Fire Safety Journal 128 (2022) 103542

Contents lists available at ScienceDirect

Fire Safety Journal

journal homepage: www.elsevier.com/locate/firesaf

Early indoor occluded fire detection based on firelight reflection characteristics

Yakun Xie, Jun Zhu *, Yukun Guo, Jigang You, Dejun Feng, Yungang Cao

Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu, 611756, China

ARTICLE INFO

Keywords:
Indoor occluded fire
Early fire detection
Firelight reflection characteristics
Foreground recognition
Multiexpert system

ABSTRACT

Indoor fire detection is a challenging task and plays a key role in disaster management. The early stage of a fire is the best stage to extinguish the fire. However, early detection of indoor fires is difficult because the early stages of fires are easily occluded in complex indoor environments. Therefore, a method based on firelight reflection characteristics is proposed for early fire detection in occluded indoor environments. First, the characteristics of fires occluded by complex environments are described by analyzing the characteristics of firelight reflection. Second, a highly sensitive method for foreground recognition is developed through use of strategic background updates and a block binarization threshold, which are suitable for detecting the weak changes caused by occluded fires in videos. Finally, a multiexpert system is established for occluded fire detection by extracting the changing characteristics of the area in which the firelight reflection occurs, including spectral variability, motion persistence, and regional expansion. The accuracy and run time of the system are evaluated based on a large dataset to verify our method. Moreover, our proposed approach is discussed in detail in terms of effectiveness and applicability, and the results show that our method can be effectively applied in indoor occluded fire detection.

1. Introduction

Early indoor fire detection can effectively prevent damage and minimize losses [1–3]. However, the places where fires often occur indoors, such as sockets, wires, fireplaces, etc., are easily occluded by other objects, such as tables, chairs, beds, etc. The existing video-based fire detection methods cannot effectively recognize fires in these situations, and the early stages of fires are easily missed due to the production of occlusion. Therefore, the development of methods that quickly and accurately detect indoor occluded fires has important theoretical and practical significance.

The existing indoor fire detection methods can be divided into two categories. The first category is sensor-based fire detection, and the second category is video-assisted fire detection [4,5]. Sensor-based fire detection provides early warnings by detecting smoke, light, and temperature [6,7]. Although the widespread deployment of these detectors facilitates early fire detection, there are still significant limitations. First, the detector has a slower detection speed; in particular, the early features of occluded fires are weak, such as smoke particles and fire size. An alarm will only be issued after much time has passed, and the fire may be too large to control at this time. Second, the detector needs to be close to the fire ignition point. In addition, there must be no barriers between the fire and the detector, and the monitoring range is limited. Usually, a small number of fire alarm devices will be installed indoors. In a complex indoor environment, the sensing ability in the early stage of a fire is greatly restricted. Finally, the detector cannot provide information about the location, size, and extent of the fire, which is not conducive to achieving effective fire rescues. Sensor-based fire detection methods are not well suited for occluded indoor fire detection in the early phase. With the popularity of video surveillance equipment and the development of pattern recognition, video-assisted fire detection methods have attracted widespread attention from researchers. Various video-based automatic fire detection technologies are being developed and have been widely used in real life [8–10]. Video-based fire detection methods can be divided into two categories. The first is traditional fire detection based on manual feature extraction. The second is video-based fire detection based on deep learning.

Image pixel statistics are used as the main reference in the first type

* Corresponding author.
E-mail addresses: yakunxie@my.swjtu.edu.cn (Y. Xie), zhujun@swjtu.edu.cn (J. Zhu), gyk@my.swjtu.edu.cn (Y. Guo), vgeyjg523@my.swjtu.edu.cn (J. You), djfeng@swjtu.edu.cn (D. Feng), yungang@swjtu.cn (Y. Cao).

https://doi.org/10.1016/j.firesaf.2022.103542
Received 19 June 2021; Received in revised form 27 December 2021; Accepted 1 February 2022
Available online 4 February 2022

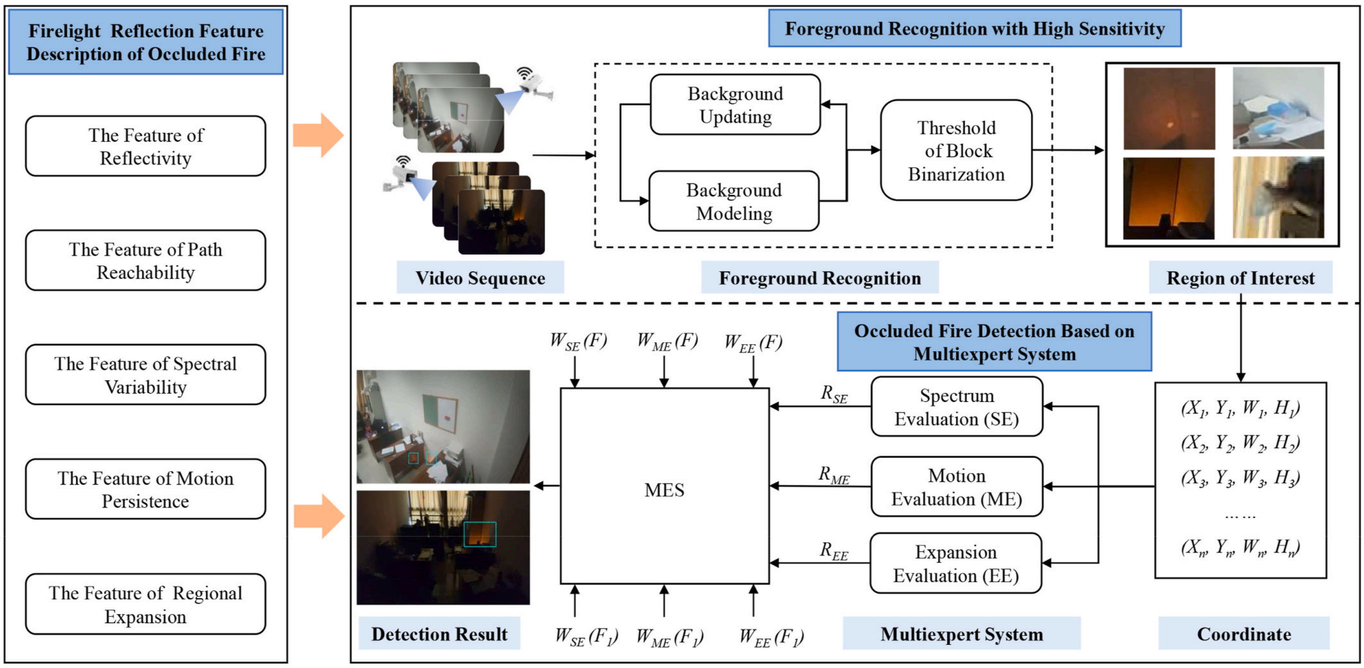

Fig. 1. The framework of the proposed approach for indoor occluded fire detection.

of fire detection method, such as those used in different color models, including RGB, YCbCr, and YUV [11–14]. However, an image fire detection algorithm with a single feature cannot meet the application requirements due to the large interference from environmental factors. This type of method is gradually being replaced by multifeatured methods, such as methods that consider the color, texture, shape, movement and growth of the fire [15–18]. In later research, researchers established rule-based algorithms or multidimensional feature vector methods on the basis of acquiring features as the input of traditional classification algorithms, such as support vector machines, AdaBoost, and hidden Markov models [19–23]. Although the method of combining multiple features prevents the interference of some environmental factors, the false alarm rate is still high, and the detection accuracy needs to be further improved. In addition to ordinary visible-light surveillance videos, fire detection methods based on special monitoring equipment are also constantly being proposed. For example, integrated infrared cameras are used as a way to supplement surveillance systems to take advantage of both visual and thermal characteristics for fire detection [24–26]. Ultrasonic cameras are also used for fire detection to overcome the limitations of visible-light cameras [27]. According to the experimental results, they achieved good results. However, the high camera cost leads to greater limitations and makes the method unsuitable for large-scale promotion and applications.

In recent years, with the development of deep learning technology, a variety of neural network models have been proposed, such as convolutional neural networks (CNNs) [28], recurrent neural networks (RNNs) [29], and deep belief networks (DBNs) [30]. These networks perform a variety of high-performance computer vision tasks, such as image processing, object detection [31–33], natural language processing [34], speech recognition [35] and other tasks [36,37]. Among them, CNNs have achieved good image classification results. On the basis of this research, many neural network algorithms have also been applied to fire detection [1]; Sharma et al., 2017; [7,38–42]. Although CNN-based methods have shown excellent performance, the existing methods that use deep learning for video-based unoccluded fire detection cannot detect fires that are completely occluded indoors. In addition, different from the spectrum, texture and dynamic characteristics of unoccluded fires, the reflection of light on the surface of one object has very different feature expressions from that on the surface of another object, and the existing training datasets are unable effectively support detection in these situations.

The above methods perform well in video-based unoccluded fire detection. However, we considered early-stage occluded fires, which is a special and common indoor situation. In this case, there are no obvious unoccluded fire features in the video, which also makes it impossible to use the previous manual features or deep learning features for fire detection. These problems are further solved in this article. The main contributions of our work are summarized below.

(1) Early occluded indoor fire detection based on firelight reflection characteristics is proposed. To the best of our knowledge, this is the first time that indoor occluded fires in videos have been studied. This is a more challenging task compared with unoccluded fire detection, and the experiments show that our method has better performance in terms of accuracy and speed.
(2) We propose a highly sensitive foreground recognition method. The method of background updating is strategically adjusted through the weak characteristic changes caused by the fire reflected from the object surface. This method solves the problem of the weak expression of changing characteristics reflected in early-stage fires.
(3) A multiexpert system (MES) was established for occluded fire detection. Expert knowledge is extracted from the three aspects of spectral variability, motion persistence, and regional expansion by analyzing the characteristic expression of the fire reflected on the surface of an object. Multiexpert knowledge is used to establish an MES to realize the effective detection of occluded fires.

The remainder of this article is structured as follows. The proposed method is introduced in Section 2, including the overall framework, the description of the reflection feature of an occluded fire, foreground recognition with high sensitivity, and occluded fire detection based on an MES. In Section 3, the dataset description, evaluation metrics, and experimental results are described in detail. The results of this paper are discussed in Section 4. Finally, conclusions and future work are presented in Section 5.

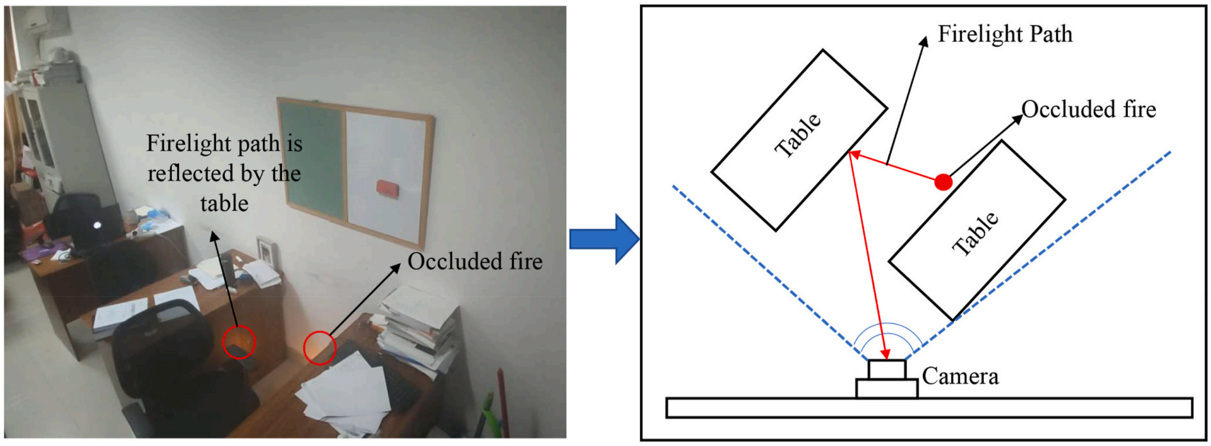

Fig. 2. Description of the firelight path change after reflection by the table.

2. Proposed method

In this section, the proposed method is described in detail. First, the overall framework of the method in this paper is introduced. Second, we describe the characteristics of the indoor firelight reflection area by analyzing light reflection features. Finally, indoor occluded fire detection is introduced in detail, including foreground recognition with high sensitivity and occluded fire detection based on an MES.

2.1. The overall framework of our method

To improve the accuracy and efficiency of indoor occluded fire detection, a study based on firelight reflection characteristics is proposed, as shown in Fig. 1. The proposed framework is divided into three main parts. First, the characteristic changes in fire reflection areas in a video are described by analyzing the reflection characteristics of the firelight. Second, a highly sensitive foreground recognition method is established to detect weak changes in the fire reflection areas in the video as much as possible. This method uses a new background update strategy and a regional binary threshold method to extract the region with weak changes in the video. In addition, the coordinates of the region of interest of the fire are also recorded in this part. Finally, expert knowledge is extracted from the three aspects of spectral variability, motion persistence, and regional expansion by analyzing the characteristic expression of the fire reflected on the surface of an object; the analysis includes spectrum evaluation (SE), motion evaluation (ME), and expansion evaluation (EE). above three types of expert knowledge are used to establish an MES. In addition, the presence of an indoor occluded fire is determined by the MES classifier, which is based on a weighted voting rule.

2.2. Description of the firelight reflection feature of occluded fires

There are inevitably many blind spots in indoor surveillance videos, which makes it impossible to directly display and record the process of a fire. In addition, blind spots also make the previous methods of video-based fire detection unusable. However, according to the principle of light reflection, the firelight path will change its propagation direction and will be reflected on the surfaces of objects, such as walls, tables and chairs, when different substances are transmitted, as shown in Fig. 2. This reflection changes the feature expression of the surface of an object in the video; this change in feature expression allows for the possibility of using videos to detect occluded fires.

Different from the characteristic expression produced by an unoccluded fire, the reflected light is complex and weak after being superimposed with the surface pixels of the object. Therefore, we performed multiple sets of occluded fire experiments in different indoor environments to explore the characteristic expression of firelight reflection as

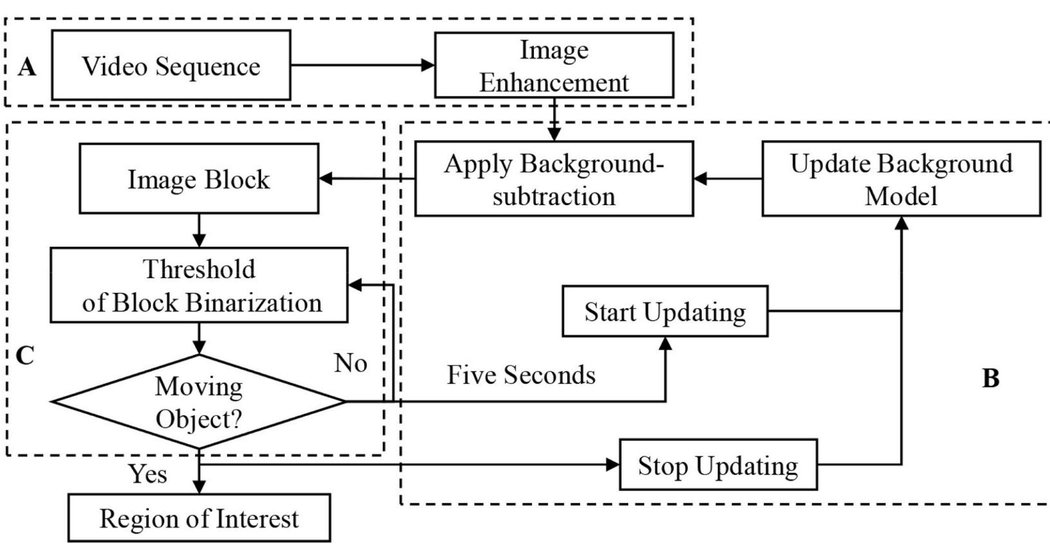

Fig. 3. Flow chart of foreground detection with high sensitivity.

extensively as possible. We described the features of occluded fires in detail according to the results of the experimental analysis, as shown below.

(1) Reflectivity of objects: objects can absorb and reflect the light generated by a fire, and most of the substances that exist in reality have a certain light reflection effect, which provides conditions suitable for occluded fire detection with videos.
(2) Reachability of the reflection route: there is generally no absolute specular reflection in an indoor environment; diffuse reflection occurs when light is projected on an object. Therefore, a fire in an occluded environment can be reflected in a video regardless of the camera angle.
(3) Spectral variability: the spectral characteristics of the surface of an object merge with the light when the light generated by a fire is projected onto the surface, which will cause the spectral characteristics of the surface of the object to change.
(4) Motion persistence: a fire produces continuous high-frequency timing changes during the combustion process, which is called the flicker characteristic. Correspondingly, fire produces light, which is projected onto the surface of an object. Although the flicker is different from the regularity of the fire, it still causes continuous movement changes when it is projected onto the surface of an object.
(5) Regional expansion: a fire is small in the early stages, and the areas of light projected on the surfaces of other objects are also small. However, the burning range of the fire will continue to expand to the surroundings as time increases, and the change in the reflective area caused by the fire on the surface of an object will also expand.

2.3. Foreground recognition with high sensitivity

Generally, foreground recognition methods, including the frame difference method (FD) [43], Gaussian mixture method (GMM) [44] and K-nearest neighbor method (KNN) [45], are used to detect moving areas in videos. Although these methods have a good effect on the detection of rigid objects, they cannot be effectively applied to detect changes caused by reflected light. Therefore, we propose a method of foreground recognition with high sensitivity by analyzing the changes caused by the reflection of a fire on the surface of an object. The method is divided into three steps, i.e., image enhancement, strategic background updating, and block binarization thresholding, as shown in Fig. 3.

2.3.1. Image enhancement

Histogram stretching is applied for image enhancement and can further improve the visual effect of a video. In addition, this method satisfies the sensitivity requirements to capture the changing areas in a video, highlights the features of interest and suppresses interference features, as shown in Fig. 3 (A).

2.3.2. Strategic background updating

Even if the background in a surveillance video is basically unchanged when there are no people, there may exist slight changes in the video due to the influence of other factors, such as light, wind and other factors, in the surrounding environment. Therefore, we made strategic dynamic updates to the background model, as shown in Fig. 3 (B).

First, the static background is updated in stages, which takes into account that the indoor surveillance video remains basically unchanged and the redundancy between frames is large. A study showed that the K-nearest neighbor (KNN) method has better performance in static scenarios [45]. In the initial stage of a video, we use KNN for background modeling. If there is no change in the foreground, no real-time background model updates will be performed for a certain period of time, which is set to 5 s. Second, the updating of the background model will be discontinued if there is a change in the foreground, and the next detection step is performed to determine whether a fire has occurred. The reason for using this strategy is that the light reflection of a fire is inherently weak. The model will be less sensitive to changing areas in foreground recognition if the background model is updated in real time. Compared with real-time updates, the strategy proposed in this article fits the subtle changes in a video better and prevents unnecessary calculations.

2.3.3. Threshold of block binarization

The foreground and the background need to be distinguished according to the preset threshold after the background is modeled. The fire reflection area is not only weak but also dynamically changing. The traditional method has many disadvantages, such as when the fixed threshold is set too high, the method easily misses small targets, and when it is set too low, global false detection occurs. We have established a threshold method for regional binarization to reduce missed detection and false detection, as shown in Fig. 3 (C). The main steps of the algorithm are as follows.

(1) The first is the image block area division, as shown in Eq. (1). The finer the granularity of the image division is, the more obvious the effect and the slower the speed.

$$w = \frac{W}{N}, \quad h = \frac{H}{N} \tag{1}$$

where W and H represent the width and height of the image, respectively, and $\lfloor\ \rfloor$ indicates rounding down. N is the number of blocks and is set to 10. For the setting of N, we considered the balance of accuracy and speed and the characteristics of the small size of the area in the early stage of a fire. Moreover, the number of blocks will include the adjacent area if there is an undivided edge area in the calculation process.

(2) The reflection of a fire not only causes changes in the spectrum but also causes changes in the original surface texture of an object, and relatively speaking, the texture changes are more obvious. Therefore, we use the inverse differential moment (IDM) to describe the image texture changes, which can reflect the homogeneity and measure the change value of textures, as shown in Eq. (2). The IDM is larger if the image texture changes slowly, and vice versa. IDM_q is used to measure the speed of regional texture changes, as shown in Eq. (3).

$$\mathrm{IDM} = \sum_{a,b} \frac{P_{\varnothing,\,d}(a,b)}{1+(a-b)^2} \tag{2}$$

$$\mathrm{IDM}_q = \frac{1}{4}\sum_{\varnothing}\sum_{a_q,b_q} \frac{P_{\varnothing,\,1}(a_q,b_q)}{1+\left(a_q-b_q\right)^2} \tag{3}$$

where the distance between a and b is d and $\varnothing$ is the angle, which represents the gray level of two pixels. $P_{\varnothing,d}(a,b)$ represents the elements of the co-occurrence matrix and is the number of occurrences of (a,b) gray level pairs.

(3) The difference probability is corrected according to the reciprocal difference moment of each area to further improve the sensitivity of the method and reduce the missed detection rate, as shown in Eq. (4).

$$\mathrm{Thr} = G^{-1}\left(\frac{\mathrm{IDM}_q}{\mathrm{IDM}_{\max}} p_G; \eta, m\right) \tag{4}$$

$$\mathrm{IDM}_{\max} = \max\left(\mathrm{IDM}_1,\ \mathrm{IDM}_2, \ldots, \mathrm{IDM}_{N^2}\right) \tag{5}$$

where p_G is the preset probability that is set to 0.9 according to some experimental results and η and m represent the scale and shape parameters, respectively. Thr is the threshold in different block binarizations.

2.4. Occluded fire detection based on multiexpert systems

It is impossible to determine whether a single change in the foreground is caused by foreground recognition, and further discrimination of the change area is needed. Three types of expert knowledge are extracted through the analysis of firelight reflection characteristics, including SE, ME, and EE. Then, an MES is established for occluded fire detection.

2.4.1. Spectrum evaluation

We describe the changes in the spectrum from three aspects by analyzing the changes produced by the reflection of firelight on the surface of an object, including changes in mutation, persistence, and uniformity characteristics. We establish the spectrum change rule and define the data structure of the spectrum change, as shown in Eq. (6).

$$Spectrum = \{Mutation, Persistence, Uniformity\} \tag{6}$$

where *Mutation* denotes a color mutation, *Persistence* denotes mutation persistence, and *Uniformity* denotes mutation consistency.

(1) Mutation. A sudden change in color will occur on the surface of an object when the light reflected by a fire is projected onto the surface of the object. The coordinate position (*x, y, w, h*) of the moving area in the video image is obtained according to the foreground recognition result. We use the equal-weighted averages of the three channel colors to calculate brightness. The brightness I_t is equal to 0.30R + 0.59G+0.11B, according to the transformation from the RGB color space into the YUV color space, to prevent floating-point calculations and reduce the number of calculations [46], as shown in Eq. (7). Furthermore, a pixel frequency oscillation matrix $SUM_t x$, y with the same size as the video motion area is created and is used to analyze the changes in brightness at the coordinates (x , y) of each pixel in the moving area in each frame. The brightness calculation method is shown in Eq. (8).

$$I_t x\ ,\ \mathrm{y} = \frac{1}{3}[\mathrm{R_t(x\ ,\ y) + G_t(x\ ,\ y) + B_t(x\ ,\ y)}] \tag{7}$$

$$SUM_T = |SUM_t x\ ,\ \mathrm{y} - \mathrm{SUM_0 x\ ,\ y} \tag{8}$$

where I_t represents the pixel brightness value at time t; R_t, G_t, and B_t represent the pixel values of each band at time t; x , y represents the coordinates of the pixel in the image; $SUM_t x$, y represents the pixel matrix of the detected area of interest; $SUM_0 x$, y represents the initial matrix of pixels before the detected area of interest changed; and SUM_T represents the pixel change value.

(2) Persistence. The sudden change caused by the fire is different from a general spectral change. This type of sudden change will always exist in the early development process of a fire and is called mutation persistence. We compared the data before foreground recognition in the region of interest with the data after the change in consecutive frames and extracted the mutation persistence characteristics by counting the number of changes in the interest area in 10 frames, as shown in Eq. (9).

$$T = \begin{cases} |SUM_t x\ ,\ y - SUM_0 x\ ,\ y,\ if(SUM_T \geq T_I) T = T + 1 \\ |SUM_t x\ ,\ y - SUM_0 x\ ,\ y,\ if(SUM_T < T_I) T = T + 0 \end{cases} \tag{9}$$

where SUM_T represents the interpolation of the initial frame of the adjacent 10 frames of the interest area matrix, T_I represents the pixel mutation threshold, and T represents the number of continuous frames that show a sudden change. An area is considered to have the characteristics of a continuous sudden change if $T \geq 5$.

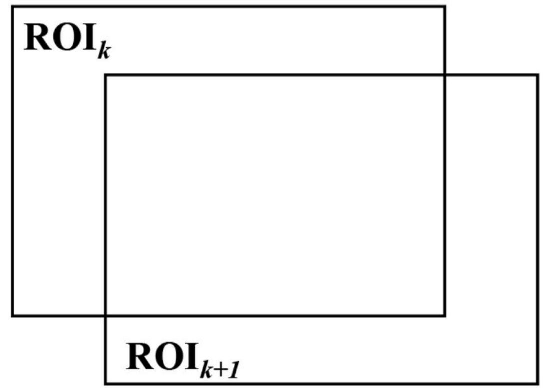

Fig. 4. Inter frame overlap.

(3) Uniformity. The color change caused by the mutation will maintain the consistency of the change in the pixel value difference for a short time, which is called the mutation consistency. We use the calculation result of Eq. (8) to extract information on the mutation consistency, which can help reduce the number of calculations, as shown in Eq. (10).

$$SUM_t = \left(\sum_{i=1}^{10} \left| (SUM_0 - SUM_i) \right| \right) \Big/ 10 \tag{10}$$

where SUM_0 represents the region-of-interest matrix of the initial frame, SUM_i represents the interest area matrix of each of the 10 frames after the mutation, and SUM_t represents the average change in the difference between 10 consecutive frames.

2.4.2. Motion evaluation

A burning fire has a development process, and similarly, its reflection area also has such a development process. This development process is a continuous process. Therefore, the position of the region of interest in consecutive frames is guaranteed to have a certain degree of overlap, as shown in Fig. 4. The degree of overlap between frames is calculated by the formula shown in Eq. (11).

$$\mathrm{Overlap} = \frac{Area_{ROI_k} \cap Area_{ROI_{k-1}}}{Area_{ROI_k} \cup Area_{ROI_{k-1}}} \tag{11}$$

where k represents the current frame, ROI_k represents the region of interest acquired in the current frame, $Area_{ROI_k}$ represents the area of interest area obtained in the current frame, and Overlap represents the degree of the overlap coefficient.

In addition, the temporal characteristics of the degree of overlap are constructed on the basis of the degree of overlap between frames. Taking 10 frames as the threshold, the existing video frame rate is generally approximately 20 frames. The region of interest is considered to have motion characteristics if there are more than 5 overlap occurrences between frames within 10 frames, as shown in Eq. (12).

$$Motion_T = \begin{cases} \mathrm{Overlap}_i + 1,\ if(\mathrm{Overlap}_i \geq \mathrm{Overlap}_T) \\ \mathrm{Overlap}_i + 0,\ if(\mathrm{Overlap}_i < \mathrm{Overlap}_T) \end{cases} \tag{12}$$

where $i \in [1, 10]$ and $\mathrm{Overlap}_i$ represents the overlap coefficient between frame number *i* and frame number *i*-1. $\mathrm{Overlap}_T$ represents the threshold of the overlap coefficient, which is set to 0.1. $Motion_T$ represents the number of exercise sequences, and the area of interest is considered to contain a fire if the value is greater than 5.

2.4.3. Expansion evaluation

The characteristics of a fire reflection area in a video are different from those of a rigid object with specific contours and features, and a fire also has a certain regional expansion expression even in the early stage.

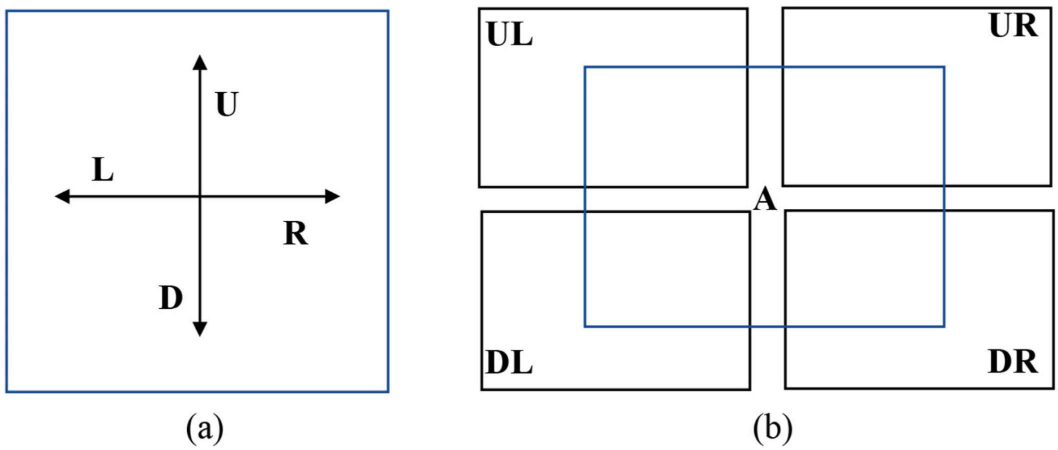

Fig. 5. Description of regional expansion.

Similarly, the area of the surface of an object impacted by the reflection of the firelight is also expanding.

A fire that is not occluded has a tendency to move upward. However, although the reflected fire appears to expand, it can expand in all directions, including up (U), down (D), left (L), and right (R), as shown in Fig. 5 (a). In addition, the direction of movement may also be in other nonpositive directions, including the upper left (UL), down left (DL), upper right (UR), and down right (DR) directions, as shown in Fig. 5 (b). Therefore, we determine the center point of the interest area from the overlapping region between frames, as shown in Eq. (8). The region of interest is considered to have the characteristics of regional expansion if the coordinates of the center point of the overlapping region do not exceed the range of the initial frame.

$$Expand = I_{k,i}\left(x_1^{k,i}, y_1^{k,i}\right) - I_{k-1,i}\left(\left(x_1^{k-1,i}\right), \left(y_1^{k-1,i}\right)\right) \tag{13}$$

where *Expand* is the expandability of the region, k is the current frame, i is the continuous region between frames obtained according to ME, and (x, y) is the coordinate of the center point.

2.4.4. Multiexpert system

MESs have been successfully applied to a variety of computer vision tasks, such as face detection. Although various new strategies have been proposed in previous studies, one of the main factors that determines the performance of an MES is the composition rules for different tasks. For the detailed principle and formula of an MES, we refer to Foggia et al. [16]. However, we have established three different expert rules that are more suitable for the occluded fire detection task.

3. Experiments and results

3.1. Dataset descriptions

The rationality of the dataset is the basis for verifying the effectiveness of the method. Many fire detection datasets can be found in the existing research [16,22,47–50]. Although these datasets contain large numbers of images, they are all unoccluded fire data and cannot be applied to the detection of occluded fires. Therefore, we established an occluded fire detection dataset. The final dataset included 20 videos, of which 10 were occluded fire videos (including 7990 images) and 10 were nonfire videos (including 23,325 images). More information about the videos is shown in Table 1. The dataset was divided into the following two parts: 20% was used for training to determine the weights of the MES, while 80% was used to test the proposed approach to further verify the effectiveness of the method [16,51].

In particular, both fire and nonfire videos take into account the complexity of real indoor scenes. First, the occluded fire videos contain different indoor scenes, such as home, office, and school scenes. In addition, the influence of light is also considered in our dataset, which contains day and night videos. Second, the nonfire videos contain more interference situations, such as movement caused by wind and people, which can lead to misclassification. In general, this is a challenging dataset that can help us test the method under a variety of conditions that may occur in a real environment. Some visual examples are shown in Fig. 6.

3.2. Evaluation metrics

To quantitatively evaluate the performance of our proposed fire detection method and to compare our results with the results of other researchers, the false positive rate (also referred to as the false alarm rate) (Eq. (14)), false negative rate (Eq. (15)) and accuracy (Eq. (16)) are used. The goal of this paper is a high accuracy rate, a low false positive rate and a low false negative rate. In addition, to evaluate and compare the effects of various foreground detection methods, we further evaluated the recall rate (Eq. (17)) of the method. In addition, frames per second (fps) are used to evaluate the detection speed, which is the average number of video frames that can be detected per second.

$$False\ positive\ rate = \frac{FP}{FP + TN} \tag{14}$$

$$False\ negative\ rate = \frac{FN}{FN + TP} \tag{15}$$

$$Accuracy\ rate = \frac{TP + TN}{TP + FN + FP + TN} \tag{16}$$

$$Recall\ rate = \frac{TP}{TP + FN} \tag{17}$$

TP, FN, FP and TN are the evaluation metrics obtained by comparing the fire detection results with the ground truth.

TP: the number of true positives, i.e., number of correctly detected occluded fire regions;

FN: the number of false negatives, i.e., number of misclassified occluded fire regions;

FP: the number of false positives, i.e., number of erroneously detected occluded fire regions;

TN: the number of true negatives, i.e., number of correctly detected nonfire regions.

3.3. Experimental results

Qualitative and quantitative analyses are used to evaluate the results

Table 1
Experimental dataset description.

Video	Resolution	Frame rate	Frames	Fire	Notes
video 1	856 × 480	20	2300	Yes	A fire blocked by a desk and chair in a classroom. The video was acquired by the authors.
Video 2	856 × 480	20	300	Yes	A fire blocked by a desk in an office during the day. The video was acquired by the authors.
video 3	856 × 480	20	180	Yes	A fire blocked by a desk in an office at night. The video was acquired by the authors.
Video 4	856 × 480	20	400	Yes	A fire blocked by a desk in a lab at night. The video was acquired by the authors.
Video 5	856 × 480	20	860	Yes	A fire blocked by a person moving in a living room. The video was acquired by the authors.
video 6	856 × 480	20	1160	Yes	A fire blocked by a desk in a bedroom. The video was acquired by the authors.
Video 7	856 × 480	20	420	Yes	A fire blocked by a wall in a living room. The video was acquired by the authors.
video 8	856 × 480	20	400	Yes	A fire blocked by a chair in a bedroom. The video was acquired by the authors.
video 9	856 × 480	20	1410	Yes	A fire blocked by a refrigerator in a living room. The video was acquired by the authors.
Video 10	856 × 480	20	900	Yes	A distant fire blocked by a desk in a classroom. The video was acquired by the authors.
video 11	800 × 600	20	380	No	A person moving in a lab with a red notebook. The video was download from Ref. [16].
video 12	320 × 240	24	936	No	Some smoke in a room. The video was acquired by the authors.
video 13	360 × 288	25	225	No	A person moving in room and the TV is playing. The video was acquired by the authors.
video 14	800 × 600	15	1485	No	A person moving in a lab holding a red ball. The video was downloaded from Ref. [16].
video 15	320 × 240	24	624	No	Some smoke in a room. The video was acquired by the authors.
video 16	640 × 368	15	3195	No	Many people moving in a vegetable market. The video was acquired by the authors.
video 17	640 × 360	15	435	No	A person patrolling a factory with a flashlight. The video was acquired by the authors.
video 18	640 × 360	25	15,000	No	A person moving in a public hall. The video was acquired by the authors.
video 19	1280 × 720	30	270	No	Video with light changes and moving people in an office. The video was acquired by the authors.
Video 20	672 × 378	25	775	No	Many people are cleaning in a factory. The video was acquired by the authors.

of this article to prove the effectiveness of the method.

3.3.1. *Qualitative analysis*

The detection results from different scene types and lighting conditions are displayed visually to fully present the test results of the method in this article in different types of scenes. The video is divided into five equal parts according to the time required to perform a fair visual display. The test result is shown in Fig. 7. The figure shows that the test results in each stage have a good effect in recognizing videos that contain occluded fires. The occluded fires in the videos can be continuously detected, and there is no missed detection. In addition, more interference, such as moving pedestrians and changes in light, is contained in the videos that do not contain occluded fires. Although there is a false detection result, the method has a good ability to filter common types of interference. The proposed approach is effective in recognizing fire in the preliminary qualitative visual display.

3.3.2. *Quantitative analysis*

The method in this paper is quantitatively analyzed to further prove the reliability of the method by using a confusion matrix. Fig. 8 shows that the TP and TN in this paper have high values, which can also intuitively show that the effectiveness of the method. In addition, to further quantify the accuracy of the method, the evaluation criteria include the accuracy rate, false positive rate, and false negative rate, as shown in Table 2. The average accuracy of our method is 92.80%, and the false positive rate and false alarm rate are 10.98% and 2.76%, respectively. According to the experimental results, our method has very promising performance considering the dataset, and because the dataset in this article contains many challenging factors, this effect is very impressive.

In addition, we conducted ablation experiments on multiple components of our method. We successively removed different key components for comparison on the basis of foreground recognition to prove that each component contributes to our method, as shown in Table 2. The experiment is divided into two groups. The first group of experiments is based on foreground recognition by adding three components separately for judgment. The false positive rate is increased by 8.96%–14.69%, the false negative rate is increased by 4.05%–7.29%, and the accuracy is reduced by 5.33%–12.18%. The second group of experiments uses the expert system to combine three components in pairs. The false positive rate is increased by 2.43%–7.28%, the false negative rate is increased by 2.11%–5.58%, and the accuracy is reduced by 1.94%–8.28%. The comparative effects of the two sets of ablation experiments show that each component's contribution to the system is very obvious. As the above analysis shows, considering the SE, ME, and EE of an occluded fire can effectively improve the accuracy of occluded fire detection and reduce the rates of false positives and false negatives. On the other hand, we highlight that this is the first study to introduce research on indoor occluded fire detection, and the proposed method is very effective.

4. Discussion

To the best of our knowledge, this is the first study on indoor occluded fire detection. We discussed the effectiveness of our method in three parts. First, the foreground recognition effect of this paper was discussed. Second, the speed of the method in this paper was evaluated in detail, including equipment with different performance results. After that, we discussed the time from the initial occurrence of the fire to the detection of the obstructed fire in view of fire discovery. Finally, we discussed the robustness of the method to disturbance, which is also one of the important criteria in proving whether the method is efficient.

4.1. *Analysis of foreground recognition with high sensitivity*

The recall rate should be increased as much as possible in the

Fig. 6. Examples of images extracted from the dataset. The top row shows frames extracted from videos in which a fire is present, while the bottom row shows frames extracted from videos with no fire present.

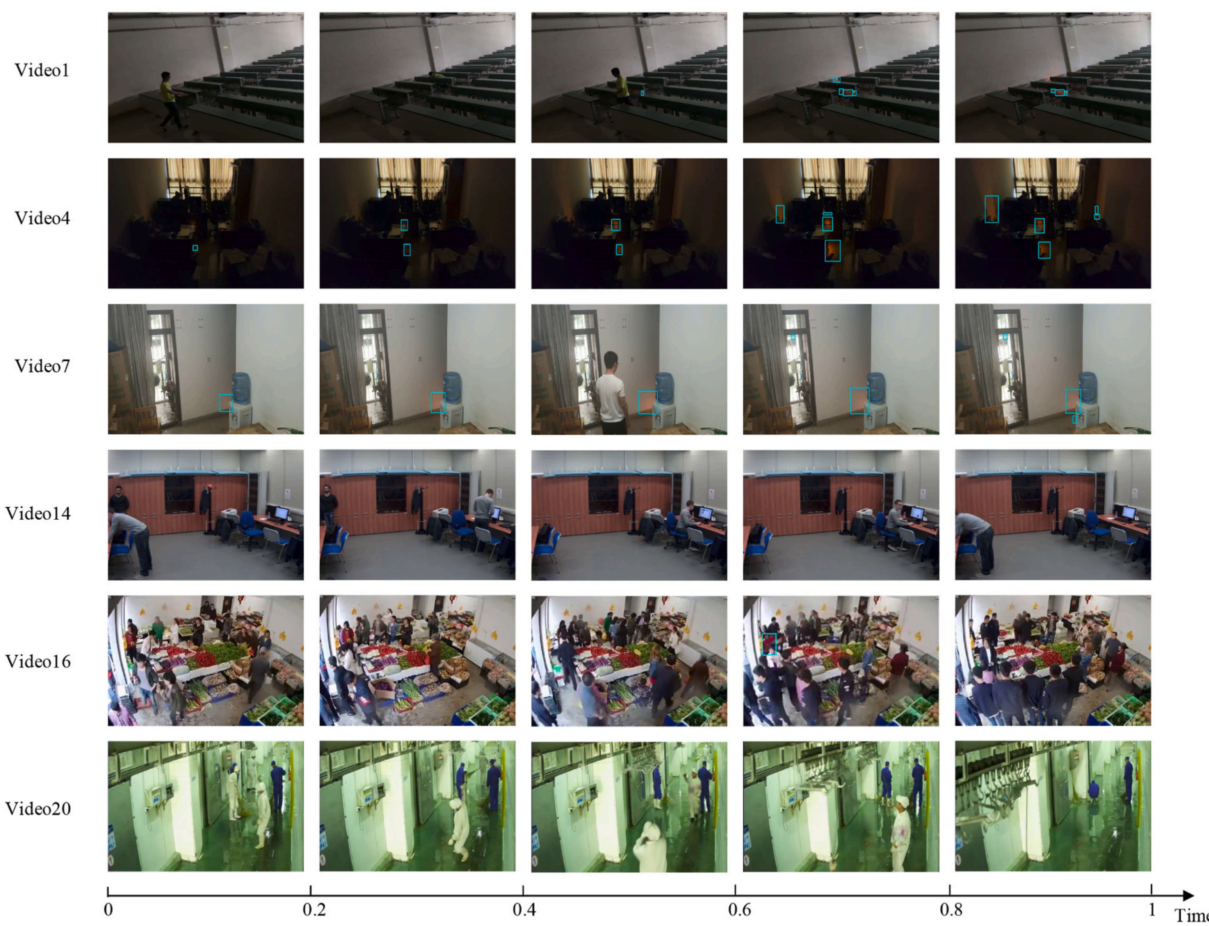

Fig. 7. Results of the proposed approach. The top three rows show frames extracted from videos with fires present, while the bottom three rows show frames extracted from videos with no fires present.

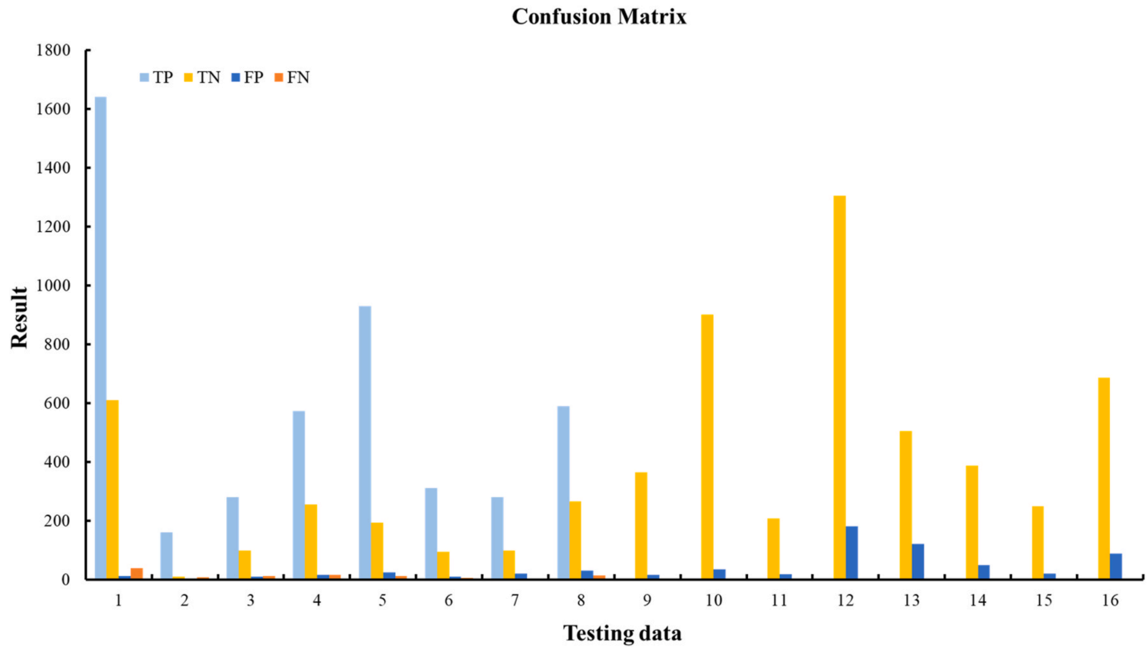

Fig. 8. The occluded fire detection confusion matrix.

Table 2
The results of the proposed methods in terms of the accuracy rate, false negative rate, and false positive rate.

Typology	Method	Accuracy rate (%)	False positive rate (%)	False negative rate (%)
Single Expert	SE	87.47	18.94	6.81
	ME	80.62	25.67	10.05
	EE	82.39	22.33	9.38
MES	SE + ME	89.65	14.59	5.91
	SE + EE	90.86	13.41	4.87
	ME + EE	84.32	19.26	8.34
Our		**92.80**	**10.98**	**2.76**

foreground recognition process to ensure that the MES can better judge whether an object is blocking a fire, which can reduce false negatives. We compare the recall rate of our method with those of the commonly used foreground recognition algorithms, including the FD [43], GMM [44], and KNN [45], to prove the effect of the foreground recognition method. The recall rate of fire videos 1–5 in the dataset is tested, and the

Table 3
Detailed comparison of the speeds of different configuration devices.

Methods	Fps	Remarks
Our method	46	Intel(R) Core(TM) i7-8750H with 16 GB of RAM
	38	Intel(R) Core(TM) i7-4810MQ with 16 GB of RAM
	27	Intel(R) Core(TM) i7-5500U with 8 GB of RAM

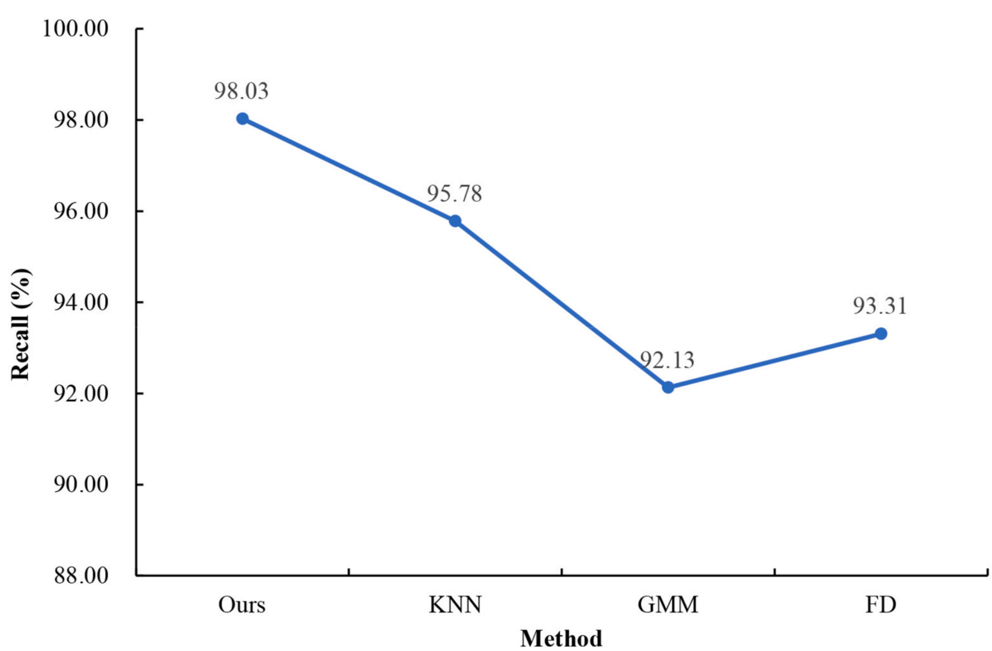

Fig. 9. Comparison of the proposed method with other famous methods in terms of recall rate.

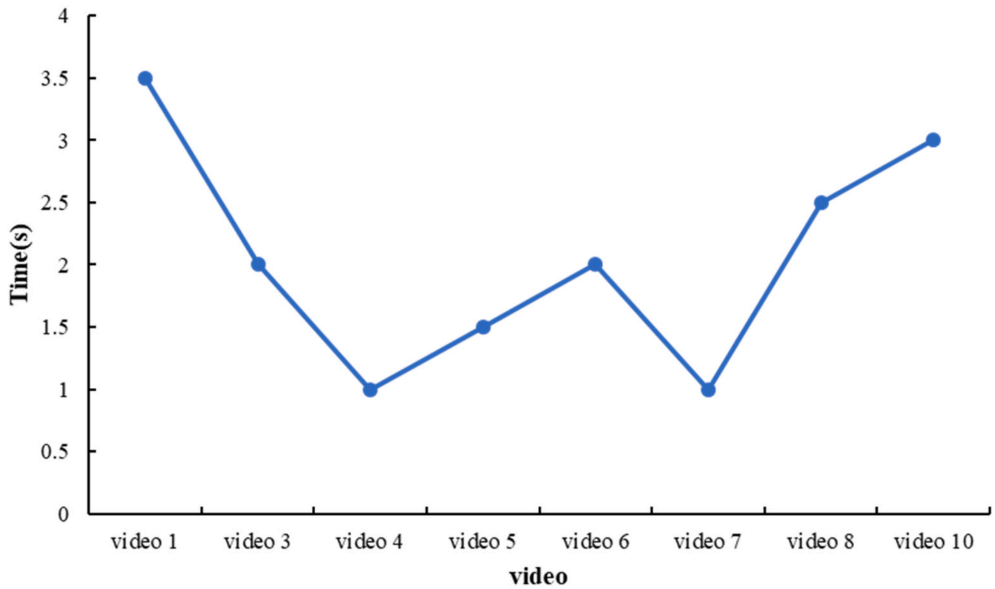

Fig. 10. Time required for occluded fire detection in each video.

(a) Disturbance in fire video (a) Disturbance in nonfire video

Fig. 11. Visual analysis of our method with disturbance.

test results are shown in Fig. 9. The figure shows that foreground recognition with high sensitivity has a higher recall rate than the traditional methods, and the recall rate is increased by 2.25%–5.90%. The recall rate can be increased through strategic background updates and the block binarization threshold, which are extremely important for reducing the missed detection rate in occluded fire detection.

4.2. Effectiveness of our method

In addition to accuracy, the detection speed is an important condition for determining the feasibility of the occluded fire detection method. We tested three different configuration devices, including an Intel (R) Core (TM) i7-8750H with 16 GB of RAM, Intel (R) Core (TM) i7-4810MQ with 16 GB of RAM, and Intel (R) Core (TM) i7-5500U with 8 GB of RAM, and the results are shown in Table 3. Real-time detection can be achieved with the three different configuration devices, and the detection speeds are 46 fps, 38 fps, and 27 fps. In addition, the frame rate of ordinary surveillance video is generally 20–25 fps, and our method can still reach 27 fps even with limited computer resources; hence, our method can achieve real-time detection.

4.3. System feasibility analysis of our method

A fire can be divided into five stages according to the changes in fire characteristics during the burning process, including ignition, development, fierce burning, decay, and extinction [41]. The early stage of development of a fire is the best time to extinguish it, and the time elapsed between the ignition of a fire and its detection is an important factor to consider when evaluating the ability of a method to achieve early fire detection. However, when a fire is occluded at the initial stage, the fire has reached the development stage if an unoccluded fire is visible in the video, which is not conducive to the control of the fire. In addition to the commonly used evaluation criteria of accuracy and run time, the time from ignition to detection was recorded, and the result is shown in Fig. 10. The amounts of time elapsed between the time the fires ignited and the occluded fires were detected in the 8 videos were 3.5 s, 2 s, 1 s, 1.5 s, 2 s, 1 s, 2.5 s, and 3 s. For all videos, it was possible to detect occluded fires within 4 s with high accuracy. This analysis reveals that the proposed approach can achieve more accurate detection in the early stages of an occluded fire.

4.4. Disturbance analysis of our method

To further analyze the antidisturbance ability of our method, we analyzed the scene disturbance in the video, which includes various types of disturbances, such as computer screens and reflection, flashing lights, personnel, and TV, as shown in Fig. 11. Fig. 11 (a) is a video containing an occluded fire, in which the red box is the disturbance. The figure shows that there are many types of disturbances in the figure, especially the spectral changes on the wall caused by the reflection of the computer screen on the wall, which is very challenging. In addition, disturbances caused by computer screens and personnel are also used to test this method. Fig. 11 (b) is a video without fire, in which disturbances, including the TV, moving personnel, and the reflection caused by the light shining on the smooth surface, occur. The results show that there is no false detection for these disturbances with our method. In addition to the strategic background updating detection method, this method also further extracts features from spectrum, motion and regional expansion, which can effectively remove such disturbances and verify the robustness of our method.

5. Conclusions and future work

Indoor fire detection has always been an important and challenging problem. In recent years, with the development of computer vision technology, video-assisted fire monitoring methods have been continuously developed. Although such applications are feasible under certain conditions, complicated indoor environments easily occlude fires in the early stages. In addition, sensor-based fire detection methods are not well suited for early detection of occluded indoor fires. Motivated by these considerations, an early indoor occluded fire detection method is proposed in this paper, and our method has achieved great success in this research task. First, we analyzed the reflection characteristics of occluded fires in detail and extracted three useable characteristics. Second, our method better adapts to the slight changes caused by indoor occluded fires at an early stage. Finally, the establishment of an MES can help us to judge the early detection of occluded fires more accurately. Our method solves a special and common problem that is very important for providing early warning when indoor fires occur. In addition, the results show that our method has good effects in terms of accuracy and speed and can be applied to early indoor occluded fire detection.

Although the experimental results prove that our exploratory research is successful, the accuracy and application scenarios of occluded fire detection still need to be further improved. In future work, we will first further explore occlusion fire detection in different scenarios, such as industrial or commercial settings. Second, we will further study the characteristics of fire reflection features, such as different feature changes reflected on different object surfaces, to further extract the features that are effective in detecting occluded fires. Finally, we will attempt to introduce a neural network into occluded fire detection to improve the accuracy of the method. We hope that our research can provide support for early intelligent fire detection in the field of public safety.

Funding statement

This paper was supported by the National Natural Science Foundation of China (Grant Nos. U2034202 and 41871289), the Sichuan Youth Science and Technology Innovation Team (Grant No. 2020JDTD0003), and the China Scholarship Council.

Author contributions

Yakun Xie (Data curation; Investigation; Methodology; Project administration; Roles/Writing - original draft), Jun Zhu (Conceptualization; Funding acquisition; Project administration; Supervision; Writing - review & editing), Yukun Guo (Data curation; Software; Investigation), Jigang You (Data curation; Visualization; Methodology), Dejun Feng (Formal analysis; Supervision; Validation, Resources), Yungang Cao (Supervision; Validation; Writing - review & editing). All authors read and approved the submitted manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors wish to thank the editors and reviewers.

References

[1] S. Frizzi, R. Kaabi, M. Bouchouicha, J.M. Ginoux, E. Moreau, F. Fnaiech, Convolutional neural network for video fire and smoke detection, in: IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society, 2016, October, pp. 877–882.
[2] O. Maksymiv, T. Rak, D. Peleshko, Real-time fire detection method combining AdaBoost, LBP and convolutional neural network in video sequence, in: 2017 14th International Conference the Experience of Designing and Application of CAD Systems in Microelectronics (CADSM), 2017, February, pp. 351–353.
[3] M. Mahbub, M.M. Hossain, M.S.A. Gazi, Cloud-Enabled IoT-based embedded system and software for intelligent indoor lighting, ventilation, early stage fire detection and prevention, Comput. Network. 184 (2021) 107673.
[4] Y. Luo, L. Zhao, P. Liu, D. Huang, Fire smoke detection algorithm based on motion characteristic and convolutional neural networks, Multimed. Tool. Appl. 77 (12) (2018) 15075–15092.
[5] P. Li, Y. Yang, W. Zhao, M. Zhang, Evaluation of image fire detection algorithms based on image complexity, Fire Saf. J. 121 (2021) 103306.
[6] K. Muhammad, J. Ahmad, S.W. Baik, Early fire detection using convolutional neural networks during surveillance for effective disaster management, Neurocomputing 288 (2018) 30–42.
[7] B. Kim, J. Lee, A video-based fire detection using deep learning models, Appl. Sci. 9 (14) (2019) 2862.
[8] S.J. Wang, D.L. Jeng, M.T. Tsai, Early fire detection method in video for vessels, J. Syst. Software 82 (4) (2009) 656–667.
[9] A.K. Wong, N.K. Fong, Experimental study of video fire detection and its applications, Procedia Eng. 71 (2014) 316–327.
[10] M. Hashemzadeh, A. Zademehdi, Fire detection for video surveillance applications using ICA K-medoids-based color model and efficient spatio-temporal visual features, Expert Syst. Appl. 130 (2019) 60–78.
[11] G. Marbach, M. Loepfe, T. Brupbacher, An image processing technique for fire detection in video images, Fire Saf. J. 41 (4) (2006) 285–289.
[12] T. Celik, H. Demirel, Fire detection in video sequences using a generic color model, Fire Saf. J. 44 (2) (2009) 147–158.
[13] A. Rafiee, R. Dianat, M. Jamshidi, R. Tavakoli, S. Abbaspour, Fire and smoke detection using wavelet analysis and disorder characteristics, in: 2011 3rd International Conference on Computer Research and Development, vol. 3, 2011, pp. 262–265.
[14] Q.T. Geng, F.H. Yu, H.W. Zhao, C. Wang, New algorithm of flame detection based on color features, J. Jilin Univ. (Sci. Ed.) 44 (6) (2014) 1787–1792.
[15] T. Celik, H. Demirel, H. Ozkaramanli, M. Uyguroglu, Fire detection using statistical color model in video sequences, J. Vis. Commun. Image Represent. 18 (2) (2007) 176–185.
[16] P. Foggia, A. Saggese, M. Vento, Real-time fire detection for video-surveillance applications using a combination of experts based on color, shape, and motion, IEEE Trans. Circ. Syst. Video Technol. 25 (9) (2015) 1545–1556.
[17] N.I. Zaidi, N.A.A. Lokman, M.R. Daud, H. Achmad, K.A. Chia, Fire recognition using RGB and YCbCr color space, ARPN J. Eng. Appl. Sci. 10 (21) (2015) 9786–9790.
[18] R. Chi, Z.M. Lu, Q.G. Ji, Real-time multi-feature based fire flame detection in video, IET Image Process. 11 (1) (2017) 31–37.
[19] B.U. Töreyin, Y. Dedeoğlu, U. Güdükbay, A.E. Cetin, Computer vision based method for real-time fire and flame detection, Pattern Recogn. Lett. 27 (1) (2006) 49–58.
[20] Y.H. Habiboğlu, O. Günay, A.E. Çetin, Covariance matrix-based fire and flame detection method in video, Mach. Vis. Appl. 23 (6) (2012) 1103–1113.
[21] Q. Yan, B. Pei, J. Zhao, Forest fire image intelligent recognition based on the neural network, J. Multimed. 9 (3) (2014) 449–455.
[22] K. Dimitropoulos, P. Barmpoutis, N. Grammalidis, Spatio-temporal flame modeling and dynamic texture analysis for automatic video-based fire detection, IEEE Trans. Circ. Syst. Video Technol. 25 (2) (2014) 339–351.
[23] L.S. Shao, Y.C. Guo, Flame recognition algorithm based on Codebook in video, J. Comput. Appl. 5 (2015) 1483–1487.
[24] I. Bosch, A. Serrano, L. Vergara, Multisensor network system for wildfire detection using infrared image processing, Sci. World J. (2013) (2013).
[25] M.M. Valero, O. Rios, E. Pastor, E. Planas, Automated location of active fire perimeters in aerial infrared imaging using unsupervised edge detectors, Int. J. Wildland Fire 27 (4) (2018) 241–256.
[26] M.J. Sousa, A. Moutinho, M. Almeida, Thermal infrared sensing for near real-time data-driven fire detection and monitoring systems, Sensors 20 (23) (2020) 6803.
[27] H. Kim, C. Song, G.J. Son, S.H. Jeong, J.H. Son, Y.D. Kim, Hyperspectral image-based night-time fire detection using NKNBD, in: 2018 7th International Congress on Advanced Applied Informatics, 2018, pp. 974–975.
[28] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Adv. Neural Inf. Process. Syst. 25 (2012) 1097–1105.
[29] A. Graves, M. Liwicki, S. Fernández, R. Bertolami, H. Bunke, J. Schmidhuber, A novel connectionist system for unconstrained handwriting recognition, IEEE Trans. Pattern Anal. Mach. Intell. 31 (5) (2008) 855–868.
[30] G.E. Hinton, S. Osindero, Y.W. Teh, A fast learning algorithm for deep belief nets, Neural Comput. 18 (7) (2006) 1527–1554.
[31] T.H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, Y. Ma, PCANet: a simple deep learning baseline for image classification? IEEE Trans. Image Process. 24 (12) (2015) 5017–5032.

[32] R. Girshick, J. Donahue, T. Darrell, J. Malik, Region-based convolutional networks for accurate object detection and segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 38 (1) (2015) 142–158.
[33] B. Jiang, J. Yang, Z. Lv, K. Tian, Q. Meng, Y. Yan, Internet cross-media retrieval based on deep learning, J. Vis. Commun. Image Represent. 48 (2017) 356–366.
[34] T. Mikolov, K. Chen, G. Corrado, J. Dean, Efficient Estimation of Word Representations in Vector Space, 2013 arXiv preprint arXiv:1301.3781.
[35] G.E. Dahl, D. Yu, L. Deng, A. Acero, Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition, IEEE Trans. Audio Speech Lang. Process. 20 (1) (2011) 30–42.
[36] D. Erhan, C. Szegedy, A. Toshev, D. Anguelov, Scalable object detection using deep neural networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2147–2154.
[37] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, A.W. M. Jeroen, B. Ginneken, C.I. Sánchez, A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017) 60–88.
[38] J.Y. Jang, K.W. Lee, Y.J. Kim, W.T. Kim, S-FDS: a smart fire detection system based on the integration of fuzzy logic and deep learning, J. Ins. Electr. Inf. Eng. 54 (4) (2017) 50–58.
[39] C. Hu, P. Tang, W. Jin, Z. He, W. Li, Real-time fire detection based on deep convolutional long-recurrent networks and optical flow method, in: 2018 37th Chinese Control Conference, CCC), 2018, pp. 9061–9066.
[40] S. Aslan, U. Güdükbay, B.U. Töreyin, A.E. Cetin, Deep Convolutional Generative Adversarial Networks Based Flame Detection in Video, 2019 arXiv preprint arXiv: 1902.01824.
[41] Y. Xie, J. Zhu, Y. Cao, Y. Zhang, D. Feng, Y. Zhang, M. Chen, Efficient video fire detection exploiting motion-flicker-based dynamic features and deep static features, IEEE Access 8 (2020) 81904–81917.
[42] R. Xu, H. Lin, K. Lu, L. Cao, Y. Liu, A forest fire detection system based on ensemble learning, Forests 12 (2) (2021) 217.
[43] T. Liu, H.H. Wang, Y.L. Xiang, P.Y. Lu, An approach of real-time vehicle detection based on improved Adaboost and frame differencing relu, J. Huazhong Univ. Sci. Technol. (Nat. Sci. Ed.) 41 (2013) 379–382.
[44] P.W. Power, J.A. Schoonees, Understanding background mixture models for foreground segmentation, Proc. Image Vision Comput. New Zealand 11 (2002).
[45] Z. Zivkovic, F. Van Der Heijden, Efficient adaptive density estimation per image pixel for the task of background subtraction, Pattern Recogn. Lett. 27 (7) (2006) 773–780.
[46] J. Chen, Y. He, J. Wang, Multi-feature fusion based fast video flame detection, Build. Environ. 45 (5) (2010) 1113–1122.
[47] D.Y. Chino, L.P. Avalhais, J.F. Rodrigues, A.J. Traina, Bowfire: detection of fire in still images by integrating pixel color and texture analysis, in: 2015 28th SIBGRAPI Conference on Graphics, Patterns and Images, 2015, pp. 95–102.
[48] B.C. Ko, S.J. Ham, J.Y. Nam, Modeling and formalization of fuzzy finite automata for detection of irregular fire flames, IEEE Trans. Circ. Syst. Video Technol. 21 (12) (2011) 1903–1912.
[49] V. Hüttner, C.R. Steffens, S.S. da Costa Botelho, First response fire combat: deep leaning based visible fire detection, in: 2017 Latin American Robotics Symposium (LARS) and 2017 Brazilian Symposium on Robotics (SBR), 2017, pp. 1–6.
[50] A. Chenebert, T.P. Breckon, A. Gaszczak, A non-temporal texture driven approach to real-time fire detection, 18th IEEE International Conference on Image Processing (2011) 1741–1744.
[51] T. Wang, L. Bu, Z. Yang, P. Yuan, J. Ouyang, A new fire detection method using a multi-expert system based on color dispersion, similarity and centroid motion in indoor environment, IEEE/CAA Journal of Automatica Sinica 7 (1) (2019) 263–275.