Image and Vision Computing 91 (2019) 103803

Contents lists available at ScienceDirect

Image and Vision Computing

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

Intelligent and vision-based fire detection systems: A survey☆

Fengju Bu [a], Mohammad Samadi Gharajeh [b,*]

[a] Shandong Agriculture and Engineering University, Jinan, Shandong, China
[b] Young Researchers and Elite Club, Tabriz Branch, Islamic Azad University, Tabriz, Iran

ARTICLE INFO

Article history:
Received 10 May 2019
Received in revised form 14 August 2019
Accepted 16 August 2019
Available online 23 August 2019

Keywords:
Fire detection system
Computer vision
Artificial intelligence
Intelligent system
Vision-based system

ABSTRACT

Fire is one of the main disasters in the world. A fire detection system should detect fires in various environments (e.g., buildings, forests, and rural areas) in the shortest time in order to reduce financial losses and humanistic disasters. Fire sensors are, in fact, complementary to conventional point sensors (e.g., smoke and heat detectors), which provide people the early warnings of fire occurrences. Cameras combined with image processing techniques detect fire occurrences more quickly than point sensors. Moreover, they provide the size, growth, and direction of fires more easily than their conventional detectors. This paper, initially, presents a glance view on the main features of various environments including buildings, forests, and mines that should be considered in the design of fire detection systems. Afterwards, it describes some of the intelligent and vision-based fire detection systems that have been presented by researchers in the last decade. These systems are categorized, in this paper, into two groups: intelligent detection systems for forest fires and intelligent fire detection systems for all of the environments. They use various intelligent techniques (e.g., convolutional neural networks, color models, and fuzzy logic) to detect fire occurrences with a high accuracy in various environments. Performances of the fire detection systems are compared to each other in terms of detection rate, precision, true-positive rate, false-positive rate, etc. under different evaluation scenarios.

1. Introduction

Fire causes much more problems to human safety, health, and property. Financial and humanistic losses in fires have been considerably increased over the past years. They are happened by some of the unpredicted conditions such as changing fuel management policies, climate change, and high development in rural areas [1]. Fire detection system is an important component in surveillance systems. It monitors various environments (e.g., buildings) as main unit of an early warning framework to report preferably the start of fire. Most of the existing fire detection systems use point sensor regarding the particle sampling, temperature sampling, and smoke analysis. They measure fire probabilities slowly, usually in minutes, and also have little applicability in outdoor environments. Upon a sufficient amount of smoke particles flows into the device or temperature is increased considerably, the smoke and heat detectors trigger to avoid false alarms. Besides, time is one of the main factors to minimize fire damages. Thus, the response time of detection systems should be decreased to increase the chances to extinguish fires and, thereby, reduce the financial and humanistic damages [2].

Researchers have attempted to achieve the slow response time of point detectors in the last decades. This issue can be observed in the alternative, fire sensor based computing for automatic fire detection systems. They have summarized that one of the latest efficient techniques to detect smoke and flame is to achieve images and/or video frames from vision sensors [3]. Vision sensor based techniques can lead response time to be decreased, probability of the detection process to be enhanced, and coverage of large areas (e.g., open environments) to be provided considerably. Vision sensors can give more information about the direction, growth, and size of fire and smoke [4]. In addition, we can upgrade the existing surveillance systems installed in various environments (e.g., factories and public sites) with low implementation costs. This process can be done to provide early warnings of fires and, also, detect flame and smoke by using video processing and machine vision techniques.

Fire detection systems, implemented by vision sensors, are generally composed of three units: motion detection, candidate object segmentation, and flame or smoke blob detection. In these systems, cameras are usually considered to be fixed but initial motion is detected based on some variations on image subtraction and background modeling techniques [5]. Motion detection process can be followed by candidate object segmentation using color information [6,7], pixel variability measures [8], texture cues [9], or optical flow field distribution [10]. Flame and smoke blobs can be identified by image separation

☆ This paper has been recommended for acceptance by Sinisa Todorovic.
* Corresponding author.
E-mail addresses: m.samadi@iaut.ac.ir, mhm.samadi@gmail.com (M.S. Gharajeh).

https://doi.org/10.1016/j.imavis.2019.08.007
0262-8856/

approaches [11,12] or by modeling the flame or smoke in mathematical terms [13–15]. After motion detection, geographical regions of fire can be identified by using morphological image processing methods as a preprocessing step. Ability of a fire system for detecting the fire and smoke depends on the viewpoint of camera field and the depth of specific scene. Moreover, smoke and flame have different dynamic behaviors and physical properties. Therefore, a fire system that locates the regions of both smoke and flame can considerably detect a fire earlier than a system that only locates the region of smoke or the region of flame [16].

Various intelligent and vision-based fire detection systems have been presented by researchers, especially in the last decade. They use some of the methods and systems such as cameras (e.g., optical cameras [17]), intelligent techniques (e.g., neural networks [18–20], particle swarm optimization [21], and fuzzy logic [22]), wireless networks (e.g., wireless sensor networks and satellite systems [17]), robotic systems (e.g., unmanned aerial systems [23,24]), and image processing techniques (e.g., RGB color [25]). Main features, advantages, and/or drawbacks of some of the latest intelligent and vision-based fire systems are described in this paper. These fire systems are designed and implemented for detecting the fire occurrences in different environments (e.g., forests [26]). It seems that the use of intelligent and vision-based techniques have had noticeable effects on the performance of fire detection systems. The most fire detection systems are organized based on various features of the fire (e.g., flame, smoke, and color) using the intelligent techniques (e.g., convolutional neural networks). That is, there are many common features between these systems. Therefore, they are categorized in this paper based on types of the environments including detection systems for forest fires and fire detection systems for all of the indoor and outdoor environments.

The rest of paper is organized as follows. Section 2 presents types of the environments that should be considered by fire detection systems. Section 3 describes some of the intelligent and vision-based fire detection systems that use various methods (e.g., convolutional neural networks) in different environments. Section 4 evaluates the simulation and/or experimental results of the latest detection systems according to different parameters such as detection rate, true-positive rate, and false alarm rate. Section 5 discusses about benchmark datasets of the intelligent fire detection systems. Finally, Section 6 concludes the paper.

2. Types of the environments

There are various types of the environments that may be destroyed by fires. Fire systems should be designed and implemented for such environments to prevent terrible loss of lives and, also, survive valuable natural and individual properties. This section presents some physical features of the three important environments including buildings, forests, and mines.

Wildland fires can destroy the exposure of building components and systems. Structures may ignite as wildland fire spreads toward the community and urban areas based on the exposure conditions. Thus, the ignition can be fully prevented or considerably reduced by understanding the mitigation strategies and the ignition potential of structures. Community design, planning practices, active/passive suppression mechanisms, and engineering-based methods are some of the mitigation strategies that can reduce the financial and property losses of fire occurrences. Hakes et al. [27] have represented the typical building components and systems that can be considered as potential pathways to ignition in urban areas. They have discussed about a structure with the main components and systems including roof, valley, dormer, gutter, eave, gable, vent, siding, deck, fence, and mulch. When a structure is revealed to radiation, firebrands, and flame contact, some components on the structure would directly begin to flame or firstly smolder and then transition to the flaming ignition. Since wood shake and shingle roofs ignite easily with many crevices and a large surface area of the flammable material, they often cause an increased fire risk [28]. The

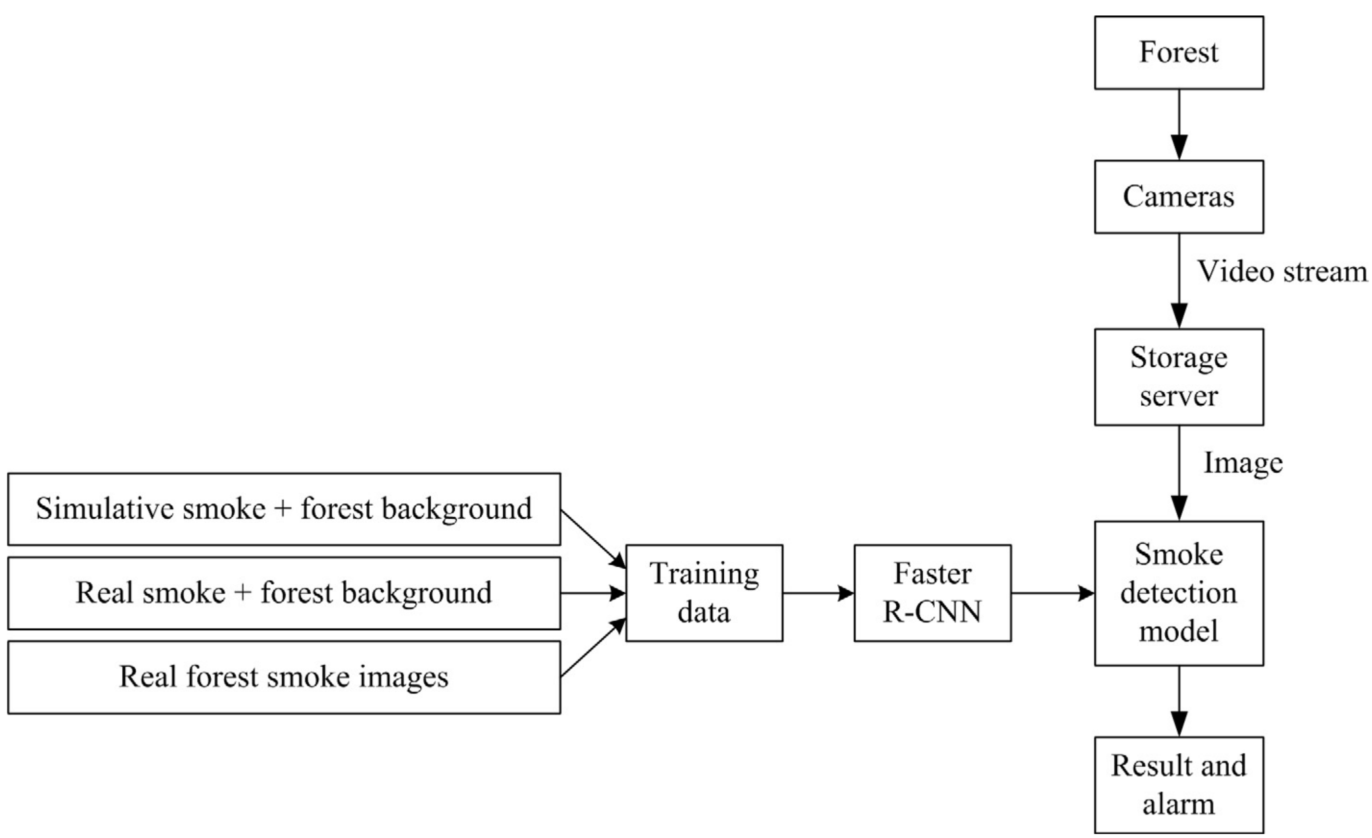

Fig. 1. Overall view of a wildland forest fire smoke detection using faster R-CNN [20].

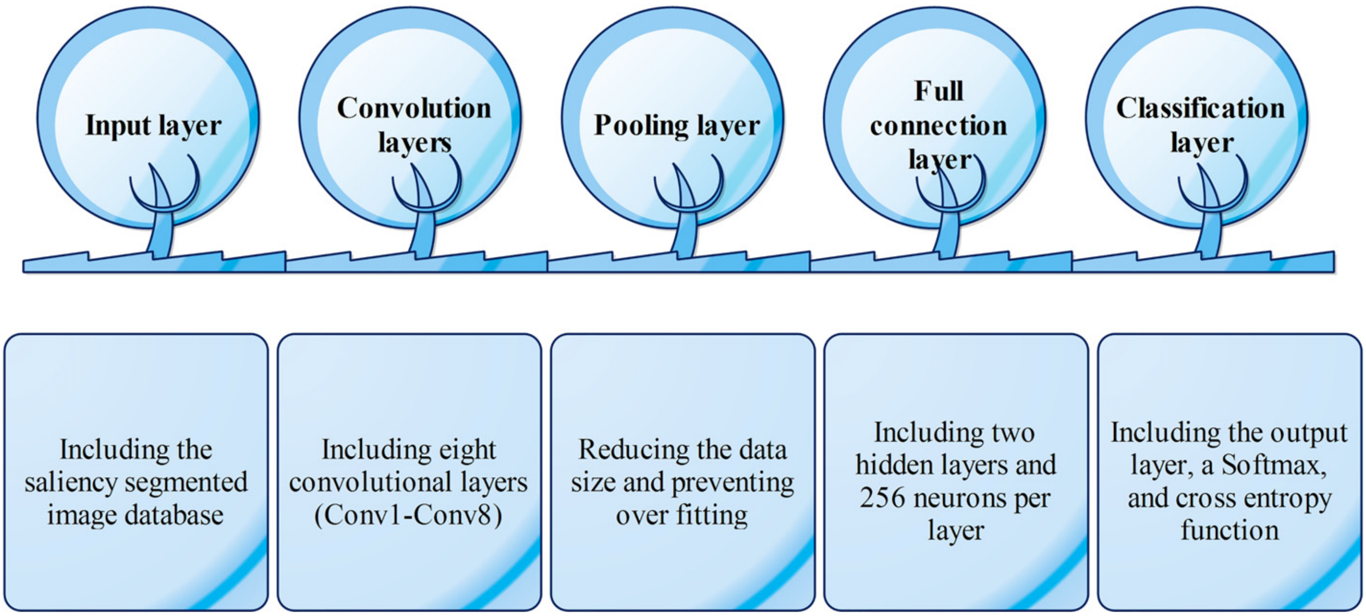

Fig. 2. Overall elements of a saliency detection and deep learning-based wildfire identification [24].

debris utilized in gutter may be ignited by firebrands, especially at roof-gutter intersection. Thus, gutters are one of the potential pathways to ignition of a home. Eaves and vents are identified as possible sources of ignition at homes in urban areas. Besides, vents cause to open a way toward the burning brands and penetrate the interior of a structure [29]. Because of radiant heat exposure or direct flame contact, siding materials have been often considered to ignite on fires. Firebrand accumulation can ignite the nearby vegetation or other fuels (e.g., wood piles and mulch) without proper clearance around the foundation of a structure. It can cause radiant heat exposure and direct flame contact in the siding materials on exterior walls. Decks, fences, and mulches are the remainder materials of a structure that may cause the most potential sources of ignition [30–32].

Forests protect the earth's ecological balance, in the most aspects. When a forest fire has spread over a large area, the fire will be notified accordingly. Thus, it is not possible to control and stop it at times. Forest fires cause a big loss and an irrevocable damage to environment and atmosphere plus an irrevocable damage to the ecology. Forest fire not only causes the tragic loss of lives, valuable natural environments, and individual properties (including hundreds of houses and thousands of hectares of forest), but also it is a big danger to ecological health and protection of environment. Researches show that forest fires increase 30% of carbon dioxide (CO_2) and huge amounts of smoke in atmosphere and ecology [33]. Furthermore, forest fires cause terrible consequences and long-term disastrous effects such as global warming, negative effect on local weather patterns, and extinction of rare species of the fauna and flora [17]. Sahin [34] has presented a sample infrastructure of the forest fire detection system, without any intelligent and vision-based technique. This system includes various elements, especially communication channels, a central classifier, and mobile biological sensors. Communication channels are provided by a satellite based system and several access points to cover all locations of the forest environment. The central classifier is used to categorize readings of the mobile biological sensors via access points and is also implemented by using a decision support system. Mobile biological sensors are the most important parts of this system, which transmit any changes of the temperature or radiation level and thereby send geographical location of the animals to access points.

Since underground coal mines contain coal dust, methane, and other toxic gases [35], they have been considered as a very hazardous and dangerous environment [36]. A consequence of coal dust or methane gas explosions is the main reason to cause approximately 33.8% of deaths in the mining sector. Recent reports represent that gas accumulation in the coal mines of salt-range region has caused approximately 38% of underground mine accidents [37]. There exist various toxic gases including methane gas, carbon monoxide (CO), CO_2, and hydrogen sulfide (H_2S) in underground coal mines that can critically lead to damage human body [38]. Consequently, an efficient monitoring system for underground mine environments can mostly guarantee the safety of miners and mine property. Jo and Khan [39] have presented an early-warning safety system for the Hassan Kishore coal mine that uses the Internet of Things (IoT) instead of using any low-cost, intelligent, vision-based system. This system uses the stationary and mobile nodes that are installed in various locations to detect phenomena data. Moreover, a router node is dedicated for every section to gather and then transmit readings data toward a base station via gateway node.

3. Fire detection systems using the intelligent and vision-based methods

Fire detection systems can be organized to use in various environments. Since every environment has some special specifications, some of the fire detections can be used in forest fires, some ones can be applied in outdoor environments, and some others can be utilized in all of the environments. Accordingly, this section describes the intelligent and vision-based fire detection systems into two groups: forest fires and all of the environments.

3.1. Intelligent detection systems for forest fires

Convolutional neural network (CNN) is considered as a high-performance technique in image classification and vision-based systems [40]. It can be used to improve detection accuracy of fire notification systems, minimize fire disasters, and reduce the social and ecological results. Zhang et al. [41] have presented a deep learning-based detection system for forest fires. They have trained this system by using a full-image and fine-grained patch fire classifier that are associated with a deep CNN. The system uses two scenarios: (i) learning the binary classifier by fire patches from scratch and (ii) learning a full-image fire classifier and then applying the fine-grained patch classifier. The authors have collected a fire detection dataset from the online

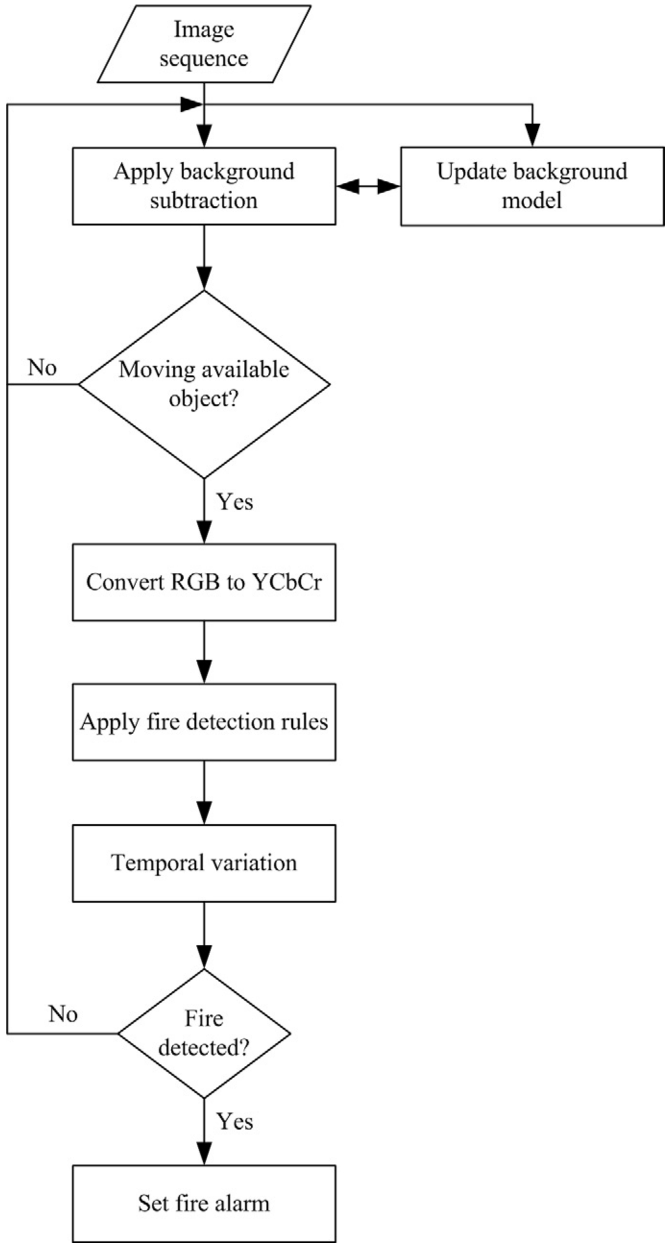

Fig. 3. Flowchart of a forest fire detection by rule-based image processing and temporal variation [26].

resources [42] that have been applied in the latest fire detection literatures. This dataset includes 25 videos obtained from forest environments. It contains 21 positive (fire) sequences and 4 negative (non-fire) sequences. Most of these literatures report image-level evaluation while some of them also report patch-level detection accuracy. The authors have used the following equations to detect a fire occurrence:

$$Accuracy = \frac{NTP + NTN}{POS + NEG}; Detection\ rate = \frac{NTP}{POS}; False\ alarm\ rate = \frac{NFN}{NEG} \tag{1}$$

where NTP is the number of truth positives, NTN is the number of truth negatives, NFN is the number of false negatives, POS is the number of positives, and NEG is the number of negatives. Support vector machine (SVM) is a machine learning algorithm, which evaluates data for the classification and regression processes. The authors have also used the SVM algorithm to evaluate performance of this detection system.

Zhang et al. [20] have presented a wildland forest fire smoke detection system that works by using faster region-based CNN (R-CNN) and synthetic smoke images. The authors have applied two approaches for synthesizing a large number of forest smoke images to prevent the process of complex feature images in traditional video smoke detection techniques. In the first approach, real smoke is extracted from green background to insert into the forest background. In the second one, synthetic smoke is generated by the rendering software to insert into the forest background, as before. In this system, input of the whole network is an original image without using any preprocessing or block segmentation. Thus, processing speed has been faster than the prior two generations. Fig. 1 shows main elements and workflow of this system. By using faster R-CNN in detecting the smoke of real scene, real forest smoke images can improve the affluence of training data. The system focuses on difference between the real smoke and synthetic smoke that are used for training process. The authors have trained smoke detection models on both of the simulative smoke + forest background dataset, namely SF dataset, and real smoke + forest background dataset, namely RF dataset.

Unmanned aerial vehicle (UAV) is a machine that is capable to fly by aerodynamic lift and guided without an onboard crew. It may be recoverable or expendable to fly semi-autonomously or autonomously. One of the important applications of this vehicle has been historically identified in the fields of surveillance and inspection [44]. Thus, we can utilize various types of aerials vehicle to collect image-based information over various areas (e.g., forests). Cruz et al. [45] have presented a forest fire detection approach for application in unmanned aerial systems using a new color index. They have focused on detection of the region of interests in which the cases are flames and smoke regarding the calculation of color indices. This system includes a transformation unit to convert images from Red, Green and Blue (RGB) three-band space to a gray-scale band. This unit uses some of the arithmetic operations between image components to enhance the desired colors. Afterwards, a thresholding process binarizes the result to separate the region of interest from the rest. Color indices are frequently collected as vegetation indices, which are used in a machine vision-based process for agricultural automatization, especially for detection of the crops and weeds. The authors have attempted to make a relationship between effective detection process and minimum execution times. Thus, they have used extra data gathered from the images (e.g., the size of fire, extension velocity, and coordinates) to perform a rapid reaction in forest fires. Generally, this work uses several units including normalization, relation between color components, fire detection index, and forest fire detection index. In the normalization unit, all of the components are normalized as follows to obtain a better robustness against different lighting conditions:

$$r = \frac{R}{S}, g = \frac{G}{S}, b = \frac{B}{S}; where\ S = R + G + B \tag{2}$$

Relation between components of the regions of interest is evaluated carefully in order to improve the desired color tones and find appropriate equations for the arithmetic operations. The colors or interest, which represents a combustion process, consists of the high red channel values rather than blue and green. Fire detection index is calculated according to the normalized colors as the following:

$$Fire\ detection\ index = 2r{-}g{-}b \tag{3}$$

The above equation can be used when there is any flame in the image. Finally, forest fire detection index is used to give more flexibility and importance to the colors present in fire and smoke. It is calculated

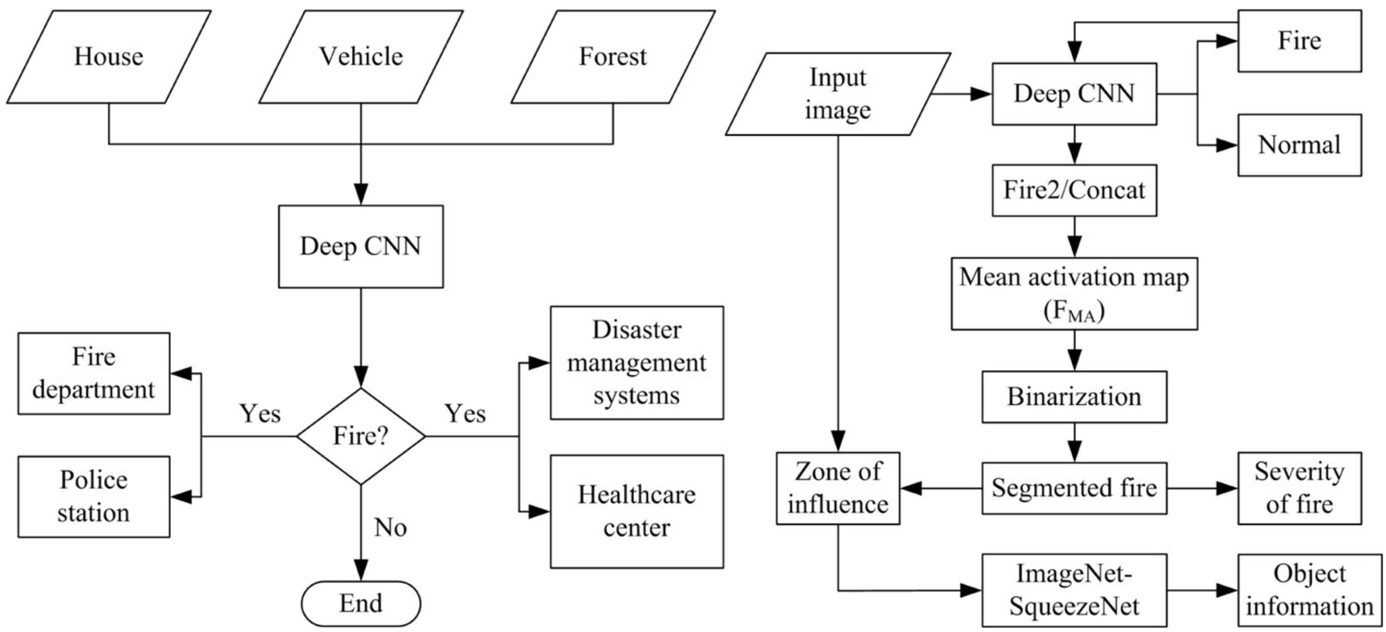

Fig. 4. General characteristics of a fire detection system using the deep CNN [19]. (a) Overview of the system; (b) fire localization process.

by using the normalized colors and a regulative weighting factor (ρ) as follows:

$$\textit{Forest fire detection index} = r(2\rho+1)-g(\rho+2)+b(1-\rho) \quad (4)$$

Different values of ρ determine various results for this index to identify a high amount of red, yellow, brown, and some other color tonalities. Afterwards, the obtained images will be binarized and also the segmented regions will be labeled by the detection process.

Yuan et al. [23] have presented a forest fire detection method that uses both color and motion features with the aid of unmanned aerial vehicles. In the first step, fire-colored pixels are extracted as fire candidate regions by using a color decision rule and the chromatic feature of fire. Afterwards, motion vectors of the candidate regions are calculated by the Horn and Schunck optical flow algorithm. Optical flow results can also predict motion feature for distinguishing the fires from other fire analogs. Binary images are determined by performing the thresholding process and morphological operations on motion vectors. Finally, geographical locations of the fires are identified in each binary image using a blob counter approach.

Zhao et al. [24] have presented a saliency detection and deep learning-based wildfire identification system using the image frames captured by UAV. They have offered a detection algorithm for the fast location and segmentation of core fire areas in aerial images. This system can considerably prevent feature loss that is caused by direct resizing. Moreover, it is applied in the formation and data augmentation of a standard fire image dataset using a 15-layered self-learning deep CNN architecture. This process causes to produce a self-learning fire feature exactor and classifier system. The authors have considered some of the non-negligible challenges, especially localization of the fire area, fixed training image size, and limited amount of aerial view wildfire images. The wildfire localization and segmentation algorithm consists of saliency detection and logistic regression classifier through two steps: region of interest proposal and region selection using machine learning. In the first step, the region of interest is extracted by the saliency detection method. Accordingly, the color and texture features of the region of interests are calculated to specify possible fire regions. In the second step, the system uses two logistic regression classifiers to identify feature vector inside the region of interest. If this vector belongs to flame or smoke then these regions will be segmented by the system. Fig. 2 illustrates overall elements of the deep CNN model that is used in this system. It consists of five layers including input layer, convolution layers, pooling layer, full connection layer, and classification layer. The wildfire features from low to high level can be extracted by this self-learning structure. Input layer feeds input data into the model, which is the saliency segmented image database having the color of three randomized RGB channels. Convolutional layers include eight convolutional layers to extract fire features from low to high level. Pooling layer captures deformable parts, reduces dimension of the convolutional output, and avoids over fitting. Full connection layer contains two hidden layers, 256 neurons per layer, and a weight initialization. It reduces the dependency between neurons and, consequently, reduces the network's overfitting. Classification layer includes output layer of the full connection neural network, a Softmax, and cross entropy function. It concludes the prior abstracted high-level features for recognition of a certain image. The network is trained until the training error is obtained equal to a negligible value.

Mahmoud and Ren [26] have presented a forest fire detection system that applies a rule-based image processing technique and temporal variation. This system uses background subtraction to follow the movements containing region detection. It converts the segmented moving regions from RGB to YCbCr color space and, also, utilizes five fire detection rules for separating the candidate fire pixels. It is worth noting that the system uses temporal variation to make differences between the fire and fire-color objects. Fig. 3 depicts main steps and workflow of this system. Background subtraction and background model are used in the system based on the input video gathered from input device. Then, the input image sequence is converted from RGB to YCbCr color space, as well as fire detection rules and temporal variation are used accordingly. If all of the detection conditions are satisfied then a fire alarm will be activated.

Particle swarm optimization [58,59] is presented to offer very simple concepts and paradigms that can be developed in few lines of the computer codes. It only needs primitive mathematical operators and is computationally inexpensive in terms of both memory requirements and processing speed. Khatami et al. [21] have presented a fire flame detection system using image processing techniques based on particle swarm optimization and K-medoids clustering. This system applies a conversion matrix to describe a new fire flame-based color space on the basis of color features. It is simply presented in terms of feature extraction and color space conversion in which color elements of every pixel are directly extracted to being applied without any extra analysis. The

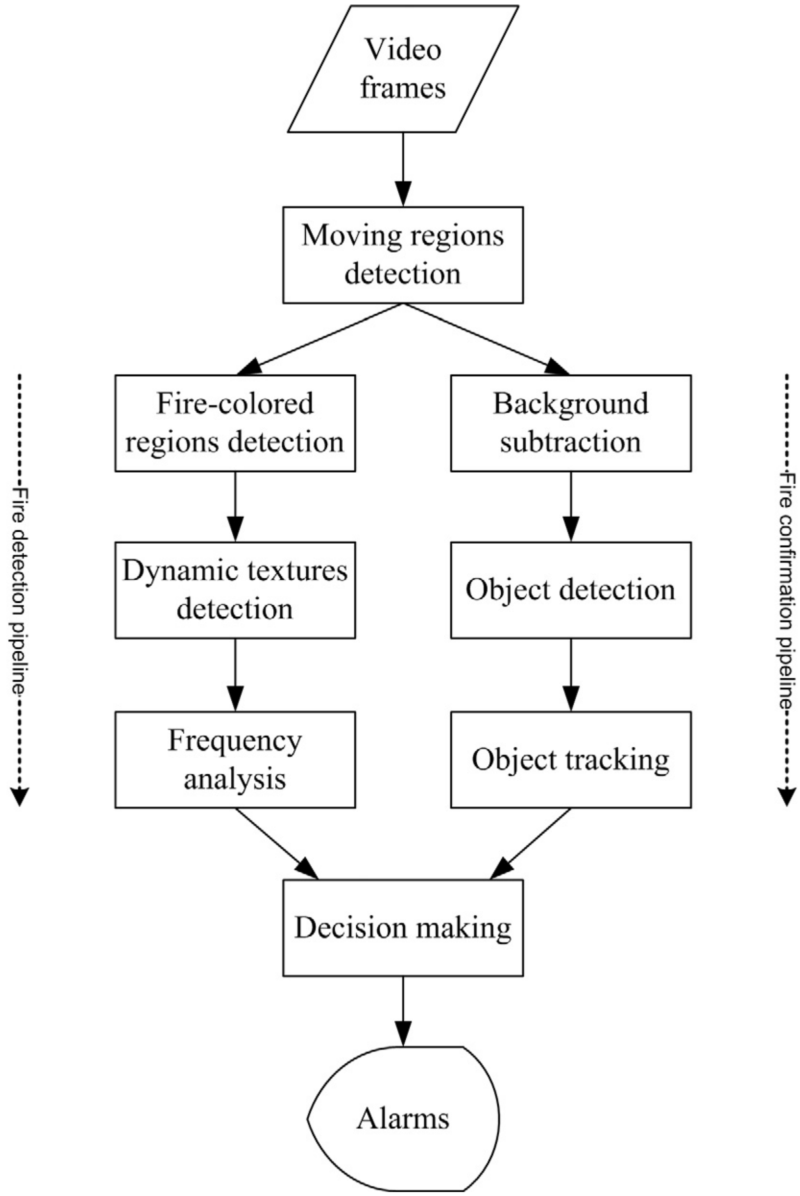

Fig. 5. Processing pipeline of a vision-based approach for fire detection [51].

authors have developed a color space based on fire flames, linear multiplication of a conversion matrix, and color properties of a sample image. The matrix multiplication generates a differentiating color space to highlight the fire part and dim the non-fire part. Weights of the color-differentiating conversion matrix are calculated by particle swarm optimization and sample pixels of an image. Furthermore, detection process uses the conversion matrix that is achieved for different fire images to implement a fast and easy fire detection system.

3.2. Intelligent fire detection systems for all of the environments

Muhammad et al. [19] have presented a fire detection system by using deep CNN and video information. They have attempted to minimize time-consuming feature of the conventional hand-crafted features for fire detection systems. Furthermore, they have designed an early fire detection system in closed-circuit television (CCTV) surveillance networks for both indoor and outdoor environments based on deep learning techniques. This system can increase fire detection accuracy, reduce the number of false alarms, and minimize fire damages. It is designed by using the AlexNet architecture [43] for fire detection process and a transfer learning strategy. The authors have offered an intelligent feature map selection algorithm sensitive to fire regions in order to select appropriate feature maps from among the convolutional layers of trained CNN. These feature maps can lead to perform a more accurate segmentation to determine essential properties of the fire such as its growth rate, the severity of fire, and/or its burning degree. Besides, they have applied a pre-trained model on 1000 classes of objects in the ImageNet dataset to specify whether the fire is in a house, a car, a forest, or any other place. This model can aid firefighters to prioritize regions with the strongest fire. Fig. 4 shows overview and fire localization process of this system. Fig. 4(a) illustrates overview of the system for fire detection process. Cameras are installed in various environments such as house, vehicle, and forest to gather environmental data. Then, a deep CNN controller analyzes the gathered images to determine whether a fire happens in the environment or not. If any fire happens, fire department, police station, disaster management systems, and healthcare center will be notified accordingly. Fig. 4

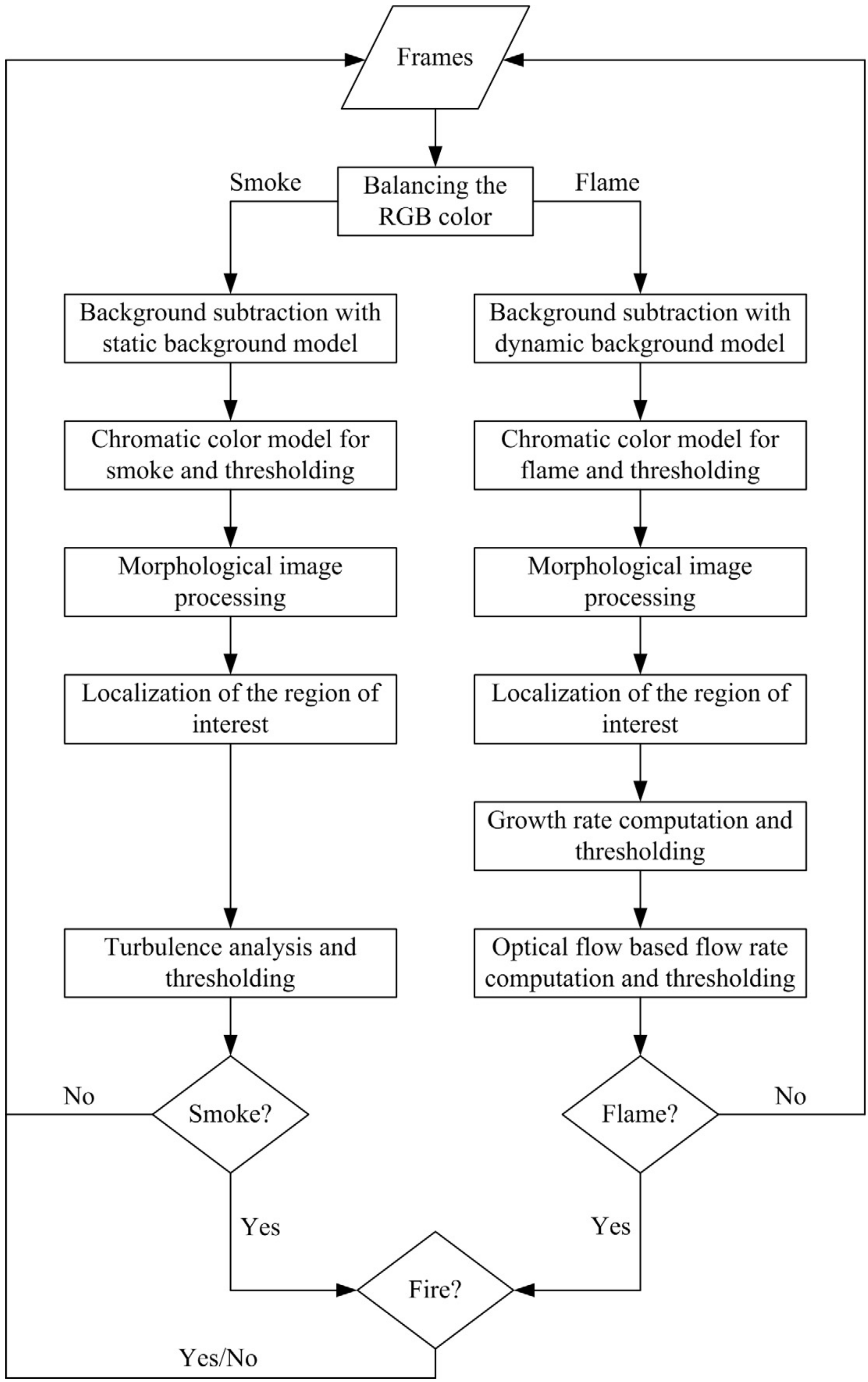

Fig. 6. Workflow of an early fire detection system using a combined video processing strategy [2].

Table 1
Evaluation results of the different patch classifiers using SVM-Raw, CNN-Raw, SVM-Pool5, and CNN-Pool5 [41].

Train type	Technique	Evaluation parameters		
		Accuracy	Detection rate	False alarm rate
Train set 1	SVM-Raw	92.2%	56.2%	3.6%
	CNN-Raw	93.1%	84.5%	3.9%
	SVM-Pool5	95.6%	76.2%	2.1%
	CNN-Pool5	97.3%	84.8%	1.2%
Train set 2	SVM-Raw	74%	23.9%	18.5%
	CNN-Raw	88.6%	59.6%	6.5%
	SVM-Pool5	89%	40.1%	3.7%
	CNN-Pool5	90.1%	39.2%	2.3%

(b) depicts fire localization process of the deep CNN technique. The deep CNN determines that a fire occurrence happens or environmental condition is normal according to the input image. Besides, the system makes the segmented fire by mean activation map and binarization to determine various features such as severity of fire and object information.

Analysis of the image and video frames is another technique in fire detection systems, which has been already studied by many researchers. There are some of the video-based flame and fire detection algorithms in the literature that apply visual information such as color, motion, and geometry of fire regions. These algorithms use various techniques such as color clues for flame detection, pixel colors and their temporal variations, and temporal object boundary pixels [4,7,46,47]. Celik et al. [48] have presented a real-time fire detector that applies both of color information and registered background scene. Statistical measurement of the sample images can determine color information of the fires. Three Gaussian distributions are used to model simple adaptive information of the scene. They make a model for the pixel values related to colored information in all of the color channels. The authors have developed the fire detection system by combining the foreground objects with color statistics and also by analysis of the output via consecutive frames. This system can detect fire occurrences in the earliest time, except in the explosive conditions where smoke is detected before the fire is started.

Celik and Demirel [13] have offered a rule-based generic color model to perform a detection process based on a flame pixel classification. This model uses YCbCr color space to separate the luminance from the chrominance. This mechanism can work more effectively than the other color spaces (e.g., RGB). It aids to make a generic chrominance model for an appropriate flame pixel classification. The rules, developed in RGB and the normalized RGB, are translated to YCbCr color space. Besides, new rules are considered to being in YCbCr color space in order to reduce harmful effects of the changing illumination and improve efficiency of the detection process. If the illumination of image is changed, fire pixel classification rules in RGB color space will not work carefully. Besides, separating the pixel's values of intensity and chrominance will not be possible. Thus, the authors have used the chrominance mechanism to make model of the color of fire more effectively than model of the intensity. Since there is a linear conversion process between RGB and YCbCr color spaces, YCbCr color space is applied to make an appropriate model of the fire pixels. The following equation formulates conversion process from RGB to YCbCr color spaces:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0.2568 \ 0.5041 \ 0.0979, \\ -0.1482 -0.2910 \ 0.4392 , \ 0.4392 \quad -0.3678 \quad -0.0714 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{5}$$

Borges and Izquierdo [50] have presented a detection system for fire occurrences in videos based on color information. They have considered visual features of the fire such as area size, color, surface coarseness, skewness, and boundary roughness to make the detection system useful for the most fire occurrences. For example, the skewness is a very applicable feature because it contains frequent occurrence of the saturation in red channel of fire zones. This system can be used for both surveillance and automatic video classification to perform retrieval process of the fire catastrophes in databases of newscast content. It evaluates frame-to-frame changes of the special low-level features to identify potential fire zones. In fact, the system considers modifications of the motion-based features and a model for probability of fire occurrences to determine location of the candidate fire zones.

Gomes et al. [51] have presented a vision-based approach for fire detection using the fixed surveillance smart cameras. They have used background subtraction and context-based learning to enhance the accuracy and robustness of fire detection process. Subsequently, a color-based model of fire's appearance and a wavelet-based model of fire's frequency signature are developed to make a fast discrimination between fire-colored moving objects and fire regions. The category and behavior of moving objects are used in decision-making process to decrease the number of false alarms in presence of the fire-colored moving objects. Besides, the camera-world mapping is predicted by using a GPS-based calibration system in order to estimate the expected size of objects and generate geo-referenced alarms. Fig. 5 shows flowchart of this approach that is offered as a processing pipeline. Workflow is started by detecting the moving regions based on input images. A dynamic threshold is employed to manage the motion detection through largeness feature of each pixel's intensity variation across three consecutive frames. Finally, decision-making process is conducted based on the results of frequency analysis and object tracking to show any alarm required by the system.

Kim et al. [52] have offered a fire detection algorithm using a color model in video sequences on wireless sensor network. This algorithm is developed by a background subtraction for foregrounds detection and a Gaussian mixture model for background modeling. Since fire has various features, the authors have applied color information to achieve useful data for the gathered video frames. They have presented the color detection algorithm for RGB color space based on area growth of the objects in consecutive. All of the fire-like colored objects, which differ from background and change in frame-by-frame boundary box, are constructed by temporal variation for each pixel. It is worth noting

Table 2
Comparison results of the fire detection system using a deep CNN [19].

Technique		Dataset1			Dataset2		
		False positives	False negatives	Accuracy	Precision	Recall	F-measure
Deep CNN [19]	After fine-tuned	8.87%	2.12%	94.50%	N/A	N/A	N/A
		N/A	N/A	N/A	0.86	0.97	0.91
	Before fine-tuned	9.99%	10.39%	89.8%	N/A	N/A	N/A
		N/A	N/A	N/A	0.84	0.87	0.85
AlexNet [43]	After fine-tuned	9.07%	2.13%	94.39%	N/A	N/A	N/A
		N/A	N/A	N/A	0.82	0.98	0.89
	Before fine-tuned	9.22%	10.65%	90.06%	N/A	N/A	N/A
		N/A	N/A	N/A	0.85	0.92	0.88

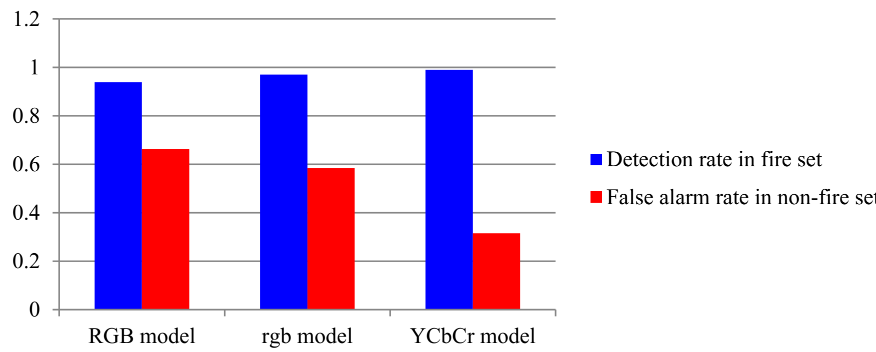

Fig. 7. Performance evaluation of the fire detection system using a video-based and generic color model [13].

that video sequences are acquired by static camera to offer an appropriate surveillance visual system based on CCTV cameras.

Dimitropoulos et al. [53] have presented a flame detection method based on combination of the features extracted from dynamic texture analysis and spatio-temporal flame modeling. This method focuses on both the identification of specific fire features (e.g., color and motion) and the powerful properties of linear dynamical systems in order to evaluate temporal evolution of the pixels' intensities and enhance robustness rate of the detection algorithm. Moreover, fire regions into a frame are identified using color analysis and background subtraction

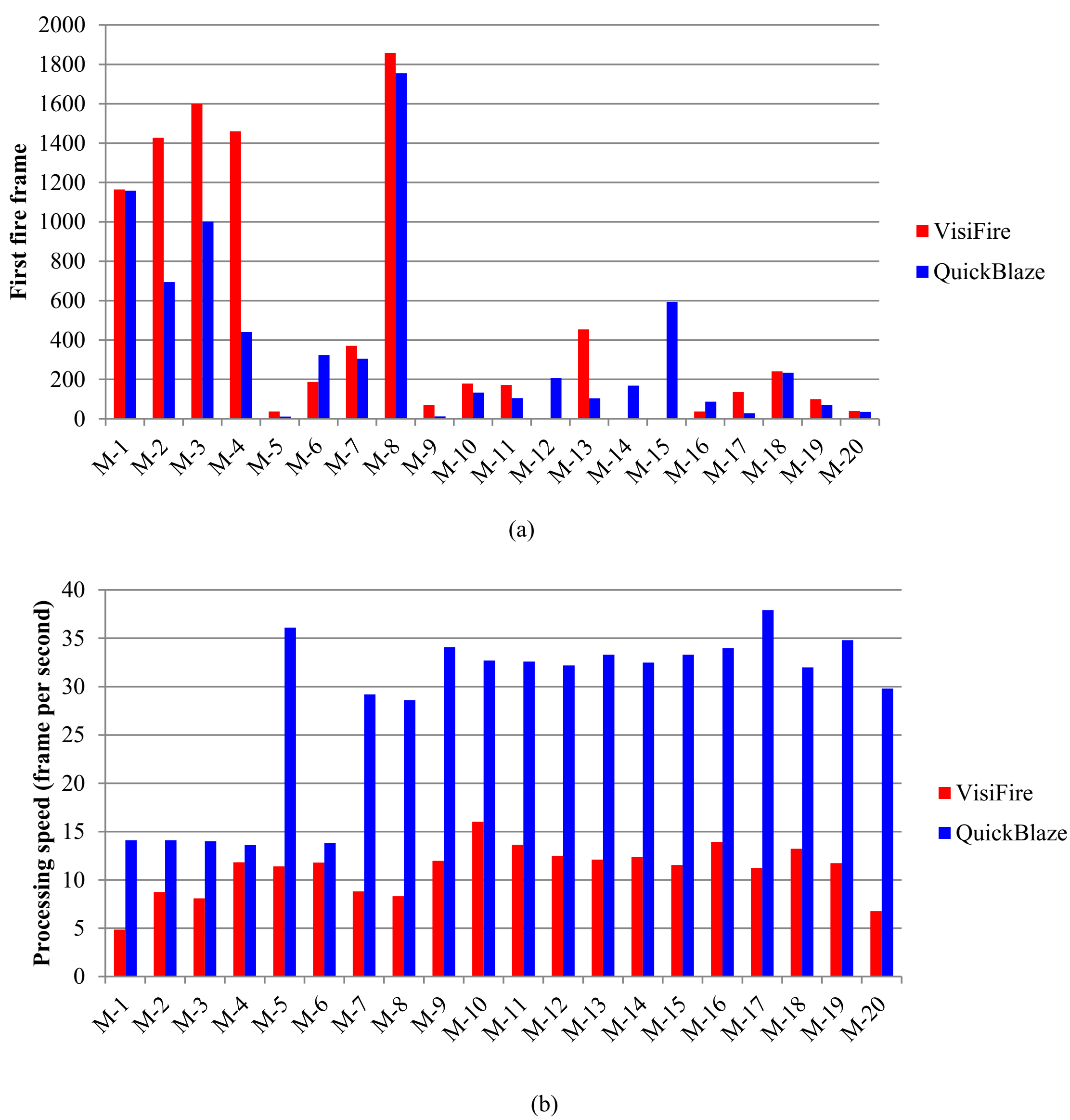

Fig. 8. Experimental results of the early fire detection system using a combined video processing strategy [2]. (a) First fire frame; (b) processing speed.

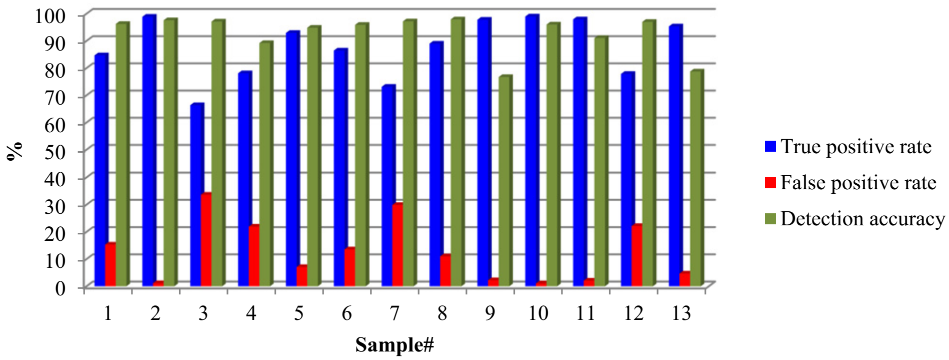

Fig. 9. Evaluation results of the fire flame detection system using K-medoids clustering and swarm intelligence [21].

regarding a non-parametric model. The spatio-temporal consistency energy of all candidate fire regions are predicted by discovering the knowledge base of possible existing fires in neighboring blocks from the current and previous videos. Besides, the authors have used a two-class classifier and SVM algorithm to determine classification of the candidate fire regions.

Foggia et al. [54] have presented a fire detection system that uses a combination of experts based on some of the essential information about color, shape, and flame movements. This system detects fire occurrences using analysis process of the videos obtained by surveillance cameras. Main purposes of the system are categorized into two phases. In the first phase, complementary information and a multi expert system are combined together based on color information, shape variation, and motion analysis. Overall efficiency of the system noticeably is increased using a relatively small effort made by designer. In the second phase, a descriptor is offered by a bag-of-words strategy to indicate movement behavior of the objects. The authors have tested the system on a wide database of fire videos captured in the real and Web environments in terms of sensitivity and specificity.

Qureshi et al. [2] have presented a real-time early fire detection system to detect both of the flame and smoke regions in video images. They have applied a parallel image processing technique for detecting these regions with high speed, low response time, and low false error rate for the open or closed indoor and outdoor environments. The system is designed by the simple image and video processing techniques in a way that the motion and color cues will be computed and segmentation of the flame and smoke candidates from background will be enabled in real-time applications. It balances the color space and separates the flame and smoke detection streams. Candidate regions are detected by both of the streams and color information to do morphological image processing on the regions. Afterwards, candidate regions are filtered via a turbulence flow rate analysis and a smoke detection system. The authors have used manual adjustment of the parameters during a calibration phase without using any offline training. Color balancing will be an essential feature for the system when object segmentation is conducted by a color model and also the system is designed under different illumination situations. It is composed of two phases: estimation of the illumination and chromatic normalization using a scaling factor. The authors have used a "gray-world" algorithm as one of the simplest illuminant prediction techniques. Furthermore, they have calculated average value of image intensities in the R, G, and B planes over every frame. The "gray-value" of image is, in fact, the ultimate vector of three intensities. Afterwards, the R, G, and B planes are independently scaled by a multiplication factor for normalizing the gray-value to average intensity of the frame. If the scaled value is greater than maximum possible intensity then it will be set to the maximum value. As shown in Fig. 6, frames are fed to the independent smoke and flame detection pipelines after balancing the RGB color.

Table 3
Effects of various model architectures on efficiency of the saliency detection and deep learning-based wildfire identification [24].

Model type index	Average processing speed	Training dataset	Validation accuracy
Model 1	31.9 ms	Original image dataset	0.952
		Saliency-based augmented dataset	0.971
Model 2	41.5 ms	Original image dataset	0.974
		Saliency-based augmented dataset	0.98
Model 3	38.6 ms	Original image dataset	0.972
		Saliency-based augmented dataset	0.978
Model 4	67.8 ms	Original image dataset	0.977
		Saliency-based augmented dataset	0.98

Li et al. [55] have presented an autonomous flame detection system in videos by a Dirichlet process Gaussian mixture color model. This system uses a flame detection framework according to the dynamics, color, and flickering features of flames. The Dirichlet process estimates the number of Gaussian components based on training data to make a model for distribution of the flame colors. Moreover, motion saliency and filtered temporal series are extracted as features to describe the dynamics and flickering features of flames. This process is conducted by using a one-dimensional wavelet transform and a probabilistic saliency analysis method. The model assigns a probability value for every pixel to identify major part of the flames according to color information. Afterwards, optical flow magnitude of every pixel is used to achieve a saliency map based on a probabilistic approach. The saliency map and results of the color model are combined together to identify candidate pixels of flames using two independent experimental thresholds.

Fuzzy logic [56,57] can be used to enhance the robustness and performance of fire detection systems. It is a generalization of the classical logic and set theory, which uses an inference process and human-based experiences. Fuzzy decision making is able to predict the complex and undefined conditions based on fuzzy rules and inference engine. Ko et al. [22] have applied stereoscopic pictures and probabilistic fuzzy logic to present a fire detection system and 3D surface reconstruction. This system applies a stereo camera to estimate distance between the given camera and fire regions, and also provide 3D surface of the fire front. Furthermore, color models and a background difference model are used to identify candidate fire regions. Since fire regions constantly change in successive frames, Gaussian membership functions are generated to determine the shape, size, and movement variation of fires. In the next step, these functions are used in fuzzy decision system to acquire real-time fire verification. After fire regions are segmented from left and right images, a matching algorithm extracts feature points to conduct the distance estimation and 3D surface reconstruction. This work not only estimates exact location of the fires, but also determines volume of fire and distance between fire region and camera via a water

Table 4
Comparison results of the forest fire detection by rule-based image processing and temporal variation [26].

Method	True-positive rate (%)	False-negative rate (%)	True-negative rate (%)	False-positive rate (%)	Recall (%)	Precision (%)	F-score (%)
T. -H. Chen et al. [7]	91.33	12	90	13.5	88.4	87.1	87.75
Wang and Ye [64]	88.67	14.33	86.5	11	86.1	89	87.53
Mahmoud and Ren [26]	95	7	93.4	7.6	93.13	92.59	92.86

cannon and an automatic fire suppression system. If a fire occurrence is detected then a warning system will transmit a pre-defined command to the water cannon placed at a remote site to suppress the fire occurrence. Besides, the authors have utilized a pre-calibrated stereo camera instead of a normal one to detect candidate fire regions. This capability will lead to track the fire spread in a 3D perspective and determine an accurate distance from fire to the camera.

4. Discussions and evaluation results

This section evaluates the functionality and efficiency of the intelligent and vision-based fire detection systems that were described in the previous section. Evaluation process is conducted based on the simulation and experimental results of works according to the most important parameters such as accuracy, precision, detection rate, false alarm rate, and processing speed.

Table 1 represents evaluation results of the different patch classifiers for forest fire detection using deep convolutional neural networks [41]. These results are categorized according to three parameters including accuracy, detection rate, and false alarm rate under two train sets and four different features: SVM-Raw, CNN-Raw, SVM-Pool5, and CNN-Pool5. They demonstrate that the CNN-Pool5 technique using Train Set 1 is more accurate than the other techniques. Besides, detection rate is enhanced and false alarm rate is reduced by this technique. Therefore, the CNN-Pool5 technique using Train Set 1 has a high performance under various evaluation scenarios.

Table 2 indicates evaluation results of the fire detection system using a deep CNN [19] compared to the AlexNet architecture [43]. Evaluation process is conducted for Dataset1 in terms of false positives, false negatives, and accuracy while it is conducted for Dataset2 in terms of precision, recall, and F-Measure. These results demonstrate that the fine-tuned feature has positive effect on both of the detection systems under different scenarios. Deep CNN has carried out favorite results in the most cases, especially for the fine-tuned feature. In summary, false positives and false negatives have been decreased and accuracy has been enhanced by deep CNN in the most cases, compared the results of AlexNet for Dataset1. Moreover, deep CNN has improved the precision and F-Measure, and also has decreased the recall in the most results of Dataset2.

Fig. 7 evaluates performance of the fire detection system that uses video sequences and a generic color model [13], compared to the other algorithms including RGB model [7] and RGB model [49]. These results compare all of the methods in terms of detection rate in fire set and false alarm rate in non-fire set. They demonstrate that YCbCr model has high performance compared to the RGB and RGB models under different evaluation parameters. The reason is that YCbCr color space includes an efficient feature to discriminate luminance from chrominance information. It has enhanced detection rate in fire set and has reduced false alarm rate in non-fire set. As mentioned before, the chrominance noticeably indicates information without the effect of luminance. Therefore, the chrominance plane including chrominance-based rules and color model is very effective in flame behavior identification.

Fig. 8 shows experimental results of the fire detection system by a combined video processing strategy, namely QuickBlaze [2], compared to another fire detection platform, namely VisiFire [42]. Comparison results are carried out based on the videos containing fire and/or smoke in terms of first fire frame and processing speed. Fig. 8(a) indicates that the number of fire frames has been decreased by QuickBlaze, comparison with another system. Besides, these results demonstrate that the response time of QuickBlaze is less than that of VisiFire. Fig. 8(b) illustrates processing speed of both the systems for various video samples. Although frame rate is varied for all of the videos regarding the number of candidate fire and smoke regions, QuickBlaze runs faster than VisiFire on all of the samples. Experimental results demonstrate that difference between the performances of these systems is very noticeable, especially in term of processing speed.

Fig. 9 evaluates experimental results of the fire flame detection that uses K-medoids clustering technique and particle swarm optimization [21] for 13 samples in terms of true-positive rate, false-positive rate, and detection accuracy. The results indicate that this detection system has enhanced true-positive rate, except for the 3rd and 7th, so it can be more robust and efficient in real applications. Furthermore, the system does not have many faults in the detection process in which it has obtained low false-positive rates, in the most samples. Since true-positive rates are high and false-positive rates are low, the system has achieved high detection accuracy in the most video samples.

Table 3 represents evaluation results of the saliency detection and deep learning-based wildfire identification [24] that uses the image frames captured by UAV. It compares the results obtained by original image dataset and those obtained by saliency-based augmented dataset. In all of the model architectures, validation accuracy of the saliency-based augmented dataset is higher than that of the original image dataset. The reason is that coefficients of the deep architectures are optimized with adequate training images while the models trained with original image database suffer from a lower training data quality. Besides, the architecture is learned by much more learning samples in the augmented database and, thereby, coefficients and weights are obtained very optimal after the training process.

Table 4 discusses about evaluation results of the forest fire detection system by rule-based image processing and temporal variation [26], compared to the other image-based fire detections [7,64]. Comparison results represent that this system has enhanced true-positive rate, true-negative rate, recall, precision, and F-score as well as it has reduced false-negative rate and false-positive rate, compared to the results of other systems. They indicate that the system is accurate and can be automatically applied in forest fire detection scenarios. It is worth noting that F-score is calculated as follows:

$$F = 2 * \frac{Pre * Rec}{Pre + Rec} \quad (6)$$

where Pre indicates precision and Rec indicates recall.

Table 5 represents error rate results of the detection system for fire occurrences in videos based on color information, compared to some of the other existing related systems. They indicate that Liu and Ahuja

Table 5
Comparison results of the color-based detection system for fire occurrences in videos [50].

Method	False positive (%)	False negative (%)
Borges and Izquierdo [50]	0.68	0.028
Borges et al. [65]	0.9	0.04
Liu and Ahuja [66]	0	0.1
Celik et al. [67]	9.9	1
Ho [68]	0.297	12.36
Töreyin et al. [4]	0.1	0

Table 6
Comparison results of the fire detection system by a combination of experts, color, shape, and flame movements, with state of the art methodologies [54].

Typology	Method		Accuracy (%)	False positive (%)	False negative (%)
Single expert	CE	Celik and Demirel [13]	83.87	29.41	0
	ME	Foggia et al. [54]	71.43	53.33	0
	SV	–	53.57	66.67	21.85
Multi expert system	CE + SV	–	88.29	13.33	9.74
	CE + ME	Di Lascio et al. [69]	92.86	13.33	0
	CE + ME + SV	Foggia et al. [54]	93.55	11.76	0
Others	RGB + shape + motion	Rafiee et al. [70]	74.20	41.18	7.14
	YUV + shape + motion	Rafiee et al. [70]	87.10	17.65	7.14
	Color + shape + motion	Habiboğlu et al. [71]	90.32	5.88	14.29
	Color + shape + motion	Chen et al. [7]	87.10	11.76	14.29

[66] and Töreyin et al. [4] include the lowest false positive compared to the other systems. However, these systems assume that the camera is stationary or use frequency transforms and motion tracking. Thus, they need much more computational processing time and are not suitable for video retrieval. Celik et al. [67] has achieved a good false negative, but the false positive obtained by this method is the highest one among false positives of all the methods. Ho [68] involves a low false-positive rate at the expense of the highest false negative.

Comparison results of the fire detection system by a combination of experts, color, shape, and flame movements [54] with state of the art methodologies in terms of accuracy, false positive, and false negative are represented in Table 6. This evaluation process is conducted based on various topologies including single expert, multi expert system, and some others. Furthermore, it considers several techniques including Expert Color Evaluation (CE), Expert Movement Evaluation (ME), Expert Shape Variation (SV), RGB, YUV, shape, motion, and color. The results indicate that this system has achieved the highest accuracy rate and an acceptable false-positive rate, compared to the other methods. Besides, they represent that this system is lack of any false negative similar to the methods that are presented by Celik and Demirel [13] and Di Lascio et al. [69].

In summary, the simulation and experimental results indicate performance of the fire detection systems that use the intelligent and vision-based techniques. In the detection systems based on CNN and deep CNN, the accuracy and detection rates have been achieved more than around 90% and false alarm rate has been obtained less than around 10%, in the most results. In the systems based on color models, detection rate has been obtained more than around 95%, false alarm rate has been achieved less than around 30%, and processing speed has been improved considerably. In the detection systems that use particle swarm optimization, true-positive rate has been obtained more than around 65%, false-positive rate has been achieved less than around 30%, and detection accuracy has been appeared more than around 75%. Therefore, the fire detection systems that use CNN and deep CNN have the highest performance in the most evaluation results under different scenarios.

5. Benchmark datasets

There are various benchmark datasets to evaluate performance of the fire detection systems, which can be available from different online resources. This section discusses about benchmark datasets of the intelligent and vision-based fire detection systems that were discussed in Section 3.

Zhang et al. [41] have mainly acquired video data from an online resource [42] to present the deep learning-based detection system for forest fires. As describe before, this dataset contains 25 video samples that are gathered in the forest environment with 21 positive (fire) sequences and 4 negative (non-fire) sequences. The images are extracted from the acquired videos with a sample rate of 5 (i.e., one sample image per every five frames). Afterwards, they are resized in the canonical size of 240 $\times$ 320 pixels. The existing fire patches are annotated with 32 $\times$ 32 bounding boxes to evaluate the fire patch localization. Table 7 represents statistics of the trained images, tested images, and the number of annotated patches. The authors have used a small dataset because time consuming for the manual annotation work was limited. They have learned a fire patch classifier (i.e., a binary classification model) using the annotated positive and negative samples via machine learning method. This system generally (i) learns the binary classifier using the annotated patches from scratch, (ii) learns a full-image fire classifier, and (iii) applies the fine-grained patch classifier if the image is classified as contains fire. The linear SVM [60] is used for the linear classifier and the CNN-based Caffe framework [61] is applied for the non-linear classifier. Since the size of the annotated dataset is small, the authors have adopted CIFAR 10 network [62].

In the wildland forest fire smoke detection system presented by Zhang et al. [20], 2800 smoke frames are used to extract smoke from the diversity of smoke images. The smoke is freely changed in shape to being inserted into 12,620 forest background images on the positions that are randomly selected with a deformation process. Every synthetic forest smoke image has one plume of smoke as well as location of the smoke is automatically determined in the inserting process. The forest smoke detection models are trained by RF dataset (Real Model) and

Table 7
Statistical information of the dataset in the deep learning-based detection system for forest fires [41].

Type	The number of images	The number of patches	The number of positive patches	The number of negative patches
Train set	178	12,460	1307	11,153
Test set	59	4130	539	3591

Table 8
Results of cross test between RF dataset and SF dataset in the wildland forest fire smoke detection using faster R-CNN [20].

Model	Test set	The number of samples	The number of smoke samples	The number of non-smoke samples	The number of false non	Detection rate (%)
Real	Real	12,620	12,582	38	8	99.94
Real	Synthetic	12,620	11,793	827	797	93.67
Synthetic	Real	12,620	12,554	66	36	99.71
Synthetic	Synthetic	12,620	12,558	62	32	99.75

Table 9
Average metrics achieved over 50 resized images in the forest fire detection approach for application in unmanned aerial systems [45].

Size (pixels)	Recall (r)	Precision (p)	Dice Index (DI)	Jaccard Index (JI)	Manhattan Index (MI)	Tp (s)	f/s
1920 × 1080	96.30%	97.39%	96.84%	93.88%	97.67%	0.1005	10
1280 × 720	96.17%	97.05%	96.61%	93.44%	97.50%	0.0639	15
960 × 540	95.95%	96.82%	96.43%	93.02%	97.33%	0.0447	22
480 × 270	95.83%	96.70%	96.25%	92.80%	97.25%	0.0275	36
240 × 135	95.71%	96.47%	96.09%	92.48%	97.12%	0.0185	54

Table 10
Details of the original and augmented image databases applied in the saliency detection and deep learning-based wildfire identification [24].

Type	Original database			Augmented database 'UAV_Fire'		
	UAV/aerial/remote-sensing	Ordinary view	Total	UAV/aerial/remote-sensing	Ordinary view	Total
Fire	197	435	632	462	1097	1559
No-fire	250	658	908	677	1325	2002
Total	447	1093	1540	1139	2422	3561

SF dataset (Synthetic Model). Table 8 indicates results of cross test between these datasets. Both the models contain very high detection rate on their own training dataset. Although the samples in SF dataset are not visually realistic, detection performance of Synthetic Model is better than that of Real Model in which Real-Synthetic classification contains 797 False Non.

In the forest fire detection approach for application in unmanned aerial systems [45], the authors have utilized a total of 50 test images from European Fire Database [45] and Forestry Images Organization Database [63] in order to analyze the key goal of this system. All the images are captured from various angles and aerial perspectives in jpeg format and different sizes in the maximum resolution of 1920 × 1080 pixels. In order to evaluate the processing time and performance, firstly they are resized to 1920 × 1080 pixels; then they are resized to 1280 × 720, 960 × 540, 480 × 270, and 240 × 135 pixels using the function “imresize” in Matlab. Table 9 represents average results for all of the test images based on the mentioned metrics in the original and reduced sizes. The last columns indicate the average processing time for each size and the frames per second. The DI is determined in the range of [0, 1] in which value ‘1’ indicates the most accurate detection. It is worth noting that this index is transformed into percentage values to being compared with the other metrics easily.

Zhao et al. [24] have tested the saliency detection and deep learning-based wildfire identification on 1105 images with both flame and smoke features from the original image dataset. The dataset is considered in a way that size of the images is small or the flame and smoke features have covered most of the area. The new created image dataset contains over 3500 images after the data augmentation process. Table 10 represents details of the original and augmented image databases.

Muhammad et al. [19] have used 31 video files to evaluate performance of their fire detection system using the deep CNN. Resolution of the videos is configured in 320 × 240, 352 × 288, 400 × 256, 480 × 272, 720 × 480, 720 × 576, or 800 × 600 pixels. Frame rates are specified as 7, 9, 10, 15, 25, or 29 as well as modality is determined as normal or fire.

In the vision-based approach for fire detection [51], the authors have used 217 images to evaluate their approach in various categories including urban, rural, indoors, and night. Furthermore, they have utilized 12 video files to calculate success rate of the fire regions in which the number of analyzed frames is determined in the range of [265, 3904].

Kim et al. [52] have used a video dataset to evaluate efficiency of the color model-based fire detection algorithm on wireless sensor network in terms of true positives, false positives, false negatives, true negatives, fire detection rate, non-fire detection rate, and overall performance. Table 11 represents more details about the applied datasets and analysis results. As seen in the evaluation results, fire detection rate is obtained more than non-fire detection rate in the first dataset. Moreover, overall performance of the algorithm in the second dataset is better than that in the first dataset.

Qureshi et al. [2] have collected a total of 30 videos to test the early fire detection system using a combined video processing strategy. In the dataset, 20 videos contain flame and/or smoke, and 10 videos are distracters containing no flame or smoke. Videos are captured in various categories such as outdoor plain box, plant pot, and warehouse fire. Duration is specified in the range of [55, 5027] frame and frames per second is determined in the set of {10, 15, 24, 25, 30} frame.

Li et al. [55] have used various testing video files to analyze their autonomous flame detection system. Table 12 represents details and specifications concerned to some of these testing videos. Burning objects are determined based on various materials such as tree, branch, and paper. There are total of 4468 frames in which the frames are different from the training ones. Lighting condition is specified as bright or dark, smoke condition is selected as thin, medium, or thick, as well as the videos are captured in the indoor or outdoor locations.

Table 13 indicates details of the fire and non-fire test videos that Ko et al. [22] have utilized them in evaluation process of the fire detection system and 3D surface reconstruction based on stereoscopic pictures and probabilistic fuzzy logic. The authors have captured 13 types of the stereoscopic videos so that the distance between the camera and fires is determined as different values. Frame rate of the videos varies from 15 to 30 Hz and size of the input images is 320 × 240 pixels. The initial five videos (videos 1–5) are captured from a 5-m distance and the remainder videos (videos 6–10) are captured from a 10-m distance.

6. Conclusions

Fire detection system is one of the surveillance and cyber-physical systems that can be used by fire centers and firefighters to detect and

Table 11
Datasets and analysis results of the color model-based fire detection algorithm on wireless sensor network [52].

Video sequences	Number 1	Number 2
Total frames	122	150
Fire frames	84	0
Non-fire frames	38	150
True positives	78	0
False positives	0	0
False negatives	6	0
True negatives	35	150
Fire detection (%)	92.8	–
Non-fire detection (%)	92	100
Overall (%)	92.6	100

Table 12
Details of the testing videos in the autonomous flame detection system by a Dirichlet process Gaussian mixture color model [55].

Video sequence	Burning objects	Distractors	Positive frames	Negative frames	Lighting condition	Smoke condition	Location
Video 1	Trees	–	230	0	Bright	Thin	Outdoor
Video 2	Trees	–	192	0	Dark	Thin	Outdoor
Video 3	Branches	A walking man	692	0	Bright	Medium	Outdoor
Video 4	Grass	–	386	0	Bright	Medium	Outdoor
Video 5	Papers	A moving light	395	0	Bright	Thin	Indoor
Video 6	Trees	–	202	0	Bright	Thick	Outdoor
Video 7	Assemble line	–	571	70	Bright	Thin	Indoor
Video 8	–	A walking person in red clothes	0	155	Bright	–	Indoor
Video 9	–	Flashing car lights	0	378	Bright	–	Indoor
Video 10	–	Moving cars and people	0	943	Bright	–	Outdoor
Video 11	–	Moving people	0	254	Bright	–	Indoor

locate fire occurrences in various environments. This paper presented overall view and main features of the intelligent and vision-based fire detection systems. Initially, it represented some types of the environments that should be considered by detection systems. Accordingly, fundamental properties of the buildings, forests, and mines are discussed to give a clear view about problems and challenges in these environments. Afterwards, two categorizes of the intelligent and vision-based fire detection systems are discussed with more details, including intelligent detection systems for forest fires and intelligent fire detection systems for all of the environments. The paper discussed about effects of various intelligent techniques such as convolutional neural network (CNN), color models, fuzzy logic, and particle swarm optimization on fire detection processes in various environments. These techniques can considerably enhance performance of the fire detection systems, in the most real scenarios. General characteristics and workflows of the CNN-based detection systems that use faster region-based CNN and deep CNN were discussed in detail. A considerable number of color model-based detection systems have been already presented by researchers. Therefore, main elements and processing pipelines of the fire detection systems applying the combined video processing and rule-based image processing were explained with more details. Furthermore, the paper discussed about effects of fuzzy decision making and particle swarm optimization on detection process of the fire frameworks.

The paper analyzed performance of the most fire detection systems under different evaluation scenarios. The simulation and evaluation results were carried out to evaluate efficiency of the intelligent and vision-based detection systems based on various important parameters such as detection rate, accuracy, precision, true-positive rate, false-positive rate, precision, recall, and processing speed. They demonstrated that these systems have considerable effects on the detection processes compared to the systems that do not use any intelligent and vision-based mechanisms. However, the systems applying the CNN and deep CNN techniques have obtained better results rather than the other fire detection systems. Finally, benchmark datasets of the intelligent fire detection systems were discussed based on various parameters such as the number of samples, location, and material type.

Table 13
Specifications of the fire and non-fire test videos in the fire detection system and 3D surface reconstruction using stereoscopic pictures and probabilistic fuzzy logic [22].

Sequence	Category	The number of frames	Location
Video 1	Fire	333	Outdoor
Video 2	Fire	497	Outdoor
Video 3	Fire	421	Outdoor
Video 4	Fire	368	Outdoor
Video 5	Fire	389	Outdoor
Video 6	Non-fire	242	Indoor
Video 7	Non-fire	200	Indoor
Video 8	Non-fire	557	Indoor
Video 9	Non-fire	650	Outdoor
Video 10	Non-fire	520	Outdoor

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.

References

[1] S.E. Caton, R.S.P. Hakes, D.J. Gorham, A. Zhou, M.J. Gollner, Review of pathways for building fire spread in the wildland urban interface part I: exposure conditions, Fire. Technol 53 (2) (2017) 429–473.
[2] W.S. Qureshi, M. Ekpanyapong, M.N. Dailey, S. Rinsurongkawong, A. Malenichev, O. Krasotkina, QuickBlaze: early fire detection using a combined video processing approach, Fire. Technol 52 (5) (2016) 1293–1317.
[3] E. Mahdipour, C. Dadkhah, Automatic fire detection based on soft computing techniques: review from 2000 to 2010, Artif. Intell. Rev. 42 (4) (2014) 895–934.
[4] 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.
[5] M. Piccardi, Background subtraction techniques: a review, Proceedings of IEEE International Conference on Systems, Man and Cybernetics, The Hague, Netherlands, 10–13 Oct., vol. 4, 2004, pp. 3099–3104.
[6] T.-H. Chen, Y.-H. Yin, S.-F. Huang, Y.-T. Ye, The smoke detection for early fire-alarming system base on video processing, Proceedings of International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP'06), Pasadena, CA, USA, 18–20 Dec. 2006, pp. 427–430.
[7] T.-H. Chen, H. Wu, Y.-C. Chiou, An early fire-detection method based on image processing, Proceedings of International Conference on Image Processing (ICIP'04), Singapore, 24–27 Oct., vol. 3, 2004, pp. 1707–1710.
[8] H.-S. Chu, C.-H. Lu, S.-K. Chang, H.-T. Chao, C.-W. Wang, K.-L. Huang, Method and System for Detecting Flame, Google Patents, 2012.
[9] Y. Cui, H. Dong, E. Zhou, An early fire detection method based on smoke texture analysis and discrimination, Proceedings of Congress on Image and Signal Processing (CISP'08), Sanya, Hainan, China, 27–30 May, vol. 3, 2008, pp. 95–99.
[10] I. Kopilovic, B. Vágvölgyi, T. Szirányi, Application of panoramic annular lens for motion analysis tasks: surveillance and smoke detection, Proceedings of 15th International Conference on Pattern Recognition (ICPR-2000), 3–7 Sept., vol. 4, 2000, pp. 714–717.
[11] H. Tian, W. Li, L. Wang, P. Ogunbona, A novel video-based smoke detection method using image separation, Proceedings of IEEE International Conference on Multimedia and Expo, Melbourne, VIC, Australia, 9–13 July 2012, pp. 532–537.
[12] H. Tian, W. Li, L. Wang, P. Ogunbona, Smoke detection in video: an image separation approach, Int. J. Comput. Vis. 106 (2) (2014) 192–209.
[13] T. Celik, H. Demirel, Fire detection in video sequences using a generic color model, Fire Saf. J. 44 (2) (2009) 147–158.
[14] B.C. Ko, S.J. Ham, J.Y. Nam, Modeling and formalization of fuzzy finite automata for detection of irregular fire flames, IEEE Transactions on Circuits and Systems for Video Technology 21 (12) (2011) 1903–1912.
[15] T. Nguyen-Ti, T. Nguyen-Phuc, T. Do-Hong, Fire detection based on video processing method, Proceedings of International Conference on Advanced Technologies for Communications (ATC 2013), Ho Chi Minh City, Vietnam, 16–18 Oct. 2013, pp. 106–110.
[16] C. Yu, Z. Mei, X. Zhang, A real-time video fire flame and smoke detection algorithm, Procedia Engineering 62 (2013) 891–898.
[17] A.A.A. Alkhatib, A review on forest fire detection techniques, International Journal of Distributed Sensor Networks 10 (3) (2014), 597368.
[18] T. Bouwmans, S. Javed, M. Sultana, S.K. Jung, Deep neural network concepts for background subtraction: A systematic review and comparative evaluation, Neural Networks 117 (2019) 8–66.
[19] K. Muhammad, J. Ahmad, Z. Lv, P. Bellavista, P. Yang, S.W. Baik, Efficient deep CNN-based fire detection and localization in video surveillance applications, IEEE Transactions on Systems, Man, and Cybernetics: Systems (99) (2018) 1–16.
[20] Q.-X. Zhang, G.-H. Lin, Y.-M. Zhang, G. Xu, J.-J. Wang, Wildland forest fire smoke detection based on faster R-CNN using synthetic smoke images, Procedia Engineering 211 (2018) 441–446.

[21] A. Khatami, S. Mirghasemi, A. Khosravi, C.P. Lim, S. Nahavandi, A new PSO-based approach to fire flame detection using K-Medoids clustering, Expert Syst. Appl. 68 (2017) 69–80.
[22] B. Ko, J.-H. Jung, J.-Y. Nam, Fire detection and 3D surface reconstruction based on stereoscopic pictures and probabilistic fuzzy logic, Fire Saf. J. 68 (2014) 61–70.
[23] C. Yuan, Z. Liu, Y. Zhang, Vision-based forest fire detection in aerial images for firefighting using UAVs, Proceedings of International Conference on Unmanned Aircraft Systems (ICUAS), Arlington, VA, USA, 7–10 June 2016, pp. 1200–1205.
[24] Y. Zhao, J. Ma, X. Li, J. Zhang, Saliency detection and deep learning-based wildfire identification in UAV imagery, Sensors 18 (3) (2018), 712.
[25] X. Jin, S. Yin, N. Liu, X. Li, G. Zhao, S. Ge, Color image encryption in non-RGB color spaces, Multimedia Tools and Applications 77 (12) (2018) 15851–15873.
[26] M.A.I. Mahmoud, H. Ren, Forest fire detection using a rule-based image processing algorithm and temporal variation, Mathematical Problems in Engineering 2018 (2018), 7612487.
[27] R.S.P. Hakes, S.E. Caton, D.J. Gorham, M.J. Gollner, A review of pathways for building fire spread in the wildland urban interface part II: response of components and systems and mitigation strategies in the United States, Fire. Technol 53 (2) (2017) 475–515.
[28] V. Babrauskas, Ignition Handbook: Principles and Applications to Fire Safety, Fire Investigation, Risk Management and Forensic Science, Fire Science Publishers, Issaquah, WA, 2003.
[29] S.L. Manzello, S. Suzuki, Y. Hayashi, Enabling the study of structure vulnerabilities to ignition from wind driven firebrand showers: a summary of experimental results, Fire Saf. J. 54 (2012) 181–196.
[30] A. Maranghides, W. Mell, A case study of a community affected by the witch and Guejito wildland fires, Fire. Technol 47 (2) (2011) 379–420.
[31] S. Quarles, P. Leschak, R. Cowger, K. Worley, R. Brown, C. Iskowitz, Lessons Learned From Waldo Canyon: Fire Adapted Communities Mitigation Assessment Team Findings, Fire Adapted Communities Coalition, IBHS Richburg, SC, 2012.
[32] S.L. Manzello, Enabling the investigation of structure vulnerabilities to wind-driven firebrand showers in wildland-urban interface (WUI) fires, Fire Safety Science 11 (2014) 83–96.
[33] An Inconvenient Truth. (Directed by Davis Guggenheim about Former United States Vice President Al Gore's Campaign [Documentary]. Los Angeles, NY, USA).
[34] Y.G. Sahin, Animals as mobile biological sensors for forest fire detection, Sensors 7 (12) (2007) 3084–3099.
[35] G.J. Joy, Evaluation of the approach to respirable quartz exposure control in US coal mines, J. Occup. Environ. Hyg. 9 (2) (2012) 65–68.
[36] H. Wang, Y. Cheng, L. Yuan, Gas outburst disasters and the mining technology of key protective seam in coal seam group in the Huainan coalfield, Nat. Hazards 67 (2) (2013) 763–782.
[37] M.A. Trevits, L. Yuan, A.C. Smith, E.D. Thimons, G.V. Goodman, The status of mine fire research in the United States, Proceedings of the 21st World Mining Congress, Taylor & Francis Group, Krakow, Poland 2008, pp. 303–308.
[38] I.O. Osunmakinde, Towards safety from toxic gases in underground mines using wireless sensor networks and ambient intelligence, International Journal of Distributed Sensor Networks 9 (2) (2013), 159273.
[39] B.W. Jo, R.M.A. Khan, An event reporting and early-warning safety system based on the internet of things for underground coal mines: a case study, Appl. Sci. 7 (9) (2017), 925.
[40] A. Gómez-Ríos, S. Tabik, J. Luengo, A.S.M. Shihavuddin, B. Krawczyk, F. Herrera, Towards highly accurate coral texture images classification using deep convolutional neural networks and data augmentation, Expert Syst. Appl. 118 (2019) 315–328.
[41] Q. Zhang, J. Xu, L. Xu, H. Guo, Deep convolutional neural networks for forest fire detection, Proceedings of the International Forum on Management, Education and Information Technology Application (IFMEITA), Guangzhou, China, January 30–31 2016, pp. 568–575.
[42] B. U. Toreyin. Fire detection dataset. Retrieved December 9, 2018, from http://signal.ee.bilkent.edu.tr/VisiFire/.
[43] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, Advances in Neural Information Processing Systems. NIPS 2012, pp. 1097–1105.
[44] S. Saripalli, J.F. Montgomery, G.S. Sukhatme, Visually guided landing of an unmanned aerial vehicle, IEEE Trans. Robot. Autom. 19 (3) (2003) 371–380.
[45] H. Cruz, M. Eckert, J. Meneses, J.-F. Martínez, Efficient forest fire detection index for application in unmanned aerial systems (UASs), Sensors 16 (6) (2016), 893.
[46] W. Phillips Iii, M. Shah, N. da Vitoria Lobo, Flame recognition in video, Pattern Recogn. Lett. 23 (1–3) (2002) 319–327.
[47] C.-B. Liu, N. Ahuja, Vision based fire detection, Proceedings of the 17th International Conference on Pattern Recognition (ICPR), Cambridge, UK, 26–26 Aug., 4, 2004, pp. 134–137.
[48] 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.
[49] T. Celik, H. Demirel, H. Ozkaramanli, Automatic fire detection in video sequences, Proceedings of the 14th European Signal Processing Conference, Florence, Italy, 4–8 Sept. 2006, pp. 1–5.
[50] P.V.K. Borges, E. Izquierdo, A probabilistic approach for vision-based fire detection in videos, IEEE Transactions on Circuits and Systems for Video Technology 20 (5) (2010) 721–731.
[51] P. Gomes, P. Santana, J. Barata, A vision-based approach to fire detection, International Journal of Advanced Robotic Systems 11 (9) (2014), 149.
[52] Y.-H. Kim, A. Kim, H.-Y. Jeong, RGB color model based the fire detection algorithm in video sequences on wireless sensor network, International Journal of Distributed Sensor Networks 10 (4) (2014), 923609.
[53] K. Dimitropoulos, P. Barmpoutis, N. Grammalidis, Spatio-temporal flame modeling and dynamic texture analysis for automatic video-based fire detection, IEEE Transactions on Circuits and Systems for Video Technology 25 (2) (2015) 339–351.
[54] 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 Transactions on Circuits and Systems for Video Technology 25 (9) (2015) 1545–1556.
[55] Z. Li, L.S. Mihaylova, O. Isupova, L. Rossi, Autonomous flame detection in videos with a Dirichlet process Gaussian mixture color model, IEEE Transactions on Industrial Informatics 14 (3) (2018) 1146–1154.
[56] M. Collotta, FLBA: a fuzzy algorithm for load balancing in IEEE 802.11 networks, J. Netw. Comput. Appl. 53 (2015) 183–192.
[57] J. Alcalá-Fdez, J.M. Alonso, A survey of fuzzy systems software: taxonomy, current research trends, and prospects, IEEE Trans. Fuzzy Syst. 24 (1) (2016) 40–56.
[58] Y.-J. Gong, J.-J. Li, Y. Zhou, Y. Li, H.S.-H. Chung, Y.-H. Shi, J. Zhang, Genetic learning particle swarm optimization, IEEE Transactions on Cybernetics 46 (10) (2016) 2277–2290.
[59] Z. Sun, Y. Liu, L. Tao, Attack localization task allocation in wireless sensor networks based on multi-objective binary particle swarm optimization, J. Netw. Comput. Appl. 112 (2018) 29–40.
[60] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, C.-J. Lin, LIBLINEAR: a library for large linear classification, Journal of Machine Learning Research 9 (Aug) (2008) 1871–1874.
[61] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, T. Darrell, Caffe: convolutional architecture for fast feature embedding, Proceedings of the 22nd ACM International Conference on Multimedia (MM '14), Orlando, Florida, USA, November 03–07 2014, pp. 675–678.
[62] A. Krizhevsky, Learning Multiple Layers of Features from Tiny Images, Master's thesis University of Toronto, April 8, 2009.
[63] Warnell School of Forestry and Natural Resources, The University of Georgia; College of Agricultural and Environmental Sciences; Center for Invasive Species and Ecosystem Health; US Forest Service; International Society of Arboriculture; USDA Identification Technology Program. Forestry Images Organization. Retrieved August 8, 2019, from http://www.forestryimages.org/browse/subimages.cfm?sub=740/.
[64] Y.L. Wang, J.Y. Ye, Research on the algorithm of prevention forest fire disaster in the Poyang Lake ecological economic zone, in advanced materials research, Trans Tech Publ 518 (2012) 5257–5260.
[65] P.V.K. Borges, J. Mayer, E. Izquierdo, Efficient visual fire detection applied for video retrieval, Proceedings of the 16th European Signal Processing Conference, Lausanne, Switzerland, 25–29 Aug. 2008, pp. 1–5.
[66] C.-B. Liu, N. Ahuja, Vision based fire detection, Proceedings of the 17th International Conference on Pattern Recognition (ICPR 2004), Cambridge, UK, 26–26 Aug., vol. 4, 2004, pp. 134–137.
[67] T. Celik, H. Ozkaramanlt, H. Demirel, Fire pixel classification using fuzzy logic and statistical color model, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'07), Honolulu, HI, USA, 15–20 April, vol. 1, 2007, pp. I-1205–I-1208.
[68] C.-C. Ho, Machine vision-based real-time early flame and smoke detection, Meas. Sci. Technol. 20 (4) (2009)(pp. 045502, 13 pp).
[69] R. Di Lascio, A. Greco, A. Saggese, M. Vento, Improving fire detection reliability by a combination of videoanalytics, Proceedings of the International Conference Image Analysis and Recognition (ICIAR), Montreal, Canada, 5–7 July 2014, pp. 477–484.
[70] A. Rafiee, R. Dianat, M. Jamshidi, R. Tavakoli, S. Abbaspour, Fire and smoke detection using wavelet analysis and disorder characteristics, Proceedings of the 3rd International Conference on Computer Research and Development, Shanghai, China, 11–13 March, vol. 3, 2011, pp. 262–265.
[71] 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.