A Novel Fire Identification Algorithm Based on Improved Color Segmentation and Enhanced Feature Data

Xijiang Chen, Qing An, Kegen Yu, *Senior Member, IEEE*, and Ya Ban

***Abstract*— In order to improve the accuracy of fire identification based on video in the Internet-of-Things environment, this article proposes a new fire identification algorithm by merging fire segmentation and multifeature fusion of fire. First, according to the relationship between R and Y channels, the improved YCbCr models are established for initial fire segmentation under reflection and nonreflection conditions, respectively. Simultaneously, the reflection and nonreflection conditions are judged by comparing the areas obtained by the two improved YCbCr models. Second, an improved region growing algorithm is proposed for fine fire segmentation by making use of the relationship between the seed point and its adjacent points. The seed points are determined using the weighted average of centroid coordinates of each segmented image. Finally, the quantitative indicators of fire identification are given according to the variation coefficient of fire area, the dispersion of centroid, and the circularity. Extensive experiments were conducted, and the experimental results demonstrate that the proposed fire detection method considerably outperforms the traditional methods on average in terms of three performance indexes: precision, recall, and $F1$-score. Specifically, compared with the deep learning method, the precision of the proposed method is slightly higher. Although the recall of the proposed method is slightly lower than the deep learning method, its computation complexity is low.**

***Index Terms*— Fire detection, multifeature fusion, regional growth, YCbCr model.**

I. Introduction

FIRE detection mainly makes use of the chemical, ionizing, or thermal sensor. One drawback of such fire detectors is that they are expensive and not able to pinpoint the place of fire formation. In recent years, with the popularization of smarter surveillance cameras in the field of real-time video analysis, i.e., objects detection [1] and object tracking [2], many scholars investigated the use of video camera images for fire detection [3]. The shape of the fire and the rate of fire spread can be obtained by using the cameras distributed at ground stations or mounted on unmanned aerial systems (UASs) [4]. Literature [5] introduced several fire detection methods, mainly based on color detection and moving object detection. The color-based fire detection is the most popular detection technique, including the algorithms based on color spaces [6], [7], HSV [8], L^*a^*b [9], YUV [10], and YCbCr [11]. The shortcoming of this method is that it can be easily affected by the fire-like objects. Due to the moving characteristic of the fire, the moving object detection is also used in the fire detection, including background subtraction method [12], [13], optical flow analysis method [14], and temporal differencing method [15]. The foreground information is extracted using an adaptive background subtraction algorithm, which is then verified using a statistical fire color model. For the background subtraction method, the Gaussian mixture model was often used to perform the background modeling [16]. However, the background subtraction method requires the contrast between the background and the moving object to be greater than a certain threshold. The optical flow analysis and the temporal differencing method can only detect the dynamic fire under the ideal condition. Even when the object does not move, the optical flow can be detected when the external light changes, and then, the object will be detected as fire. Also, the dynamic fire cannot be detected when lacking sufficient gray gradient variation area. In addition, there are many dynamic fire-like objects in outdoor scene, so they are not suitable for outdoor fire detection. The flame flickering of uncontrolled fire can be used to distinguish the fire from ordinary objects according to the variation characteristics of flame flicker [13]. A lot of smoke will be produced in the presence of fire, so the fire can be detected through analysis of the temporal behavior of smoke by the wavelet domain energy [17]. However, this approach cannot distinguish objects that appear gray in color from smoke or provide objective experimental results in dynamic fire situations. Meanwhile, if the fire does not produce a lot of smoke, the method cannot accurately detect it.

As introduced above, fire detection is mainly based on the characteristics of fire and smoke, which include the characteristics of color, motion, and geometrical contour of flame and smoke. Due to the difference of color between the flame and the objects around the fire, the color information was

Manuscript received March 19, 2021; accepted April 4, 2021. Date of publication April 26, 2021; date of current version May 6, 2021. This work was supported in part by the Chongqing Technological Innovation and Application Development under Program cstc2019jscx-msxmX0051 and in part by the CRSRI Open Research Program under Program CKWV2019758/KY. The Associate Editor coordinating the review process was Dr. Lei Zhang. *(Corresponding author: Xijiang Chen.)*

Xijiang Chen is with the School of Artificial Intelligence, Wuchang University of Technology, Wuhan 430223, China, and also with the School of Safety Science and Emergency Management, Wuhan University of Technology, Wuhan 430070, China (e-mail: cxj_0421@163.com).

Qing An is with the School of Artificial Intelligence, Wuchang University of Technology, Wuhan 430223, China (e-mail: 490756729@qq.com).

Kegen Yu is with the School of Environment Science and Spatial Informatics, China University of Mining and Technology, Xuzhou 221116, China (e-mail: kegen.yu@cumt.edu.cn).

Ya Ban is with the Chongqing Academy of Metrology and Quality Inspection, Chongqing 401121, China (e-mail: 573469578@qq.com).

Digital Object Identifier 10.1109/TIM.2021.3075380

used to judge the fire. For example, Philips, III, *et al.* [18] performed the fire identification according to the temporal variation of color information of images extracted from the video sequences. Chen *et al.* [19] presented the RGB model for extracting fire and smoke. They used an RGB/hue, intensity, saturation (HIS) color model and a dynamic analysis of flames that matches the disordered characteristic of flames to the growth of pixels to check for the existence of fire. However, because they measure the frame difference between two consecutive frames, the decision rule cannot distinguish real fire regions from moving regions or from noise. As a result, this kind of fire identification method is not stable. In addition to the color information, the flame flicker that can be detected by the hidden Markov model is used to segment the fire [20]. On this basis, the color model, motion information, and the fire flicker analysis based on the Markov model are combined together to detect the fire [21]. The shortcoming of this type of method is that they do not analyze the characteristic of variation of fire. Therefore, it is easy to regard the fire-like moving object as fire, such as dynamic light, swaying leaves, and so on. To improve the detection performance, a real-time detection method is proposed that combines the foreground object information with the color pixel of fire [6]. This kind of method has high accuracy in identifying the fire that is rapidly spreading. However, it is easily affected by the fire-like regions and dynamic objects. When using video stream for fire monitoring, the fire intensity is determined by a visual image sensor. The fire can be detected automatically mainly based on the temporal variation of fire intensity [10]. However, some intensity of fire-like regions is similar to that of fire, which affects the judgment of fire by the method. In order to decrease the influence of fire-like region and improve the precision of fire detection, the HSI color space model is used to distinguish the fire from fire-like region for brighter environments [22]. Celik and Demirel [11] used the YCbCr color space model to separate the luminance from the chrominance, which segments fire more effectively than RGB color spaces. Simultaneously, the CIE L^*a^*b color space is proposed to detect the fire pixels [9]. Although this type of method can distinguish a fire-like region from the identification area, it is still affected by the object whose color is very similar to the fire.

The fire detection based on the color space model is easily influenced by the ordinary fire-colored objects. In view of this, the variation of spatial color was used to distinguish ordinary fire-colored objects from uncontrolled fires [23]. However, this method is not applicable in the field of smoke detection because the high variation of spatial color is not shown in the smoke regions. In addition, these features are difficult to define and depend largely on the kind of fire. Accordingly, these methods will lead to low-precision detection. In view of this, a convolutional neural network (CNN) is proposed to perform feature extraction to identify fire with video images [24]. Khan *et al.* [25] also used the deep learning method for smoke detection in both normal and foggy environments. Because deep learning requires large memory space, a computationally efficient CNN architecture based on the SqueezeNet architecture was proposed for fire detection [26]. Considering the characteristic of fire, all the fire color rules are used as the input of machine learning, and the detection results are compared with the rule-based method and the machine learning method performs better than the other rule-based methods [27]. The deep learning-based fire detection method using faster region-based CNN (Faster-RCNN) was proposed to detect suspected regions of fire [28]. A number of methods used the neural network to detect the fire in the road and rail tunnels [29]. However, the efficiency of detection is mainly influenced by the distance of the vehicle and charge-coupled device (CCD) color camera. The disadvantage of this method is that a large fire data set needs to be constructed, and it requires a large memory space. Simultaneously, deep learning does not consider the characteristic of fire. Consequently, it cannot determine the trend of fire development and distinguish the fire and the fire-like objects (such as candle flame, light, and so on).

The main contributions of this article are summarized as follows.

1) The relationship model between the R channel in RGB and the Y channel in YCbCr is used to improve the YCbCr model, so as to avoid the influence of reflection and nonreflection on fire segmentation.
2) The improved region growing algorithm is developed by weighting the centroid of each segmented image, and this improved method not only accurately segments the fire but also eliminates nearly all the noises.
3) The fusion of the variation of area, the dispersion of centroid, and the circularity is proposed for fire identification. Furthermore, the accurate fire identification can be realized without the requirement of large memory space due to its video-clip-based processing.

The remainder of this article is organized as follows. Section II presents the proposed fire detection method in detail, including the flowchart of the proposed method, the fire segmentation, and fire identification based on multifeature description. Section III provides the performance evaluation results using a large amount of experimental data. Finally, Section IV concludes this article.

II. Proposed Method

This article proposed the fire identification method based on the fire segmentation and multifeature fusion in surveillance videos. First, the initial segmentation of fire by improved YCbCr is obtained. The segmentation of fire can be achieved under the condition of reflection and nonreflection according to the comparison between the R channel in the RGB color model and the Y channel in the YCbCr color model. Second, the seed points are accurately determined by weighting the centroid of each connected region. The improved region growing algorithm is developed according to the analysis of seed points and the eight adjacent connected pixels. Third, the fire identification is conducted according to the variation coefficient of area, the dispersion of centroid, and the circularity. The flowchart of fire identification in the proposed method is shown in Fig. 1.

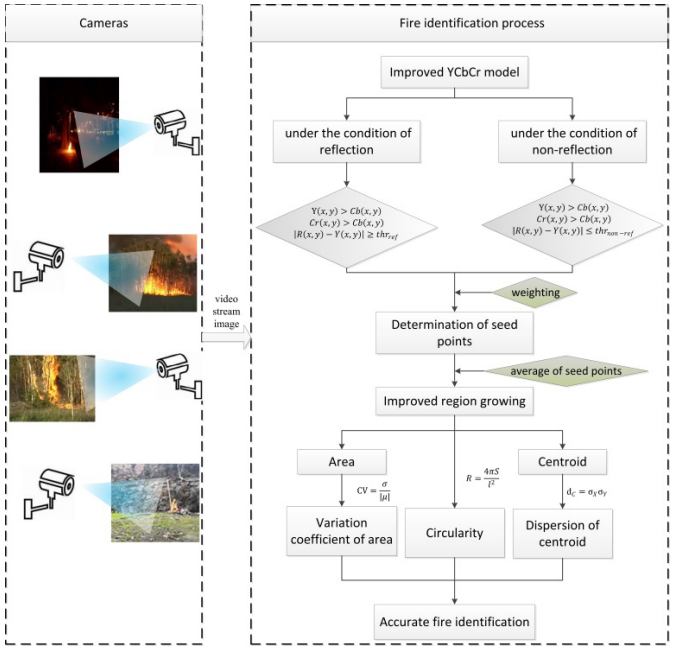

Fig. 1. Flowchart of fire identification.

Algorithm 1 Initial Segmentation of Fire by the Improved YCbCr Model

Input: Fire image

1. Improved YCbCr under the condition of reflection:
$$\begin{cases} Y(x,y) > Cb(x,y) \\ Cr(x,y) > Cb(x,y) \\ |R(x,y) - Y(x,y)| \le thr_{ref} \end{cases}$$
2. Improved YCbCr under the condition of non-reflection
$$\begin{cases} Y(x,y) > Cb(x,y) \\ Cr(x,y) > Cb(x,y) \\ |R(x,y) - Y(x,y)| \ge thr_{non-ref} \end{cases}$$
3. Judgement of non-reflection and reflection

We use the improved YCbCr under the condition of reflection and non-reflection to segment fire, and obtain the segmented fire area S_{ref} and S_{non}.

If $S_{ref} > S_{non}$, fire non-reflection

If $S_{ref} < S_{non}$, fire reflection

4. Judgement of non-reflection and reflection

Output: Initial fire segmentation results

A. Initial Segmentation of Fire

The segmentation of fire image is often influenced by the surrounding reflection light. Therefore, we consider the segmentation of fire under the condition of reflection and nonreflection.

YCbCr is a digital color system; it is one kind of color space that is used in digital video [30]. This kind of family space can be elicited from RGB by using the following transformation matrix [31]:

$$\begin{bmatrix} \mathrm{Y} \\ \mathrm{Cb} \\ \mathrm{Cr} \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (1)$$

Fig. 2. Original fire. (a) Fire without reflection. (b) Fire with reflection.

The YCbCr color model not only describes luminance information but also considers the color saturation. The three components of the YCbCr color space, which are Y, Cb, and Cr components, represent luminance, blue, and red component, respectively. The segmentation model based on YCbCr color space is described as [31]

$$\begin{aligned} &\text{Rule1}: \mathrm{Y}(x,y) > \mathrm{Cb}(x,y) \\ &\text{Rule2}: \mathrm{Cr}(x,y) > \mathrm{Cb}(x,y) \\ &\text{Rule3}: \mathrm{Y}(x,y) > \mathrm{Y}_{\mathrm{mean}} \\ &\text{Rule4}: \mathrm{Cb}(x,y) < \mathrm{Cb}_{\mathrm{mean}} \\ &\text{Rule5}: \mathrm{Cr}(x,y) > \mathrm{Cr}_{\mathrm{mean}} \\ &\text{Rule6}: |\mathrm{Cb}(x,y) - \mathrm{Cr}(x,y)| \ge \tau. \end{aligned} \quad (2)$$

In order to verify that the YCbCr model can be used in the fire segmentation under different environments, this article constructs the fire segmentation model under the condition of reflection and nonreflection, as shown in Fig. 2. It can be seen that there is no reflection around the flame in Fig. 2(a), but there is reflection around the fire in Fig. 2(b).

The two fire images in Fig. 2 are segmented and the results are shown in Fig. 3. The flame foreground region extracted by the YCbCr color model is still greatly influenced by noise. Fig. 3(a) shows a better fire segmentation under the condition of nonreflection and has eliminated most noise components. However, it is greatly affected by the red object near the fire. The red wooden bench is misidentified as fire, as shown in the red solid rectangle of Fig. 3(a). The possible reason for this phenomenon is that the characteristics of the flame are not considered. Fig. 3(b) shows that the extraction effect of fire edge contour is not good. The segmentation of fire in the reflection region is greatly influenced by the luminance, and many surrounding luminance areas are identified as fire. The reason for this phenomenon is that the relationship between luminance and fire color is not considered.

In view of this, we need to distinguish the luminance and fire color information in the process of fire image segmentation. The luminance of reflection region is high, and the nonreflection is low. Therefore, we consider the segmentation of fire image under the condition of reflection and nonreflection. We use the components of R channel in the RGB model minus the components of the Y channel in the YCbCr model and then apply the different relationship between the R channel in RGB and the Y channel in YCbCr to improve the YCbCr model. We consider the improved YCbCr under the condition of reflection and nonreflection.

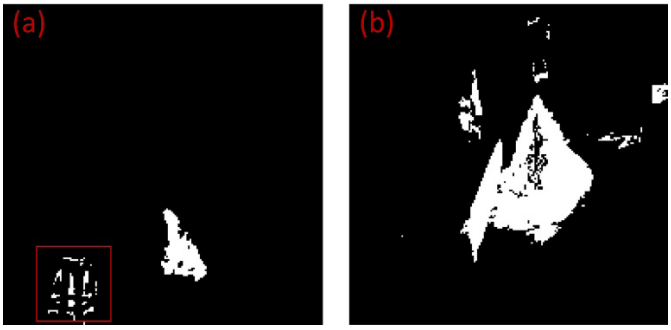

Fig. 3. YCbCr color space model. (a) Fire segmentation under the condition of nonreflection in Fig. 2(a). (b) Fire segmentation under the condition of reflection in Fig. 2(b).

In the presence of reflection, the red component (R) of fire is close to the luminance component (Y), and R and Y take a large proportion in the color model under the condition of reflection. Therefore, the segmentation model of fire can be described as

$$\begin{cases} \mathrm{Y}(x,y) > \mathrm{Cb}(x,y) \\ \mathrm{Cr}(x,y) > \mathrm{Cb}(x,y) \\ |R(x,y) - \mathrm{Y}(x,y)| \le \mathrm{thr}_{\mathrm{ref}} \end{cases} \tag{3}$$

where $\mathrm{thr}_{\mathrm{ref}}$ is the threshold value of difference between the R and Y channels under the condition of reflection and (x, y) is the spatial location of the fire pixel.

In the absence of nonreflection, the red component (R) of fire can be greater or smaller than the luminance component (Y). Only R or Y but not both take a large proportion in the color model. Therefore, the segmentation model of fire can be described as

$$\begin{cases} \mathrm{Y}(x,y) > \mathrm{Cb}(x,y) \\ \mathrm{Cr}(x,y) > \mathrm{Cb}(x,y) \\ |R(x,y) - \mathrm{Y}(x,y)| \ge \mathrm{thr}_{\mathrm{non-ref}} \end{cases} \tag{4}$$

where $\mathrm{thr}_{\mathrm{non-ref}}$ is the threshold value under the condition of nonreflection.

The segmentation of fire image can be achieved under the condition of reflection or nonreflection according to (3) and (4). In order to determine the threshold, the different segmentation accuracies under the condition of different thresholds are calculated

$$P = 100\% - \frac{|S_1 - S_2|}{S_1} \times 100\% \tag{5}$$

where S_1 is the actual fire area and S_2 is the segmented fire area.

Fig. 4 shows the segmentation accuracy with respect to the threshold, indicating that the segmentation accuracy is greatly influenced by the threshold value. The segmentation accuracy is smaller than 60% when the threshold is less than 35 and greater than 55 under the condition of nonreflection. On the other hand, the segmentation accuracy is smaller than 60% when the threshold is less than 30 and greater than 45 under the condition of reflection. The segmentation results can be regarded as reliable segmentation when the segmentation accuracy is greater than 60% [32]. Therefore,

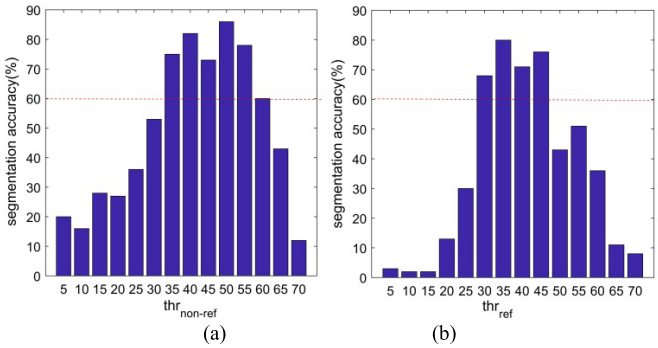

Fig. 4. Segmentation accuracy with different thresholds. (a) Under the condition of nonreflection. (b) Under the condition of reflection.

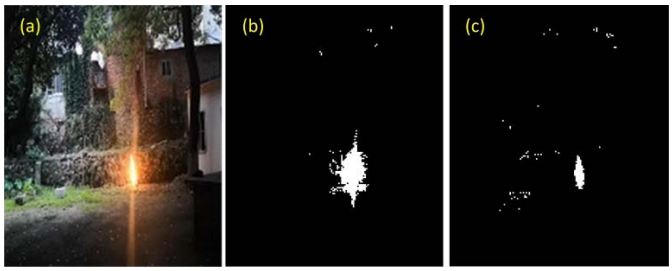

Fig. 5. Fire segmentation under the condition of reflection. (a) Original fire image. (b) Segmentation result by the nonreflection-based YCbCr method. (c) Segmentation result by the reflection-based YCbCr method.

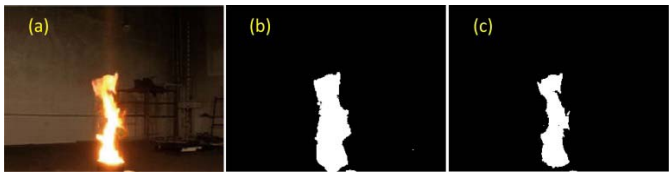

Fig. 6. Fire segmentation under the condition of nonreflection. (a) Original fire image. (b) Segmentation result by the reflection-based YCbCr method. (c) Segmentation result by the nonreflection-based YCbCr method.

the threshold of reliable segmentation is in the range from 35 to 55 under the condition of nonreflection and in the range from 30 to 45 under the condition of reflection in these two cases. Generally, the intermediate value is considered to be the value that achieves the best results.

To judge whether there is fire reflection or no reflection, we use the improved YCbCr method to conduct the fire segmentation under the condition of nonreflection and reflection. In the presence of fire reflection, we first use the improved nonreflection-based YCbCr method to segment the fire, and Fig. 5(b) shows the segmented fire area $S_{\mathrm{non\_ref}}$ Then, we use the improved reflection-based YCbCr method to segment the fire, and Fig. 5(c) shows the segmented fire area $S_{\mathrm{ref\_\_ref}}$. It can be seen that the relationship between $S_{\mathrm{ref\_non}}$ and $S_{\mathrm{ref\_ref}}$ satisfies

$$S_{\mathrm{non\_ref}} > S_{\mathrm{ref\_ref}}, \quad \text{under reflection condition.} \tag{6}$$

Similarly, in the absence of fire reflection, we first use the improved reflection-based YCbCr method to segment the fire, and the segmented fire area $S_{\mathrm{ref\_non}}$ is shown in Fig. 6(b). Also, the improved nonreflection-based YCbCr method is used to segment the fire, producing the segmented fire area $S_{\mathrm{non\_non}}$ as shown in Fig. 6(c). In this case, the two segmented fire areas

Algorithm 2 Fine Segmentation of Fire

Input: Initial segmentation results

1. Determine the seed points after initial segmentation of fire
2. Determine optimal threshold thr_{gray} according to objective function
$\mathrm{H}(thr) = -\sum_{i=1}^{K}\left(\tilde{D}_i ln\left(\tilde{D}_i\right) + \left(1-\tilde{C}_i\right) ln\left(1-\tilde{C}_i\right)\right)$
3. Conduct the region growing according to difference between seed point and adjacent points

$$\begin{cases} \left|\tilde{h}-h_i\right| \le thr_{gray}, h_i \in \tilde{h} \\ \left|\tilde{h}-h_i\right| > thr_{gray}, h_i \notin \tilde{h} \end{cases} \quad i = 1.2, \ldots, l$$

Output: Fire segmentation image

satisfy

$$S_{\text{ref\_non}} > S_{\text{non\_non}}, \text{under nonreflection condition.} \tag{7}$$

B. Fine Segmentation of Fire Based on the Improved Region Growing

The luminance of fire is significantly different from the surroundings during combustion. In view of this, the initial segmentation region obtained by the improved YCbCr model is divided into several connected regions. The centroid (X_j, Y_j) of the jth segmented image is calculated by

$$\begin{cases} X_j = \dfrac{1}{N_j}\displaystyle\sum_{i=1}^{N_j} x_{j,i} \\ Y_j = \dfrac{1}{N_j}\displaystyle\sum_{i=1}^{N_j} y_{j,i} \end{cases} \quad \left(x_{j,i}, y_{j,i}\right) \in C_j \tag{8}$$

where C_j is the jth segmented image and N_j is the number of pixels of the jth segmented image.

The seed point is defined as the weighted average of centroids of the connected regions, which is given by

$$\begin{cases} X = \dfrac{\sum_{j=1}^{n} w_j X_j}{\sum_{j=1}^{n} w_j} \\ Y = \dfrac{\sum_{j=1}^{n} w_j Y_j}{\sum_{j=1}^{n} w_j} \end{cases} \tag{9}$$

where w_j is the weight of the jth centroid and n is the number of centroids.

Calculate the number of pixels of segmented image, and the maximum is regarded as the suspected fire region. The scope of the weight of suspected fire region is usually set as $[0.7, 0.9]$, and the interference region of flame is set as $[0, 0.3]$. The seed points are determined as follows. First, the fire region in the foreground as shown in Fig. 7(a) is segmented according to the procedure described in Section II-A, as shown in Fig. 7(b). Second, the segmented image is determined. Equation (8) is used to calculate the centroid of segmented image, marked with colored stars as shown in Fig. 7(c). The weight for the interference region of fire is set as 0.1, and the weight for the suspected flame region is set as 0.9. The seed points are then obtained according to (9), marked with a gray star as shown in Fig. 7(d).

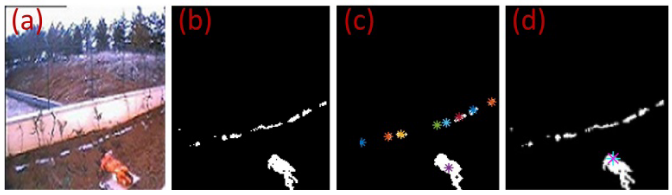

Fig. 7. Determination of seed point. (a) Original fire image. (b) Segmentation result by the improved YCbCr. (c) Centroid of segmented image. (d) Seed point.

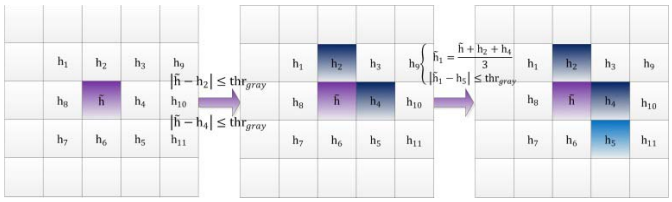

Fig. 8. Improved region growing algorithm. (a) Pixels of seed and adjacent points. (b) Merge of initial seed and adjacent pixel points. (c) Direction of the region growth.

The differences between the gray value of seed point and the adjacent points are calculated and compared with threshold value $\mathrm{thr}_{\mathrm{gray}}$ as described in (10). If the difference is smaller than $\mathrm{thr}_{\mathrm{gray}}$, the seed and adjacent point are merged together

$$\begin{cases} \left|\tilde{h}-h_i\right| \le \mathrm{thr}_{\mathrm{gray}}, & h_i \in \tilde{h} \\ \left|\tilde{h}-h_i\right| > \mathrm{thr}_{\mathrm{gray}}, & h_i \notin \tilde{h} \end{cases} \quad i = 1.2, \ldots, l \tag{10}$$

where h_i is the gray value of the ith adjacent point and $\tilde{h}$ is the gray value of the seed point. Note that a point here is actually a pixel of an image.

The steps of the improved region growing algorithm are shown in Fig. 8. Denote $\tilde{h}$ as the initial seed point and $h_1, h_2, \ldots, h_8$ as the eight adjacent connected pixels. If h_2 and h_4 meet the first condition of (10), then they are regarded as another seed points and merged with the initial seed point, as shown in Fig. 8(b).

After growth of initial seed point, we calculate the mean of the gray values that belong to the seed points, as described in the following equation:

$$\tilde{h}_1 = \frac{\tilde{h} + h_2 + h_4}{3}. \tag{11}$$

According to (11), the gray value of new seed point is calculated. Similarly, we compare $h_1, h_3, h_5, h_6, h_7, h_8, h_9, h_{10}$, and h_{11} with $\tilde{h}_1$ to determine the new growth region of seed point, as shown in Fig. 8(c). Ultimately, the improved region growth method is implemented according to the growth of seed points, and then, the fire is accurately segmented.

The threshold value $\mathrm{thr}_{\mathrm{gray}}$ will determine the segmentation effect of region growth. In order to obtain the optimal threshold $\mathrm{thr}_{\mathrm{gray}}$ in the region growth algorithm, an objective function is constructed by combining the evaluation index of image segmentation [33]. The optimal threshold can be determined by seeking the optimal value of the objective function. The homogeneity inside the region is regarded as the common index of quality of image segmentation. Assume

that R_i and R_j are the two adjacent connected regions after segmentation according to the threshold thr_{gray}. The variance of the connected regions of the segmented image and the mean of the absolute difference between the average grayscale of adjacent connected regions are obtained

$$\begin{cases} C_i = \dfrac{1}{A_i} \displaystyle\sum_{x \epsilon R_i} \left[f(x) - \mu_{R_i} \right]^2 \\ D_i = \dfrac{1}{k_i} \displaystyle\sum_{j=1}^{k_i} \left| \mu_{R_i} - \mu_{R_j} \right| \end{cases} \tag{12}$$

where A_i is the area of region R_i, μ_{R_i} and μ_{R_j} are the mean of the grayscale of region R_i and R_j, respectively, $f(x)$ is the grayscale value of image, and k_i is the number of regions adjacent to region R_i.

The maximum variance of the connected regions of the segmented image and the maximum difference between the average grayscale of adjacent connected regions are obtained as

$$\begin{cases} C_{\max} = \max \left(\dfrac{1}{A_i} \displaystyle\sum_{x \epsilon R_i} \left[f(x) - \mu_{R_i} \right]^2 \right) \\ D_{\max} = \max \left| \mu_{R_i} - \mu_{R_j} \right|. \end{cases} \tag{13}$$

According to (13), the normalization of C_T and D_T is given by

$$\begin{cases} \tilde{C}_i = \dfrac{1}{C_{\max} A_i} \displaystyle\sum_{x \epsilon R_i} \left[f(x) - \mu_{R_i} \right]^2 \\ \tilde{D}_i = \dfrac{1}{D_{\max} k_i} \displaystyle\sum_{j=1}^{k_i} \left| \mu_{R_i} - \mu_{R_j} \right| \end{cases} \tag{14}$$

where $\tilde{C}_i$ reflects the uniformity within region R_i and $\tilde{D}_i$ reflects the average contrast between region R_i and its adjacent regions.

The good image segmentation means the better uniformity within connected regions and greater contrast between adjacent regions. However, the uniformity and contrast are always contradictory. In view of this, this article uses entropy to balance uniformity and contrast and then obtain more image feature information after segmentation. According to the definition of entropy, a binary entropy model of uniformity and contrast is constructed as

$$H = -\sum_{i=1}^{K} \left(\tilde{D}_i \ln \left(\tilde{D}_i \right) + \left(1 - \tilde{C}_i \right) \ln \left(1 - \tilde{C}_i \right) \right) \tag{15}$$

where K is the number of connected regions in the segmented image.

According to (15), different threshold thr_{gray} corresponds to different entropies. Generally, the range of threshold thr_{gray} is from 10 to 50 based on statistics. Therefore, we use the search method to determine the maximum entropy within the threshold range. Ultimately, the optimal threshold thr_{gray} can be obtained.

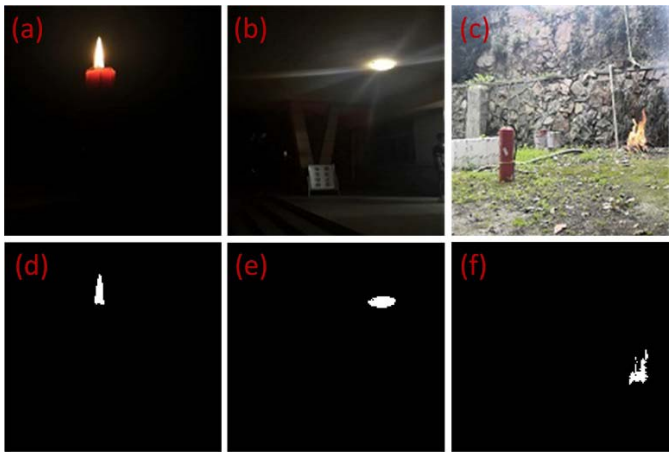

Fig. 9. Segmentation of fire and interference source. (a)–(c) Original image of candle flame, light, and fire, respectively. (d)–(f) Segmentation results of candle flame, light, and fire, respectively.

C. Identification of Fire Based on the Multifeature Description

According to the above analysis, the fire can be segmented by the improved region growth algorithm. Nevertheless, the segmentation method cannot be directly used for the fire identification because of the influence of interference source. For example, Fig. 9(a)–(c) shows the original images: candle flame, bulb light, and a fire outdoor. The improved region growth algorithm is used to segment the three types of images, as shown in Fig. 9(d)–(f). It can be seen that the three objects are totally segmented and they are all regarded as fire. Therefore, the segmentation method gives the false identification of the two objects whose color and illuminance are close to the fire. In order to improve the accuracy of fire identification, we need to analyze the variation of flame character to further judge the fire. Here, three quantitative indicators are proposed to identify fire, which are the variation coefficient of area, the dispersion of centroid, and the circularity.

1) Variation Coefficient of Area: The area of fire is always changeable as the fire getting smaller or larger during the process of combustion. Therefore, we can use this character to further judge the fire. The segmented image is binarized so that each pixel of the image has a gray value of either one or zero. Assuming that the area of each pixel is s, the area of segmented image is calculated by

$$S = s * p \tag{16}$$

where p is the number of pixels with a value being one. Also, the rate of variation of area is calculated by

$$S' = \frac{\Delta S}{\Delta t} = \frac{\left| S_{i+j} - S_j \right|}{t_{i+j} - t_j} \tag{17}$$

where S_j and S_{i+j} are the areas of jth and $(i + j)$th frame image, respectively, and $t_{i+j} - t_j$ is the time interval between t_j and t_{i+j}

For different image sequences, the area and the rate of area variation of segmented images are calculated, as shown in Fig. 10. Fig. 10(a) shows the areas of different image

Algorithm 3 Identification of Fire

Input: Segmented fire image

Variation coefficient of area

1. Calculate fire area:

$S = s * p$

2. Obtain the variation coefficient of area

$C_S = \frac{\sigma_S}{\mu_S}$

If $C_S > 27.63$ the object is possibly identified as fire

Dispersion of centroid

1. Calculate fire centroid:

$$\begin{cases} X_j = \dfrac{1}{n}\sum_{i=1}^{n} x_i \\ Y_j = \dfrac{1}{n}\sum_{i=1}^{n} y_i \end{cases}$$

2. Obtain the dispersion of centroid

$d_C = \sqrt{|D_{XY}|} = \sigma_X \sigma_Y$

If $d_C > 150$ the object is possibly identified as fire

Circularity

$R = \frac{4\pi S}{l^2}$

If the circularity of a continuous multi-frame image is less than 0.3, the object is possibly identified as fire

Output: Fire identification results

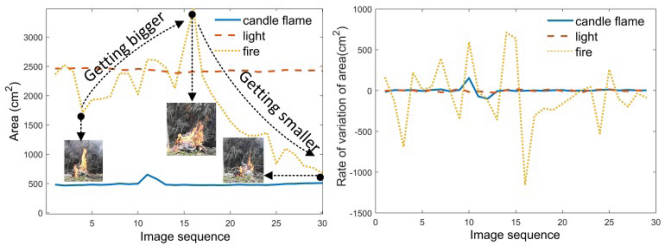

Fig. 10. Variation of the area of segmented image. (a) Area of different segmented objects. (b) Rate of area variation of different segmented objects.

sequences. The areas of candle flame and light remain approximately the same with different image sequences, but the area of fire is always changeable. The combustion goes through three phases of initial, developed, and decay of combustion, although Fig. 10(a) mainly shows the developed and decay phases, with the flame area getting bigger and then smaller. During the stage of development, the burning area is enlarged and the scope of fire is getting bigger, as shown in the first and second fire image of Fig. 10(a). During the stage of decay, the burning area reduces and the flame area is getting smaller, as shown in the second and third fire images of Fig. 10(a). Fig. 10(b) shows that the rates of area variation of bulb light and candle flame area are almost zero, although the candle flame area has one great fluctuation from the 10th to 11th image sequence. Therefore, we can perform the fire identification according to the variation trend of segmented image area. The variation coefficient of area is defined to describe the variation trend of area, as shown in the following equation:

$$C_S = \frac{\sigma_S}{\mu_S} \tag{18}$$

where σ_S and μ_S are the standard deviation and mean of area of segmented image sequence, respectively. Table I shows the

TABLE I

VARIATION COEFFICIENT OF AREA OF DIFFERENT OBJECTS

Objects	Candle flame	Light	Fire
Variation coefficient	5.760	1.849	27.069

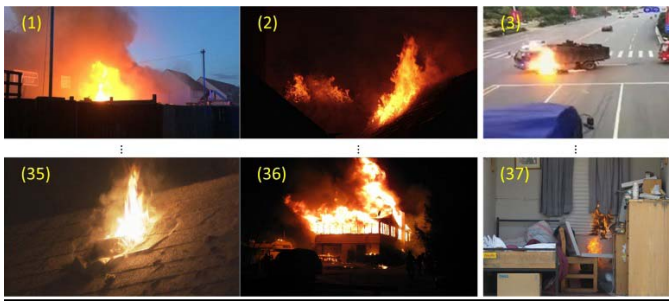

Fig. 11. Thirty-seven different types of fire videos.

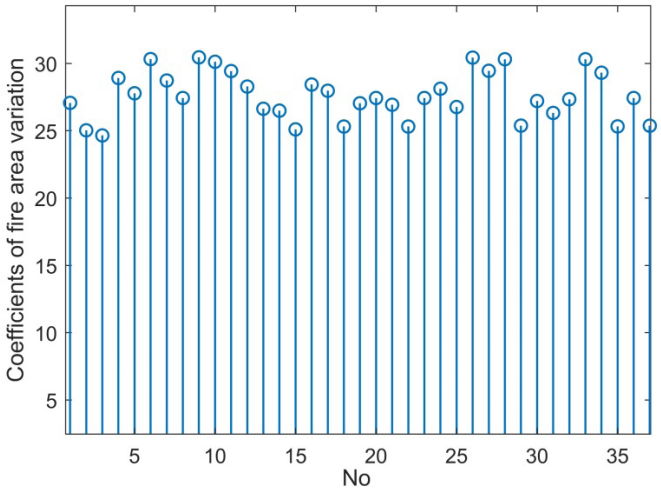

Fig. 12. Coefficients of area variation of different fires.

variation coefficient of area variation of three sequences of 30 images for candle flame, light, and fire.

It can be seen from Table I that the coefficient of area variation of fire is much greater than the candle flame and light and that of light is the lowest among the three objects. The reason for this result is that the luminance variation of fire is greater than the light and candle flame, and the luminance variation of light is rather marginal over the image sequence. Therefore, we can identify fire according to the magnitude of coefficient of area variation. In order to determine the magnitude of coefficient of area variation of fire, we collect many different types of fire videos, which include big and small fire, as shown in Fig. 11.

Use the proposed segmentation method to segment different fire images read from 37 fire videos. According to (18), we calculate the variation coefficients of 37 types of fire areas, as shown in Fig. 12.

From Fig. 12, it can be seen that coefficients of area variation of different fires are around 27, and they are not affected by the size of fire. The main reason for this result

is that the size of fire will directly influence the standard deviation and the mean of fire image sequence simultaneously. The mean coefficient is calculated to be 27.63. Therefore, the magnitude of coefficient of area variation of fire is defined as 27.63. If the coefficient of area variation of segmented images is much greater than 27.63, the object is possibly identified as fire.

2) Dispersion of Centroid: For each image, we define the left bottom of the image as the origin of pixel coordinate system. Consequently, each pixel coordinate of segmented image is obtained. The jth centroid coordinates of the segmented image are then calculated by

$$\begin{cases} X_j = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i \\ Y_j = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} y_i \end{cases} \tag{19}$$

where n is the number of pixels in the segmented image and (x_i, y_i) are the ith pixel coordinates of segmented image.

Fig. 13 shows the distribution of centroids of three different segmented image sequences associated with the three objects, and it can be seen that the centroids of the segmented fire images have a wider spread than the light and candle flame. Apparently, the centroids of segmented light image sequence are almost on the same position, so they have the highest aggregation. The reason for this phenomenon is that the luminance of segmented light image sequence is almost invariable. Also, the centroids of segmented fire images change irregularly due to the variation in combustion. Accordingly, the dispersion of fire centroid is greater than the candle flame and bulb light. Here, we use the standard deviation to describe the dispersion of the centroid. The centroid coordinates can be expressed as a random vector of 2-D

$$C = \begin{bmatrix} X \\ Y \end{bmatrix}. \tag{20}$$

The standard deviations of X and Y are obtained by

$$\begin{cases} \sigma_X = \sqrt{\dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}\left(X_j - \bar{X}\right)^2} \\ \sigma_Y = \sqrt{\dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}\left(Y_j - \bar{Y}\right)^2} \end{cases} \tag{21}$$

where (X_j, Y_j) are the centroid coordinates calculated by (19) and $(\bar{X}, \bar{Y})$ are the mean centroid coordinates. If the random vector components X and Y are uncorrelated, then the variance matrix of the random vector of 2-D is expressed as

$$D_{XY} = \begin{bmatrix} \sigma_X^2 & 0 \\ 0 & \sigma_Y^2 \end{bmatrix}. \tag{22}$$

The determinant of D_{XY} is then computed by

$$|D_{XY}| = \sigma_X^2 \sigma_Y^2. \tag{23}$$

In this article, we will use the square root of determinant of D_{XY} to determine the dispersion of centroid of different image sequences

$$d_C = \sqrt{|D_{XY}|} = \sigma_X \sigma_Y. \tag{24}$$

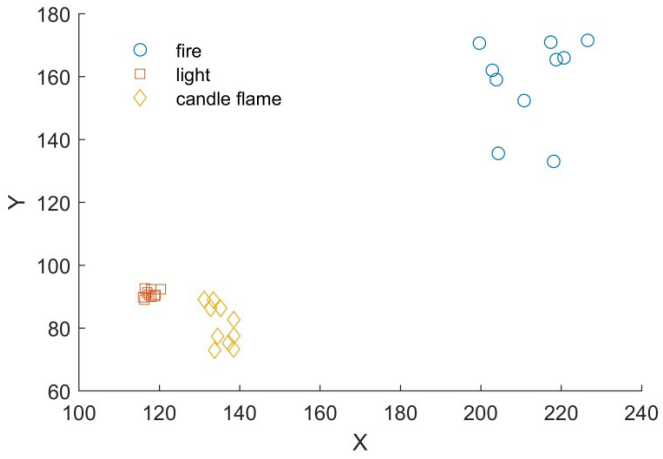

Fig. 13. Centroids of segmented images with different time series.

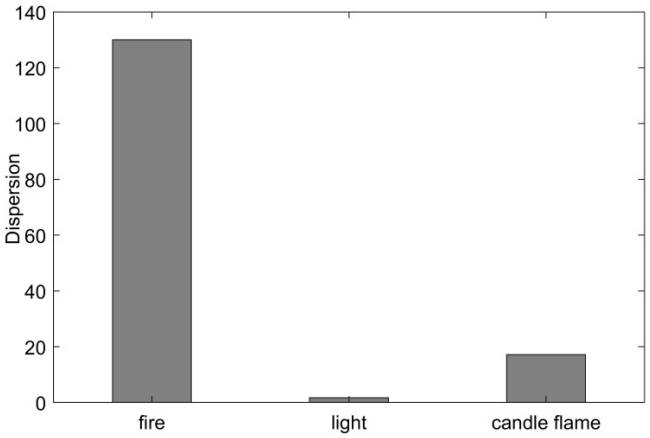

Fig. 14. Dispersion of centroid of candle flame, light, and fire images.

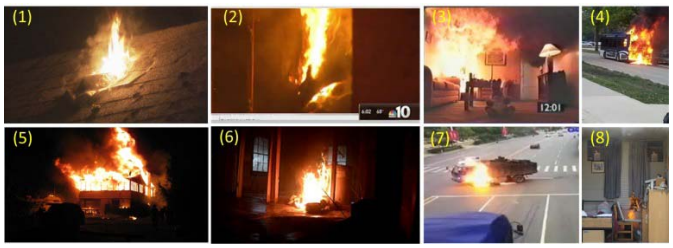

Fig. 15. Video stream images of fires.

Fig. 14 shows the calculated dispersion of centroid of candle flame, light, and fire images.

In Fig. 14, the dispersion of centroid of fire is almost 130, which is about eight times of candle flame and more than 30 times of light. Thus, the fire can be readily identified by distinguishing the magnitude of dispersion of centroid of the segmented images. Nevertheless, we need to determine the threshold of centroid of the fire images. We select eight different types of fire videos, which include small, large, indoor, outdoor, far, and near fire, as shown in Fig. 15.

Each fire video consists of 30 images, and we read an image every 9 s, so the video lasts 3 min. According to (24),

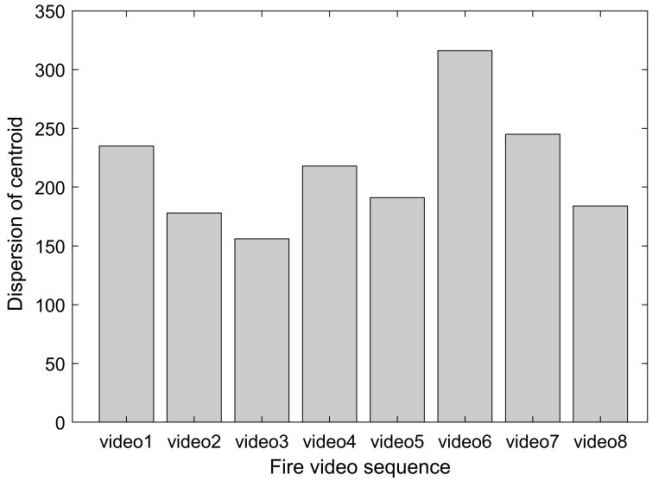

Fig. 16. Dispersion of centroid of different fire videos.

the dispersion of centroid of images of every fire video is calculated, as shown in Fig. 16.

Fig. 16 shows that the maximum dispersion of centroid of the eight different fire videos is greater than 300, and the minimum is almost greater than 150. Therefore, we set the threshold of the dispersion of centroid to be 150. If the dispersion of centroid of segmented images is greater than a threshold, the object can be identified as fire.

3) Circularity: Circularity describes the degree of irregularity of fire. Therefore, circularity can be used as one of the quantitative indicators for fire identification, which is defined as

$$R = \frac{4\pi S}{l^2} \tag{25}$$

where S and l are the area and perimeter of segmented object, respectively. According to the definition, the more circularly segmented object, the greater the circularity.

Fig. 17 shows the circularity of candle flame, light, and fire in three 30-image sequences, calculated by (25). It can be seen that the circularity of fire is around 0.2 and basically varies between 0.15 and 0.33. The circularity of candle flame is greater than fire, and its scope is between 0.34 and 0.45. The light has the greatest circularity, which ranges between 0.74 and 0.91 and is much greater than fire. The reason for this phenomenon is that the segmented light is close to the circle, as shown in Fig. 9(e).

In order to fully evaluate the effect of circularity on fire identification, the circularity of 12 different types of fire videos is analyzed, and part of the video stream frame image is shown in Fig. 18.

The 12 types of fire videos include not only large and small fires but also indoor and outdoor fires. Also, these videos contain long- and short-distance observations. Therefore, these fire videos almost cover all types of fires. For these, fire images were extracted from each of the 12 video streams of different types of fire. The circularities of fire image of different fire videos are calculated, as shown in Fig. 19.

For each type of fire video, Fig. 19 shows that the trend of circularity variation of different fire images is stable. The

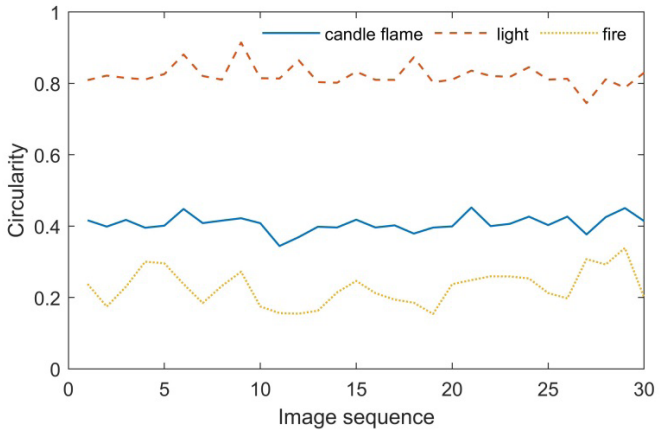

Fig. 17. Circularity of candle flame, light, and fire.

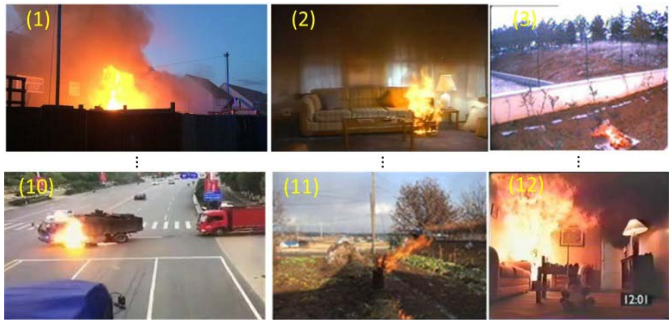

Fig. 18. Video stream images of different types of fires.

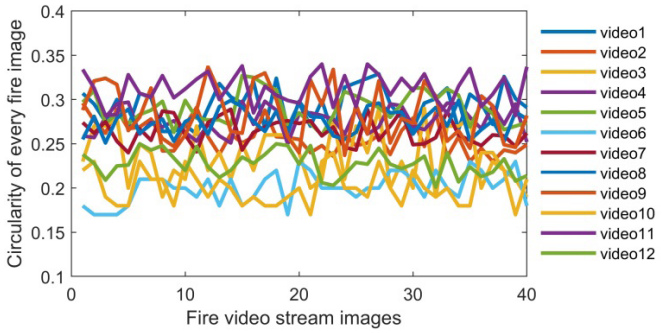

Fig. 19. Circularities of fire image of 12 fire videos.

difference between different fire videos in circularity is small. The scope of circularities of different fire videos is from 0.17 to 0.34, and the maximum circularity of fire image is no more than 0.34.

According to Fig. 19, we obtain the maximum, mean, and standard deviation of circularity of every fire video stream image under different scenarios, as shown in Fig. 20.

Fig. 20 shows that the maximum of circularity of 12 types of fire is mainly around 0.3, and the maximum is less than 0.35. The mean of circularity of every video stream image under different scenarios ranges between 0.2 and 0.3. The maximum and the mean circularity of videos 6, 10, and 12 are much smaller than the other videos. The reason for this result is that the fires of videos 6, 10, and 12 present long

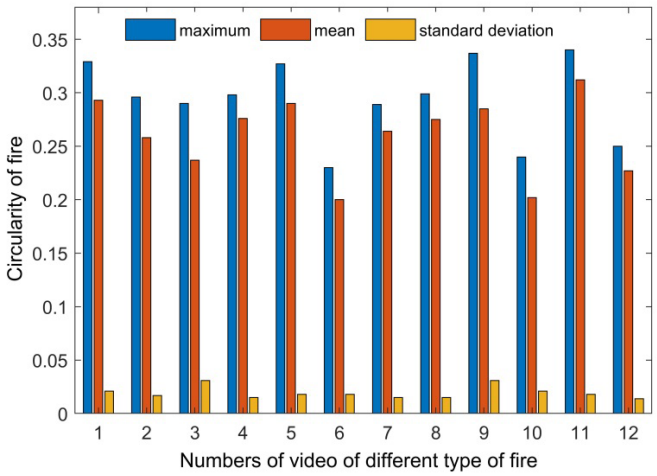

Fig. 20. Maximum, mean, and standard deviation of circularity.

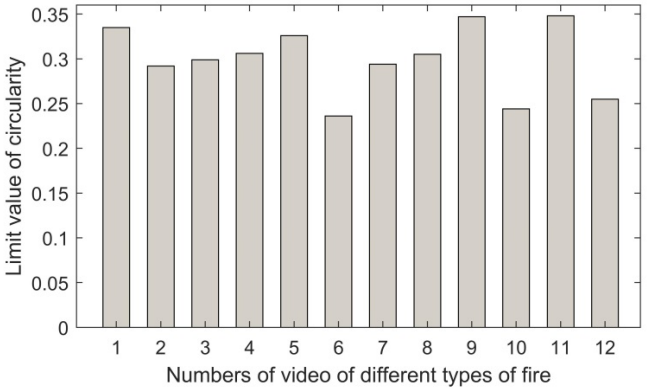

Fig. 21. Limit value of circularity of 12 types of fire.

shape. The standard deviation of circularity is about between 0.02 and 0.03. According to limit theorems [34], the limit value of circularity of every video stream image under different scenarios is calculated by

$$\rho = \mu + 2\sigma \tag{26}$$

where μ is the mean of circularity and σ is the standard deviation of circularity.

According to (26), the limit value of circularity of every video stream image is obtained, as shown in Fig. 21.

Fig. 21 shows that the maximum of limit value of circularity in 12 different types of fires is 0.34 and the minimum is 0.24. The limit value of circularity of most of the fires is mainly around 0.3. Therefore, the object is possibly identified as fire if the circularity of a continuous multiframe image is less than 0.3.

III. Performance Evaluation Results

A. Performance Analysis of Fire Segmentation

1) Comparison of Different Methods: In order to verify the advantage of the proposed fire segmentation method, it is compared with the threshold segmentation algorithm [35] and color segmentation algorithm [6], as shown in Fig. 22.

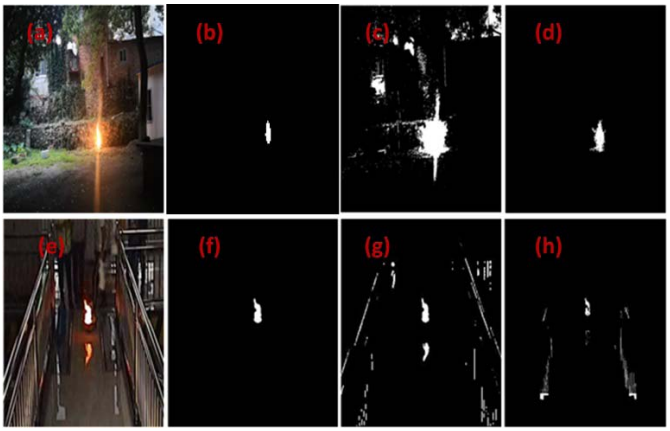

Fig. 22. Segmentation results of different methods. (a) and (e) Original image. (b) and (f) Proposed fire segmentation method. (c) and (g) Threshold segmentation algorithm. (d) and (h) Color segmentation algorithm.

TABLE II
Segmentation Accuracy of Different Methods

Original image	Threshold segmentation	Color segmentation	Proposed fire segmentation method
a	85.46%	92.56%	96.13%
e	49.04%	75.57%	97%

From Fig. 22(c) and (g), it can be seen that the segmented fire region by the threshold segmentation algorithm is obviously larger than the actual flame, and it includes a lot of noises. The color segmentation algorithm is superior to the threshold segmentation algorithm, but it still contains some noises, especially as shown in Fig. 22(h). Consequently, the segmented fire region is greater than the actual fire area. From Fig. 22(b) and (f), it can be seen that the segmented fire region by the proposed fire segmentation method is very similar to the actual fire region.

In order to evaluate the segmentation efficiency, the evaluation index of segmentation accuracy is constructed as

$$C = \frac{A \cap B}{\max(A, B)} \tag{27}$$

where A is the actual fire area and B is the segmented fire area. Table II shows the segmentation accuracy of the three methods.

From Table II, it can be seen that the segmentation accuracy of the proposed fire segmentation method is obviously higher than the other two methods. This means that the area of segmented fire by the proposed fire segmentation method is very similar to the area of actual fire. For the original image shown in Fig. 22(a), the segmentation accuracy of color segmentation algorithm is greater than 90%, but its segmentation fire area is greater than the actual fire. Conversely, the segmentation fire area of the proposed fire segmentation method is slightly smaller than the actual fire. The reason for this result is that the proposed fire segmentation method considers the influence of luminance, but the color segmentation algorithm only considers the color. The segmentation accuracy of threshold

Fig. 23. PASCAL VOC data set.

segmentation method is lower than that of the other two methods, achieving the worst segmentation efficiency.

Therefore, it is clear that the proposed fire segmentation method segments the fires accurately and the segmented fires are almost the same as the actual fires. It can be applied to flame segmentation for scenarios either with nonreflection region or with reflection region. The proposed fire segmentation method not only retains almost all the fire information but also eliminates nearly all the noises.

2) Performance Evaluation Based on the PASCAL VOC Data Set: The proposed segmentation method is mainly used to segment the fire. In order to analyze the influence of nonfire objects on the fire segmentation, the metrics of PASCAL VOC data set is used to evaluate the overall performance of the proposed fire segmentation method. In the PASCAL VOC data set, we select indoor and outdoor images that include light, fire-like, and nonfire object, as shown in Fig. 23.

Use the proposed fire segmentation method to segment the outdoor image and compare the results with actual segmentation results, as shown in Fig. 24.

From Fig. 24(e) and (f), it is clearly visible that there is no object in the image after segmentation with the proposed fire segmentation method. Nevertheless, the actual segmentation results include objects, such as horse and train, as shown in Fig. 24(a) and (b). The possible reason for the segmentation results is that the proposed fire segmentation method is only applicable to the segmentation of fire scene. No object in the image after segmentation illustrates that there is no fire. Fig. 24(g) shows that the car body is segmented by the proposed fire segmentation method, and the two ship objects are not segmented. From Fig. 24(h), three sheep cannot be segmented by the proposed method, but the sun is accurately segmented. The reason for these segmentation phenomena is that the color of the car body and the luminance of sun are similar to fire.

In view of indoor scene, some objects are segmented by the proposed fire segmentation method, as shown in Fig. 25(e)–(g). The proposed method mainly segments fire, as shown in Fig. 25(e), and segments lamplight, as shown in Fig. 25(f) and (g), but it does not segment other indoor objects. Therefore, if the segmentation result is empty, there is no fire or fire-like object in the environment of interest.

B. Performance Analysis of Fire Identification

1) Evaluation Metrics: A number of indexes or metrics can be used to evaluate the fire identification performance, including correct detection rate [6], false positive (FP) rate [23],

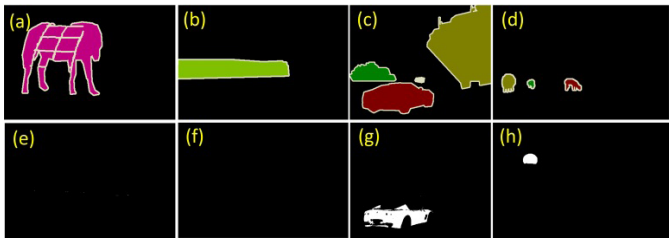

Fig. 24. Segmentation results of outdoor scene of the PASCAL VOC data set. (a)–(d) Actual segmentation results and (e)–(h) segmentation results of the proposed fire segmentation method.

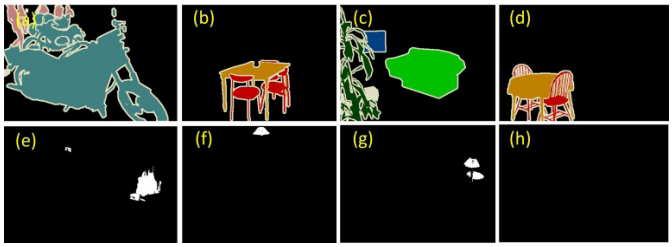

Fig. 25. Segmentation results of indoor scene of the PASCAL VOC data set. (a)–(d) Actual segmentation results and (e)–(h) segmentation results of the proposed fire segmentation method.

and accuracy [36]. To achieve accurate evaluation, multiple evaluation metrics may be used.

Originally introduced in the field of information retrieval, precision, recall, and $F1$ score are the three common measures for evaluating the effectiveness of a system [37]. Precision represents the percentage of correctly retrieved elements produced by the considered system, whereas recall indicates the percentage of reference truth data that are correctly retrieved. The former is sensitive to the existence of spurious elements, whereas the latter is sensitive to the existence of reference data that are not recognized by the system. Finally, the $F1$ score balances precision and recall and is commonly used as a unique measurement of a system's overall effectiveness. The use of the tripartite measures of precision, recall, and $F1$ score is an effective way for assessing the performance in object detection. Therefore, the tripartite measures are used to evaluate the effect of fire identification in this article.

Given an identification method and a correct hypothesis, there are four possible identification outcomes. If the correct hypothesis of video steam is fire and it is identified as fire, then it is counted as a true positive (TP). If it is identified as nonfire, it is then counted as false negative (FN). If the correct hypothesis of video stream is nonfire and it is identified as nonfire, it is then counted as true negative (TN). If it is identified as fire, then it is counted as FP. All the four possible outcomes of fire identification are listed in Table III.

The identification performance is quantitatively measured using precision, recall, and $F1$-score, which are defined as

$$\text{Precision} = \frac{|\text{TP}|}{|\text{TP}| + |\text{FP}|} \tag{28}$$

$$\text{Recall} = \frac{|\text{TP}|}{|\text{TP}| + |\text{FN}|} \tag{29}$$

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}. \tag{30}$$

TABLE III
FOUR POSSIBLE OUTCOMES OF FIRE IDENTIFICATION

	Positive (presence of fire)	Negative (absence of fire)
True	True Positive(TP)	True Negative (TN)
False	False Positive(FP)	False Negative(FN)

Fig. 26. Examples of images extracted from the videos used for testing the methods. (a), (b), (d), (f), and (g) Images are taken from fire videos of fire. (c) and (h) Images are taken from nonfire videos. (e) Image is taken from candle flame video.

Note that, similar to precision, a higher recall or a higher $F1$-score means a better identification performance.

2) Performance Comparison: The proposed method is evaluated against other state-of-the-art approaches, based on the mentioned three performance indexes. The statistical color model in video sequences [6] is probably the most popular method for fire identification, which constructs a real-time fire detector by combining color information with the registered background scene. GMM [38], ViBe [39], and RPCA algorithm are mainly used to detect the dynamic object in video, such as fire. One disadvantage of these methods is that it does not distinguish fire from other dynamic objects, such as moving vehicles and leaves blown by the wind. Another popular fire detection method is the deep learning-based method that uses a video sequence. For instance, the method in [28] uses Faster-RCNN to detect the suspected regions of fire and of nonfire based on their spatial features. The HoG-SVM model and the deep learning model, YOLO-v4, were also used to conduct the object detection [40]. Thus, these seven state-of-the-art fire identification methods are selected for performance comparison against the proposed method.

In order to show the advantages of the proposed method against other state-of-the-art approaches, we collected and recorded eight video sequences. Each video sequence has 10 000 frames of size 176 × 144 with a frame rate of 10 frames/s. These video sequences include images containing nonfire and fire videos, as shown in Fig. 26. Fig. 26(a), (b), and (f) shows the images with outdoor fire, and Fig. 26(d) and (g) shows the images with indoor fire. Fig. 26(c) shows the dynamic object bulb light. Fig. 26(e) and (h) shows the images with bulb light and candle flame, which have color and luminance similar to fire. Therefore, they will disturb the identification of fire. Each video is divided into 30 video clips, each of which contains 20 images. Therefore, there are 600 images in each video.

The statistical color model, GMM, ViBe, RPCA, HoG-SVM, Faster-RCNN, and YOLO-v4 are used to identify fire from each image, but the proposed method only needs to identify fire for each video clip. Thus, there are 240 video clips for the proposed method to identify and 4800 images for the other methods to identify. For the GMM, ViBe, and RPCA algorithm, the detected dynamic objects are regarded as fire. For the HoG-SVM, Faster-RCNN, and YOLO-v4 deep learning model, a large number of images are required to train these models. However, current small-scale image/video fire databases cannot meet the needs, some of which are listed in Table IV. Therefore, 15 000 images are collected from small public fire image/video databases, large public images/video data sets, previous experiment data from research institutions, and the Internet [41]. These images are selected as the training set of the Faster-RCNN, HoG-SVM, and YOLO-v4 model.

In order to ensure that the proposed method can be used in different scenes, the optimal threshold thr_{gray} of each video needs to be determined in advance according to the objective function. The determination of the optimal threshold mainly ensures that the fire and fire-like objects can be segmented accurately. According to (15), the corresponding optimal thresholds of eight individual videos named (a)–(h) are obtained, as shown in Table V.

According to the optimal threshold in Table V, the fire in the videos can be segmented, as described in Section II-B. Then, the fire identification is implemented according to the method described in Section II-C.

Among the 4800 images, 3000 images are associated with fire, while fire is absent in the remaining 1800 images. As mentioned earlier, there are four possible identification outcomes (TP, TN, FP, and FN), and only one is accepted for one image by each of the six methods. Table V shows the identification results of the seven algorithms using the 4800 images.

The Faster-RCNN and the proposed method have very similar counts of TP and small numbers of FN, while the TP and TN of HoG-SVM are less than those of Faster-RCNN and YOLO-v4. In addition, the FP and FN of HoG-SVM are more than those of the other two deep learning methods. This indicates that Faster-RCNN and YOLO-v4 methods perform fire detection better than the other deep learning methods. YOLO-v4 has the largest TP and the smallest FN, indicating that the method can accurately identify the fire. However, the method can easily misidentify the candle flame in Fig. 26(e) as fire, resulting in a relatively large FN value. Faster-RCNN and HoG-SVM also showed the similar results.

These three deep learning methods also have low accuracy in identifying nonfire object because of the influence of candle flame, resulting in low TN value. Table VI shows that the TP values of GMM, ViBe, and RPCA methods are lower than those of other methods. The main reason for the identification results is that the fire dynamic change is not obvious, as shown in Fig. 26(a) and (d). The FP values of the three methods are greater than other methods. This is because these methods can accurately detect the dynamic object but cannot distinguish the difference between the dynamic object and fire. Since the statistical color model only considers color information, the TP value obtained by the method is high, whereas the TN value is low. The proposed method considers not only the

TABLE IV
SMALL-SCALE FIRE IMAGE/VIDEO DATABASES

Institutions	Format	Object
Bilkent University	Video	Fire, disturbance
CVPR Lab. At Keimyung University	Video	Fire, disturbance
UMR CNRS 6134 SPE， Corsica University	Dataset	Fire
National Fire Research Laboratory, NIST	Video	Fire

TABLE V
OPTIMAL THRESHOLD FOR FIRE SEGMENTATION OF DIFFERENT VIDEOS

Different videos	a	b	c	d	e	f	g	h
thr_{gray}	44	32	25	26	22	31	42	22

TABLE VI
NUMBER OF FOUR POSSIBLE OUTCOMES

Different methods	TP	TN	FP	FN
Statistical color model	2880	1220	580	120
GMM	2205	1204	596	795
ViBe	2654	1184	616	346
RPCA	2104	1188	612	896
HoG-SVM	2832	1196	604	168
Faster-RCNN	2923	1227	573	77
YOLO-v4	2981	1298	502	19
Proposed method	2940	1720	80	60

TABLE VII
TRIPARTITE MEASURES OF DIFFERENT METHODS

Different methods	Precision (%)	Recall (%)	F1 score (%)
Statistical color model	83.2	96.0	89.2
GMM	78.7	73.5	76.0
ViBe	81.2	88.5	84.7
RPCA	77.5	70.1	73.6
HoG-SVM	82.4	94.4	88.0
Faster-RCNN	83.6	97.4	90.0
YOLO-v4	85.6	99.4	92.0
Proposed method	97.4	98.0	97.7

segmentation of fire but also the variation characteristics of fire. As a consequence, both the TP and TN values of the proposed method are high.

According to the TP, TN, FP, and FN, the tripartite measures are calculated, as shown in Table VII. It can be seen that the precision of the YOLO-v4 is similar to that of HoG-SVM, Faster-RCNN, and statistical color model, but higher than that of dynamic detection models (GMM, ViBe, and RPCA) that are easily affected by moving objects. However, the precisions of the YOLO-v4, Faster-RCNN, HoG-SVM, and statistical color model are lower than that of the proposed method. The statistical color model and the deep learning algorithms are easily affected by light, candle flame, and other factors, whereas the proposed method can effectively distinguish fire from light and candle flame. In addition, the recall of the

TABLE VIII
TRIPARTITE MEASURES OF THE DEEP LEARNING AND THE PROPOSED METHOD

Different methods	Precision (%)	Recall (%)	F1 score (%)
Deep CNN-based fire detection model	95.8	99.2	97.5
Proposed method	97.4	98.0	97.7

TABLE IX
TRAIN TIME OF THREE DIFFERENT DEEP LEARNING METHODS

	Deep CNN-based model	YOLO-v4	Faster-RCNN
Training time(h)	48	43	56

TABLE X
IDENTIFICATION TIME OF THREE DIFFERENT METHODS

	Proposed method	Deep learning method		
		Deep CNN-based model	YOLO-v4	Faster-RCNN
Identification time(s)	5214	6834	6112	6456

proposed method is almost the same with Faster-RCNN but slightly lower than the YOLO-v4. This is because the recall is only related to TP and FN as indicated by (29), for both of which the YOLO-v4 performs a bit better than the HoG-SVM, Faster-RCNN, and proposed method. The precision and $F1$-score of the proposed method are both considerably higher than that of the other methods. Therefore, the experimental results show that the proposed method can successfully improve the fire identification accuracy compared with the statistical color model, dynamic detection models, or deep learning algorithms by reducing the false identifications.

However, the identification results of deep learning model in Table VII are affected by model training because deep learning model training is parameter tuning dependent. In order to avoid the influence of parameter tuning, the constructed deep learning (deep CNN-based fire detection) model [26] is used to conduct the fire identification, and the identification results are compared with our proposed method, as shown in Table VIII.

Table VIII shows that the proposed method is almost the same as the deep CNN-based fire detection model in terms of $F1$ score. Although the precision of the proposed method is slightly higher than that of deep learning method, the recall of the deep learning method is slightly higher, which may be more important than precision in fire identification. Thus, from the perspective of recall, the proposed method is a bit inferior to deep learning method. However, the deep learning model requires training and the hours of the training time for the three deep learning methods are shown in Table IX.

From Table IX, it can be seen that the three different deep learning methods cost a lot of training time. However, the proposed method does not need to train the fire identification model, so the proposed method saves a lot of training time. In view of five fire video sequences, the proposed method needs to detect 150 fire video clips and the three

deep learning methods (deep CNN-based model, YOLO-v4, and Faster-RCNN) need to detect the 3000 fire images, and the identification time periods of the four methods are obtained, as shown in Table X.

Table X shows that the identification time length of the proposed method is less than those of the deep learning methods. It illustrates that the proposed method is better than the deep learning method in computation complexity.

According to the above analysis of performance comparison, the proposed method has similar performance with deep learning method in fire identification. However, the fire identification effect of the proposed method is mainly affected by the threshold value thr_{gray}. In order to ensure that the proposed method has a better robustness in real application for new data sets, the optimal threshold of each camera needs to be determined in advance.

IV. Conclusion

Fire is one of the most precarious events and it is very essential for disaster management to identify it in its initial stages. The traditional fire identification method does not consider the influence of fire-like objects and reflection of fire. The state-of-the-art deep learning methods are computationally expensive and require a lot of time to train model. With these motivations, a novel fire identification algorithm based on improved color segmentation and multifeature description has been proposed in this article. Our proposed method presents a judgment of reflection and nonreflection environments and constructs the determination model of optimal threshold value in fine fire segmentation. The variation coefficient of area, the dispersion of centroid, and the circularity of segmented images are proposed to be the statistics for fire identification. Experimental results show that the precision, recall, and $F1$-score of the proposed method are all over 97%. The proposed method does not need to train the model, and the identification time is lower than the deep learning methods. Therefore, the proposed method is computationally inexpensive for fire identification of surveillance video streams.

A limitation of the proposed method is that the optimal threshold parameter needs to be determined in advance. Simultaneously, the proposed method mainly focused on early fire identification. Future study will focus on conducting the fire identification automatically and localizing the fire according to the extraction of detailed information such as distance from the camera.

References

[1] L. Hu and Q. Ni, "IoT-driven automated object detection algorithm for urban surveillance systems in smart cities," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 747–754, Apr. 2018.

[2] S. Ramnath, A. Javali, B. Narang, P. Mishra, and S. K. Routray, "IoT based localization and tracking," in *Proc. Int. Conf. IoT Appl. (ICIOT)*, May 2017, pp. 1–4.

[3] G. Lu, Y. Yan, and M. Colechin, "A digital imaging based multifunctional flame monitoring system," *IEEE Trans. Instrum. Meas.*, vol. 53, no. 4, pp. 1152–1158, Aug. 2004.

[4] T. Qiu, Y. Yan, and G. Lu, "An autoadaptive edge-detection algorithm for flame and fire image processing," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 5, pp. 1486–1493, May 2012.

[5] A. E. Cetin *et al.*, "Video fire detection–review," *Digit. Signal Process.*, vol. 23, no. 6, pp. 1827–1843, 2013.

[6] A. K. Bhandari, I. V. Kumar, and K. Srinivas, "Cuttlefish algorithm-based multilevel 3-D Otsu function for color image segmentation," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 5, pp. 1871–1880, May 2020.

[7] B. C. Ko, K.-H. Cheong, and J.-Y. Nam, "Fire detection based on vision sensor and support vector machines," *Fire Saf. J.*, vol. 44, no. 3, pp. 322–329, Apr. 2009.

[8] C.-B. Liu and N. Ahuja, "Vision based fire detection," in *Proc. 17th Int. Conf. Pattern Recognit., (ICPR)*, vol. 4, Aug. 2004, pp. 134–137.

[9] C. Turgay, "Fast and efficient method for fire detection using image processing," *ETRI J.*, vol. 32, no. 6, pp. 1–12, Dec. 2010.

[10] G. Marbach, M. Loepfe, and T. Brupbacher, "An image processing technique for fire detection in video images," *Fire Saf. J.*, vol. 41, no. 4, pp. 285–289, Jun. 2006.

[11] T. Çelik and H. Demirel, "Fire detection in video sequences using a generic color model," *Fire Saf. J.*, vol. 44, no. 2, pp. 147–158, Feb. 2009.

[12] S. Calderara, P. Piccinini, and V. Cucchiara, "Smoke detection in video surveillance: A MoG model in the wavelet domain," in *Proc. 6th Int. Conf. Comput. Vis. Syst. (ICVS)*, May 2008, pp. 119–128.

[13] B. Ko, K.-H. Cheong, and J.-Y. Nam, "Early fire detection algorithm based on irregular patterns of flames and hierarchical Bayesian networks," *Fire Saf. J.*, vol. 45, no. 4, pp. 262–270, Jun. 2010.

[14] S. Cai, J. Liang, Q. Gao, C. Xu, and R. Wei, "Particle image velocimetry based on a deep learning motion estimator," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, pp. 3538–3554, Jun. 2020.

[15] B. D. L. Han, "Real-time fire detection using camera sequence image in tunnel environment," in *Proc. Int. Conf. Intell. Comput.*, Aug. 2007, pp. 1209–1220.

[16] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 1999, pp. 246–252.

[17] B. U. Toreyin, Y. Dedeoglu, and A. E. Cetin, "Contour based smoke detection in video using wavelets," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2006, pp. 1–5.

[18] W. Phillips III, M. Shah, and N. da Vitoria Lobo, "Flame recognition in video," *Pattern Recognit. Lett.*, vol. 23, nos. 1–3, pp. 319–327, Jan. 2002.

[19] T.-H. Chen, P.-H. Wu, and Y.-C. Chiou, "An early fire-detection method based on image processing," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, vol. 4, Oct. 2004, pp. 1707–1710.

[20] B. U. Toreyin, Y. Dedeoglu, and A. E. Cetin, "Flame detection in video using hidden Markov models," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2005, pp. 1230–1233.

[21] B. U. Töreyin, Y. Dedeoğlu, U. Güdükbay, and A. E. Çetin, "Computer vision based method for real-time fire and flame detection," *Pattern Recognit. Lett.*, vol. 27, no. 1, pp. 49–58, Jan. 2006.

[22] W.-B. Horng, J.-W. Peng, and C.-Y. Chen, "A new image-based real-time flame detection method using color analysis," in *Proc. IEEE Netw., Sens. Control*, Mar. 2005, pp. 100–105.

[23] P. V. K. Borges and E. Izquierdo, "A probabilistic approach for vision-based fire detection in videos," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 5, pp. 721–731, May 2010.

[24] S. Frizzi, R. Kaabi, M. Bouchouicha, J.-M. Ginoux, E. Moreau, and F. Fnaiech, "Convolutional neural network for video fire and smoke detection," in *Proc. 42nd Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Florence, Italy, Oct. 2016, pp. 877–882.

[25] S. Khan, K. Muhammad, S. Mumtaz, S. W. Baik, and V. H. C. de Albuquerque, "Energy-efficient deep CNN for smoke detection in foggy IoT environment," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 9237–9245, Dec. 2019.

[26] K. Muhammad, J. Ahmad, Z. Lv, P. Bellavista, P. Yang, and S. W. Baik, "Efficient deep CNN-based fire detection and localization in video surveillance applications," *IEEE Trans. Syst. Man, Cybern., Syst.*, vol. 99, no. 7, pp. 1–16, Jul. 2018.

[27] T. Toulouse, L. Rossi, T. Celik, and M. Akhloufi, "Automatic fire pixel detection using image processing: A comparative analysis of rule-based and machine learning-based methods," *Signal, Image Video Process.*, vol. 10, no. 4, pp. 647–654, Apr. 2016.

[28] B. Kim and J. Lee, "A video-based fire detection using deep learning models," *Appl. Sci.*, vol. 9, no. 14, pp. 28–42, 2019.

[29] T. Ono *et al.*, "Application of neural network to analyses of CCD colour TV-camera image for the detection of car fires in expressway tunnels," *Fire Saf. J.*, vol. 41, no. 4, pp. 279–284, Jun. 2006.

[30] R. Starosolski, "New simple and efficient color space transformations for lossless image compression," *J. Vis. Commun. Image Represent.*, vol. 25, no. 5, pp. 1056–1063, Jul. 2014.
[31] R. M. A. Eshaq, E. Hu, M. Li, and M. S. Alfarzaeai, "Separation between coal and gangue based on infrared radiation and visual extraction of the YCbCr color space," *IEEE Access*, vol. 8, pp. 55204–55220, 2020.
[32] T. Lin, "Research on video-based face recognition and tracking algorithm," *Comput. Modernization*, vol. 3, no. 1, pp. 121–129, 2013.
[33] H. J. Sun, T. Q. Deng, and Y. C. Li, "Image segmentation algorithm based on the improved watershed algorithm," *J. Harbin Eng. Univ.*, vol. 35, no. 7, pp. 856–894, 2014.
[34] T. M. Łapiński, "Multivariate Laplace's approximation with estimated error and application to limit theorems," *J. Approximation Theory*, vol. 248, no. 1, pp. 1–13, 2019.
[35] T. Nakaichi *et al.*, "Accuracy of metabolic volume and total glycolysis among six threshold-based target segmentation algorithms," *Ann. Nucl. Med.*, vol. 6, no. 4, pp. 321–333, 2017.
[36] P. Foggia, A. Saggese, and M. Vento, "Real-time fire detection for video-surveillance applications using a combination of experts based on color, shape, and motion," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 9, pp. 1545–1556, Sep. 2015.
[37] H. Schütze, C. D. Manning, and P. Raghavan, *Introduction to Information Retrieval*. Cambridge, U.K.: Cambridge Univ. Press, 2008.
[38] N. Martel-Brisson and A. Zaccarin, "Moving cast shadow detection from a Gaussian mixture shadow model," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, San Diego, CA, USA, Jun. 2005, pp. 643–648.
[39] O. Barnich and M. Van Droogenbroeck, "ViBe: A universal background subtraction algorithm for video sequences," *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1709–1724, Jun. 2011.
[40] M. M. Bilal and M. S. Hanif, "Benchmark revision for HOG-SVM pedestrian detector through reinvigorated training and evaluation methodologies," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 1–11, Mar. 2020.
[41] P. Li and W. Dazhao, "Image fire detection algorithms based on convolutional neural networks," *Case Stud. Thermal Eng.*, vol. 19, no. 1, pp. 1–11, 2020.

Xijiang Chen was born in Anhui, China. He received the Ph.D. degree from the School of Geodesy and Geomatics, Wuhan University, Wuhan, China, in 2014.

He was a Visiting Scholar with the University of Calgary, Calgary, AB, Canada, from 2018 to 2019, under the supervision of Derek Lichti. He is currently an Associate Professor with the School of Artificial Intelligence, Wuchang University of Technology, Wuhan, and the School of Safety Science and Emergency Management, Wuhan University of Technology, Wuhan.

Qing An was born in Hubei, China. He received the M.S. degree from the School of Artificial Intelligence, Huazhong University of Science and Technology, Wuhan, China, in 1995.

He is currently a Professor with the School of Artificial Intelligence, Wuchang University of Technology, Wuhan.

Kegen Yu (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Sydney, Sydney, NSW, Australia, in 2003.

He was with the Jiangxi Geological and Mineral Bureau, Nanchang, China; Nanchang University, Nanchang; the University of Oulu, Oulu, Finland; the CSIRO ICT Center, Sydney; Macquarie University, Sydney; the University of New South Wales, Sydney; and Wuhan University, Wuhan, China. He is currently a Professor with the School of Environmental Science and Spatial Informatics, China University of Mining and Technology, Xuzhou, China. He has coauthored the book *Ground-Based Wireless Positioning* (Wiley and IEEE Press, 2009, a Chinese version of the book is also available) and another book *Wireless Positioning: Principles and Practice* (Springer, 2018) and has authored the book *Theory and Practice of GNSS Reflectometry* (Springer, 2021). He has authored or coauthored more than 100 refereed journal articles, including more than 50 IEEE journal articles. He edited the book *Positioning and Navigation in Complex Environments* (IGI Global, 2018) and another book *Indoor Positioning and Navigation* (Science Press, 2018). His research interests include global-navigation-satellite-system reflectometry, ground-based and satellite-based positioning, and remote sensing.

Dr. Yu served on the Editorial Board of the *EURASIP Journal on Advances in Signal Processing*, the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS, and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY from 2013 to 2017. He is also on the Editorial Board of IEEE ACCESS. He was the Lead Guest Editor for a Special Issue of the Physical Communication on "Indoor Navigation and Tracking" and for a Special Issue of the *EURASIP Journal on Advances in Signal Processing* on "GNSS remote sensing." He is also the Lead Guest Editor of a Special Issue on Recent Advances in GNSS Reflectometry of *Remote Sensing*, which is currently open for submissions.

Ya Ban was born in Henan, China. He received the Ph.D. degree from the School of Geodesy and Geomatics, Wuhan University, Wuhan, China, in 2017.

He is currently a Senior Engineer with the Chongqing Academy of Metrology and Quality Inspection, Chongqing, China.