segmentation, and the superpixel algorithm based on clustering has higher efficiency. Then the mask of object is inversely selected by combining the position information of bounding box with the result of image segmentation. The bounding box is covered on the block of image segmentation,

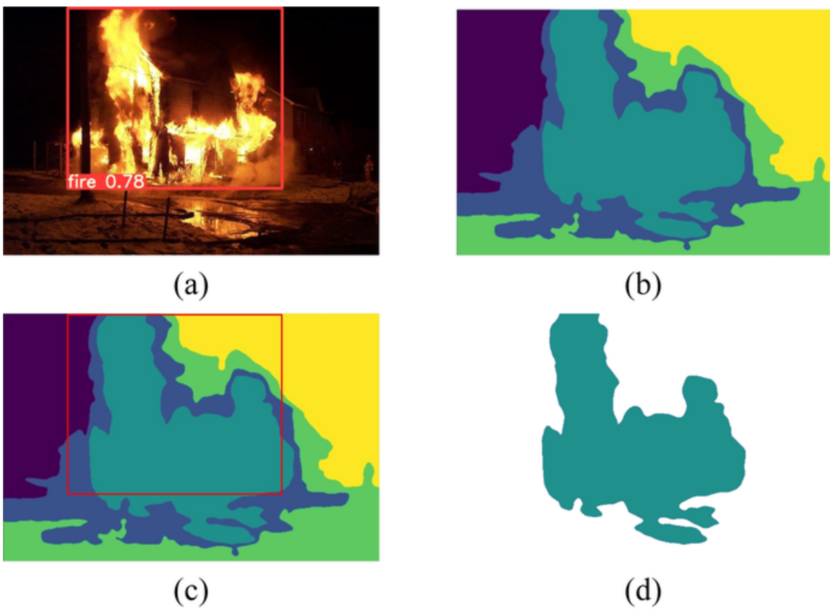

FIGURE 3. An example of obtaining the object mask by YOLO: (a) The detection result from YOLO, (b) the segmentation result from SLIC, (c) the combination of (a) and (b), and (d) the object mask from (c)

counting the number of pixels of each segmentation in the range of bounding box, and the block with the most pixels is taken as the mask of the object. In this way, we get the object masks of every two adjacent frames in video, and calculate the IoU between two masks. If the IoU is greater than the threshold, it is considered as a static target. Otherwise, it is retained as a dynamic target. An example is shown in Figure 3.

4.4. **Threshold.** Considering the continuity of video, if the static threshold is not changed during the whole video, it will surely make the whole task ineffective. Here, we use the

Table 4. Comparison of the video detection result with different static object elimination algorithms

Algorithms	**Backbone**	**Test video**	**Not eliminate static object**	**Eliminate with box**	**Eliminate with mask**
R-CNN	ResNet-50	Static fire	97.30%	65.30%	35.60%
		Static smoke	92.20%	45.90%	39.30%
		Real fire	95.50%	53.80%	66.80%
SDD	VGG-16	Static fire	94.50%	54.60%	27.20%
		Static smoke	96.70%	43.90%	23.60%
		Real fire	99.60%	54.30%	87.40%
YOLOv3	DarkNet53	Static fire	97.80%	12.40%	5.30%
		Static smoke	99.20%	50.80%	36.40%
		Real fire	98.10%	71.30%	89.20%
YOLOv4	DarkNet53	Static fire	98.70%	26.30%	0
		Static smoke	100%	30.80%	25.50%
		Real fire	100%	63.90%	90.20%
YOLOv5s	Focus+CSP∗5	Static fire	100%	0	0
		Static smoke	100%	44.20%	67.50%
		Real fire	100%	54%	94%
YOLOv5m	Focus+CSP∗5	Static fire	94.10%	0	0
		Static smoke	100%	5.80%	47.40%
		Real fire	100%	51.90%	74.80%
YOLOv5l	Focus+CSP∗5	Static fire	92.80%	3.90%	0
		Static smoke	94.80%	7.80%	6.50%
		Real fire	99.20%	77.80%	81.50%
YOLOv5s6	Focus+CSP∗6	Static fire	100%	7.40%	0
		Static smoke	0	0	0
		Real fire	100%	68.90%	77%
YOLOv5m6	Focus+CSP∗6	Static fire	99.30%	8.50%	1.30%
		Static smoke	100%	9.10%	16.20%
		Real fire	100%	77%	84.40%
YOLOv5l6	Focus+CSP∗6	Static fire	100%	0	0
		Static smoke	6.50%	5.80%	3.90%
		Real fire	100%	65.20%	81.50%

In the above table, we use three different types of video for testing, namely static flame (candle), static smoke (cloud) and dynamic flame (real fire). We hope that the recognition rate of static flame and smoke should be as low as possible, while that of real fire should be as high as possible.

dynamic threshold selection mechanism, first we design an initial threshold value based on experience, and then use the dynamic threshold judgment mechanism. The appearance of dynamic and static targets has certain continuity in the video. The IoU between two adjacent frames is always relatively small when detecting a dynamic object, and relatively large when detecting a static object. Two queues of size N are maintained, and the IoU judged to be dynamic or static objects is pushed into the queues, and the average value of the two queues' mean values is taken as the new threshold.

In the application, considering that the occurrence of real fire is rare, it is difficult to dynamically correct the threshold value. Once a target judged to be a dynamic flame or smoke appears, an alarm will be issued immediately.

5. **Conclusion and Future Work.** In this paper, we use the traditional object detection algorithm to detect smoke and flame in fire, and propose a new method to eliminate static objects in fire detection. In order to eliminate the interference of candles, clouds or other static objects, this paper considers the continuity of video frames, calculates the IoU of objects in adjacent frames, and determines whether the object is static or not by comparing it with a threshold. Our basic idea is calculating the IoU between bounding boxes in adjacent frames. Considering that the bounding box is just a rough description, this paper uses masks to calculate the IoU of adjacent frames. In order to extract objects' mask quickly and accurately, this paper adds the machine learning method into YOLO, and uses SLIC to extract objects' mask, achieving better results.

This paper combines machine learning with deep learning and proposes a method to detect dynamic targets in video by traditional object detection algorithm. We believe that the methods in this paper are not limited to fire detection tasks, but can be extended to other tasks, and this method can achieve the same excellent results for other static objects.

Acknowledgment. This work is supported in part by the National Key Research and Development Program of China under Grant No. 2020AAA0140004. This research work is also partially supported by Ningbo Science and Technology innovation 2025 major project under Grant No. 2021Z010.

REFERENCES

[1] J. R. Hall, *The Total Cost of Fire in the United States*, National Fire Protection Association, Quincy, MA, 2014.
[2] S. Saponara, A. Elhanashi and A. Gagliardi, Real-time video fire/smoke detection based on CNN in antifire surveillance systems, *Journal of Real-Time Image Processing*, vol.18, no.3, pp.889-900, 2021.
[3] X. Zhu, Y. Wang, J. Dai, L. Yuan and Y. Wei, Flow-guided feature aggregation for video object detection, *Proc. of the IEEE International Conference on Computer Vision*, pp.408-417, 2017.
[4] X. Zhu, J. Dai, X. Zhu, Y. Wei and L. Yuan, Real-time video fire/smoke detection based on CNN in antifire surveillance systems, *Journal of Real-Time Image Processing*, vol.18, no.3, pp.889-900, 2021.
[5] W. Yang, B. Liu, W. Li and N. Yu, Tracking assisted faster video object detection, *2019 IEEE International Conference on Multimedia and Expo (ICME)*, Shanghai, China, pp.1750-1755, DOI: 10.1109/ICME.2019.00301, 2019.
[6] K. Kang, H. Li, J. Yan, X. Zeng, B. Yang, T. Xiao and W. Ouyang, T-CNN: Tubelets with convolutional neural networks for object detection from videos, *IEEE Trans. Circuits and Systems for Video Technology*, vol.28, no.10, pp.2896-2907, 2017.
[7] H. Zhu, H. Wei, B. Li, X. Yuan and N. Kehtarnavaz, A review of video object detection: Datasets, metrics and methods, *Applied Sciences*, vol.10, no.21, 7834, 2020.
[8] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu and A. C. Berg, SSD: Single shot multibox detector, *European Conference on Computer Vision*, pp.21-37, 2016.

[9] S. Ren, K. He, R. Girshick and J. Sun, Faster R-CNN: Towards real-time object detection with region proposal networks, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.39, no.6, pp.1137-1149, 2016.
[10] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, You only look once: Unified, real-time object detection, *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.779-788, 2016.
[11] S. Bell, C. L. Zitnick, K. Bala and R. Girshick, Inside-Outside Net: Detecting objects in context with skip pooling and recurrent neural networks, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp.2874-2883, DOI: 10.1109/CVPR.2016.314, 2016.
[12] J. Redmon, and A. Farhadi, YOLOv3: An incremental improvement, *arXiv Preprint*, arXiv: 1804. 02767, 2018.
[13] A. Bochkovskiy, C. Y. Wang and H. Y. M. Liao, YOLOv4: Optimal speed and accuracy of object detection, *arXiv Preprint*, arXiv: 2004.10934, 2020.
[14] S. Saponara, L. Pilato and L. Fanucci, Early video smoke detection system to improve fire protection in rolling stocks, *Real-Time Image and Video Processing*, vol.9139, pp.14-22, 2014.
[15] M. Wang, X. Liu, Y. Gao, X. Ma and N. Q. Soomro, Superpixel segmentation: A benchmark, *Signal Processing: Image Communication*, vol.56, pp.28-39, 2017.
[16] T. L. Dang, G. T. Nguyen and T. Cao, SLIC superpixels compared to state-of-the-art superpixel methods, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.34, no.11, pp.2274-2282, 2012.
[17] T. L. Dang, G. T. Nguyen and T. Cao, Object tracking using improved deep_sort_YOLOv3 architecture, *ICIC Express Letters*, vol.14, no.10, pp.961-969, 2020.

Author Biography

Hao Chen received the B.S. degree in communication engineering from Shandong University, China, 2021. He is currently pursuing the master's degree in Zhejiang University, China. His current research interests include natural language processing, self-supervised learning, and computer vision.

Zhe-Ming Lu received the B.S. and M.S. degrees in electrical engineering and the Ph.D. degree in measurement technology and instrumentation from the Harbin Institute of Technology (HIT), Harbin, China, in 1995, 1997, and 2001, respectively. He became a Lecturer with HIT in 1999. Since 2003, he has been a Professor with the Department of Automatic Test and Control, HIT. He is currently a Full Professor with the School of Aeronautics and Astronautics, Zhejiang University, Hangzhou, P. R. China. In the areas of multimedia signal processing and information hiding, he has published more than 350 papers, seven monographs in Chinese, two monographs in English and three book chapters in English. His current research interests include multimedia signal processing, information security, and complex networks.