Neurocomputing 288 (2018) 30–42

Contents lists available at ScienceDirect

Neurocomputing

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

Early fire detection using convolutional neural networks during surveillance for effective disaster management

Khan Muhammad, Jamil Ahmad, Sung Wook Baik*

Intelligent Media Laboratory, Digital Contents Research Institute, Sejong University, Seoul, Republic of Korea

A R T I C L E I N F O

Article history:
Received 23 December 2016
Revised 21 March 2017
Accepted 30 April 2017
Available online 29 December 2017

Keywords:
Machine learning
Image classification
Learning vision
Deep learning
Surveillance networks
Fire detection
Disaster management

A B S T R A C T

Fire disasters are man-made disasters, which cause ecological, social, and economic damage. To minimize these losses, early detection of fire and an autonomous response are important and helpful to disaster management systems. Therefore, in this article, we propose an early fire detection framework using fine-tuned convolutional neural networks for CCTV surveillance cameras, which can detect fire in varying indoor and outdoor environments. To ensure the autonomous response, we propose an adaptive prioritization mechanism for cameras in the surveillance system. Finally, we propose a dynamic channel selection algorithm for cameras based on cognitive radio networks, ensuring reliable data dissemination. Experimental results verify the higher accuracy of our fire detection scheme compared to state-of-the-art methods and validate the applicability of our framework for effective fire disaster management.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Disaster management, as a hybrid research area, has attracted the attention of many research communities such as business, computer science, health sciences, and environmental sciences. According to federal emergency management agency policy, there are two main categories of disaster: (1) Technological such as emergencies related to hazardous materials, terrorism, and nuclear power plants etc., and (2) Natural such as floods, earth quakes, and forest fires etc. Regardless of the nature of the disaster, certain characteristics are necessary for effective management of almost all of them. These features include prevention, advance warning, early detection, early notification to the public and concerned authorities, response mobilization, damage containment, and providing medical care as well as relief to affected citizens [1]. Disaster management has four main phases including preparedness, mitigation, response, and recovery, each of which requires different types of data, which are needed by different communities during disaster management. Such data can be processed using data analysis technologies such as information extraction, information retrieval, information filtering, data mining, and decision support [2,3]. An overview of this data flow in disaster management is shown in Fig. 1.

Fig. 1 shows that data is gathered from different sources during disaster management, which are helpful for the detection of disaster, the response of concerned authorities against the disaster and its recovery. Among the given resources, online streaming data from CCTV cameras can be helpful for early detection of different disasters such as fire [4] and flood [5], which in turn can facilitate disaster management teams in quick recovery and reducing the loss of human lives.

Fire disasters mainly occur due to human error or the failure of a system, causing economic as well as ecological damage along with endangering human lives [6]. According to [7], wildfire disasters alone in the year 2015 resulted in 494,000 victims and caused damage worth US$ 3.1 billion. Each year, an area of vegetation of 10,000 km^2 is affected by fire disasters in Europe. The statistic for fire damage is about 100,000 km^2 in Russia and North America. Other examples of fire disasters include (1) the disaster of Arizona (USA, June 2013) which ruined 100 houses and killed 19 firefighters, and (2) the forest fire of California (August 2013) which burned an area of 1042 km^2 and damaged around 111 structures, incurring a firefighting cost of $127.35 million [8]. Considering these examples of damage, early detection of fire is of paramount interest to disaster management systems, so as to avoid such disasters. In this context, researchers have explored different approaches to fire detection including conventional fire alerting systems and visual sensors based systems. The systems belonging to the first category are based on ion or optical sensors, needing close proximity to the fire, and thus failing to provide additional information such as the fire size, location, and degree of burning. In addition to this, such sys-

* Corresponding author.
E-mail addresses: khanmuhammad@sju.ac.kr (K. Muhammad), jamilahmad@sju.ac.kr (J. Ahmad), sbaik@sejong.ac.kr (S.W. Baik).

https://doi.org/10.1016/j.neucom.2017.04.083
0925-2312/© 2017 Elsevier B.V. All rights reserved.

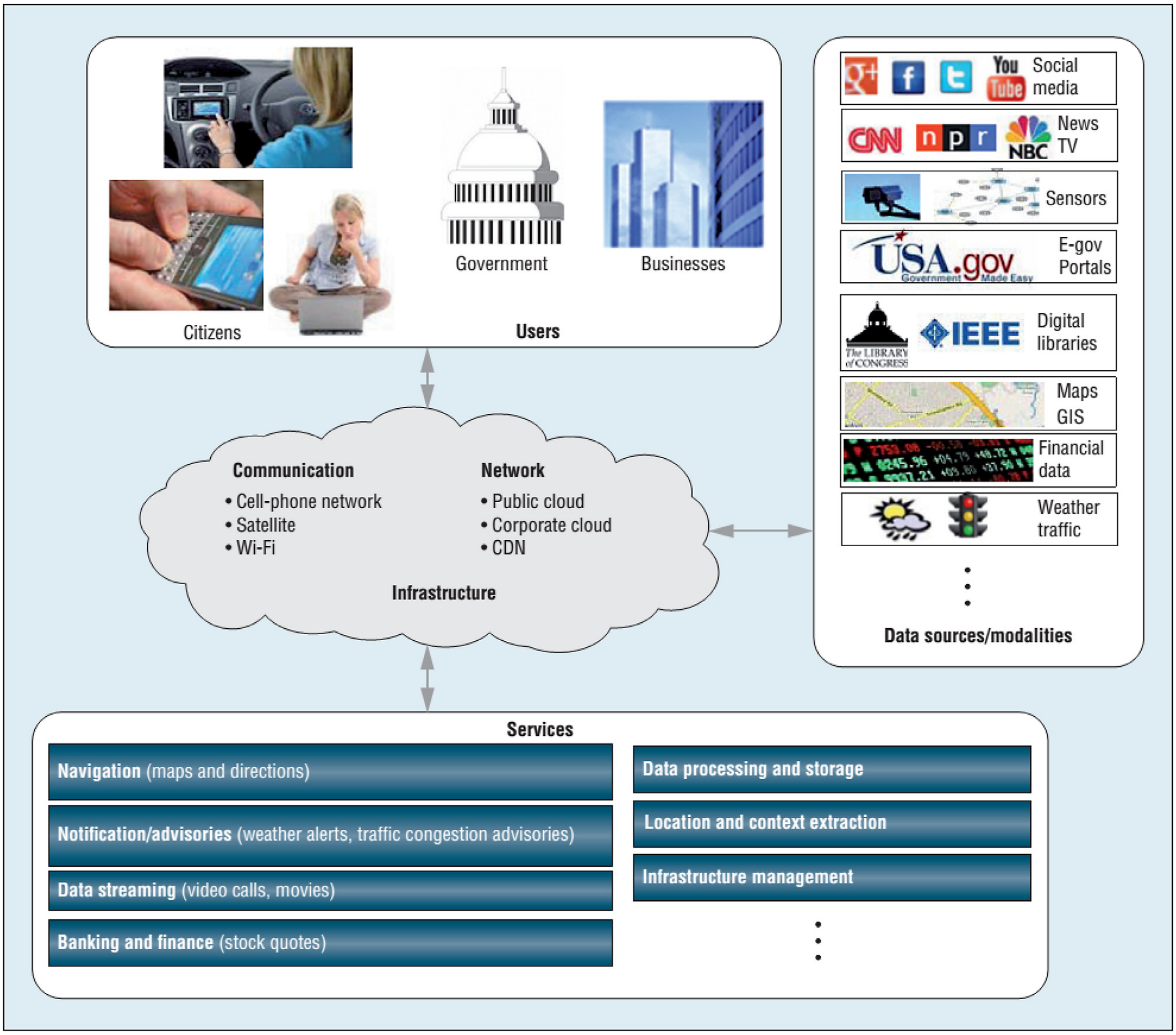

Fig. 1. Flow of data in disaster management system.

tems involve heavy human intervention, such as visiting the fire location to confirm the fire in the event of any fire alarm. To cope with these limitations, many fire detection systems based on visual sensors have been presented [9–12].

Visual sensors based fire detection systems are motivated by several encouraging advantages including: (1) low cost due to the existing setup of installed cameras for surveillance, (2) monitoring of larger regions, (3) comparatively fast response time due to the elimination of waiting time for heat diffusion, (4) fire confirmation without visiting the fire location, (5) flexibility for the detection of smoke and flames through adjustment of certain parameters, and (6) the availability of fire details such as size, location, and degree of burning. Due to these characteristics, they have attracted the attention of many researchers and as a result, many fire detection methods [12–17] have been investigated based on numerous visual features, achieving good performance. But still such methods encounter several problems such as the complexity of the scenes under surveillance due to people and objects looking like fire, the irregularity of lighting (night, day, artificial, shadows, light reflections, and flickering), and the low quality of the captured images, their lower contrast, and lower transmission of signals. These problems demand urgent solutions from the concerned research communities due to their importance to disaster management systems. Further, sending all the streaming data of multiple cameras during surveillance is impractical due to network constraints. In addition to this, an alert of fire and its associated keyframes need an autonomous and reliable communication medium for transmission, to enable the disaster management team to handle it as early as possible.

To address the aforementioned problems, we propose an early fire detection framework using convolutional neural networks (CNNs) and the internet of multimedia things (IoMT) for disaster management. To this end, the major contributions of this study can be summarized as follows:

(1) Unlike traditional hand-engineered features, which are not suitable for the detection of several types of fire, we incorporate deep features of CNNs in our fire detection framework, which can detect fire at an early stage under varying conditions. For this purpose, we used Alexnet as a baseline architecture and fine-tuned it according to our problem, considering the accuracy and complexity.
(2) Due to the emergency nature of fire for disaster management, we propose an adaptive prioritization mechanism for cameras in the surveillance system, which can adaptively switch the sta-

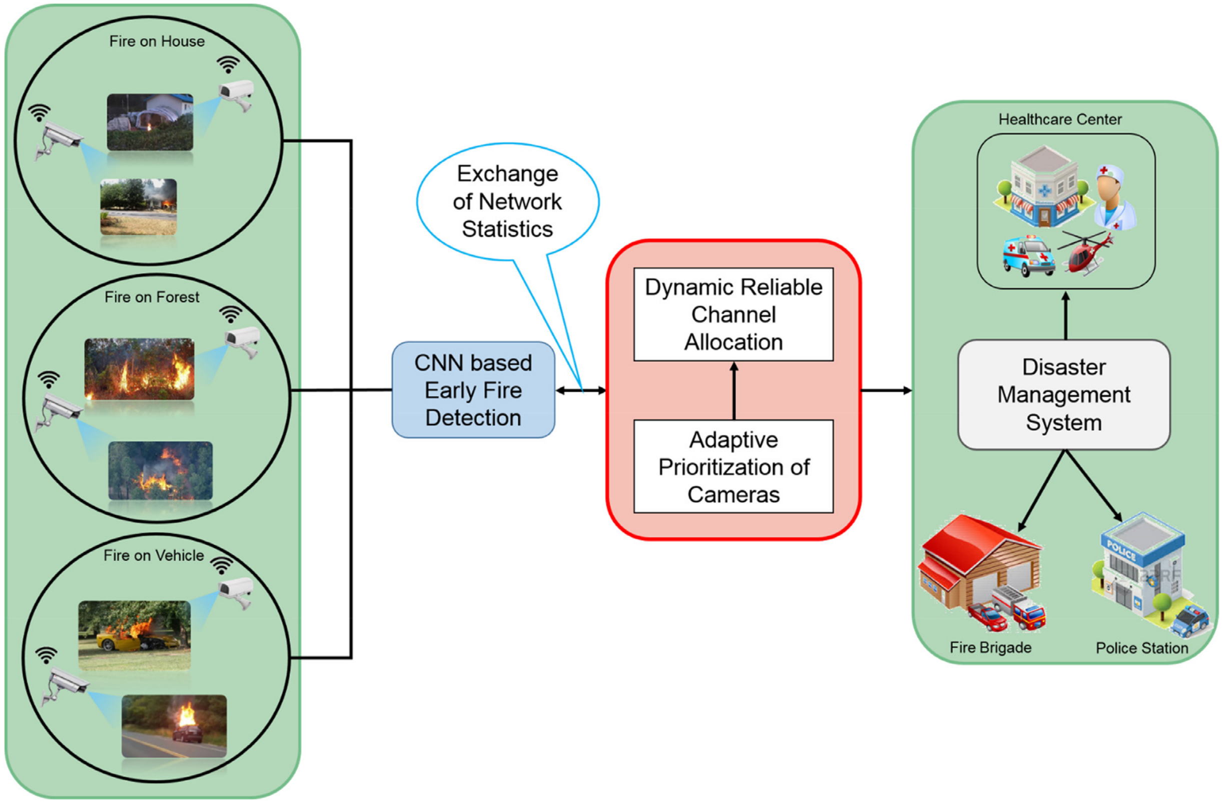

Fig. 2. Early fire detection using CNN with reliable communication for effective disaster management.

tus of camera nodes based on their importance. Furthermore, our system contains a high-resolution camera that can be activated for capturing the important scenes when fire is detected. This can be helpful for disaster management systems in confirming the fire and analyzing the disaster data in real time.

(3) We propose a dynamic channel selection algorithm for high-priority cameras based on cognitive radio networks, ensuring reliable data dissemination and an autonomous response system for disaster management.

The rest of the paper is structured as follows: Related work on fire detection and disaster management is presented in Section 2. Our proposed work is explained in Section 3. Experimental results are provided in Section 4. Finally, our work is concluded in Section 5.

2. Related work

In this section, we first critically discuss the fire detection methods reported in the current literature along with their strengths and weaknesses. Next, we briefly highlight our approach to solving the problems of some of the current methods for early fire detection. Finally, we discuss how early fire detection can be used in effective disaster management systems. Recent advance in technology have resulted in a variety of sensors for different applications such as wireless capsule sensors for visualization of the interior of the human body [18], vehicle sensors for obstacle detection [19], and fire alarm sensors [20]. Current fire alarm sensors, such as infrared, ion, and optical sensors, need close proximity to the heat, fire, radiation or smoke for activation, hence such sensors are not considered good candidates for environments of a critical nature [12]. As an alternative to these sensors, vision-based sensors are widely used, which provide many advantages compared to the traditional sensors, such as lower cost, fast response time, larger coverage of surveillance area, and less human intervention, avoiding the need to visit the location where the fire alert has been triggered [21]. Although vision-based sensors have several encouraging properties, they still encounter some problems with varying lighting conditions, scene complexity, and the lower image quality of cameras due to network constraints. Thus, researchers have made attempts to address these issues. For instance, the authors in [15] explored temporal as well as spatial wavelet analysis and pixels in dynamic regions. Their method achieved good results but it is based on several heuristic thresholds, making it impractical for real-world fire detection applications.

Liu et al. [10] investigated three different models, including spectral, spatial, and temporal, for fire regions in images. However, their method is based on an assumption considering the irregular shape of fire, which is not always the case as moving objects can also change their shape. Another fire detection approach is presented in [22] for forests using contours based on wavelet analysis and FFT. The authors in [23] investigated the YCbCr colour model and devised new rules for the effective separation of luminance and chrominance components, which led them to a rule-based pixels classification of flame. Another colour model YUV along with motion was explored by the authors in [24] for classification into candidate pixels for fire or non-fire. Besides the investigation of color models, specific low-level features of fire regions such as skewness, color, roughness, area size etc. have also been used for determining the frame-to-frame changes, which in combination with a Bayes classifier can recognize fire [17]. Another method is presented in [25] considering a lookup table for detection of fire regions and their confirmation using temporal variation. This method is based on heuristic features, decreasing the certainty of getting the same results while changing the input data.

Considering the heuristic features of [25], the authors in [6] presented a decision rule-based fire detection method using the dynamic analysis of fire along with RGB/HSI color space. Their

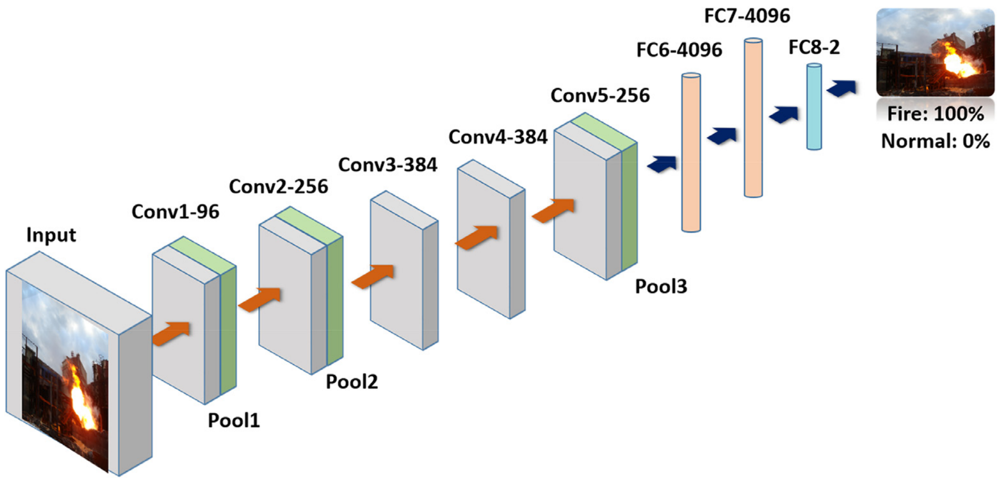

Fig. 3. Architecture of the proposed CNN.

method considers the growth of pixels with the disordered properties of fire for detection. However, it fails to differentiate between moving regions and fire as it is based on frame-to-frame difference. In [26], a fire detection method is proposed by comparison of a normal image with its color information for tunnels. The method is suitable only for static fire situations due to its use of many ad hoc parameters.

Analyzing the above mentioned fire detection techniques, it is observed that the color-based fire detection methods generate more false positives due to their sensitivity to variation in brightness and shadows. For instance, such methods may interpret red-colored vehicles or people wearing red clothes as fire due to its dominant amount. Later on, a possible solution was introduced based on the fact that fire changes its shape continuously, which can differentiate it from moving rigid objects. An example of such methods is presented in [14], where a feature vector is extracted using the optical flow and physical characteristics of fire and can differentiate the flame from moving rigid objects. Another similar method based on dynamic textures and shape features is investigated in [27].

Considering the aforementioned fire detection methods, it can be observed that some of them are too naïve; their execution time is fast but such methods compromise on accuracy, producing a large number of false alarms. Conversely, some methods have achieved good fire detection accuracies but their execution time is too high, hence they cannot be applied in real-world environments especially in critical areas where a minor delay can lead to a huge disaster. Therefore, for more accurate and early detection of fire, we need a robust mechanism that can detect fire during varying conditions and can send the important keyframes and alert immediately to disaster management systems.

3. The proposed framework

Early fire detection in the context of disaster management systems during surveillance of public areas, forests, and nuclear power plants can result in the saving of ecological, economic, and social damage. However, early detection is a challenging problem due to varying lighting conditions, shadows, and the movement of fire-colored objects. Thus, there is a need for an algorithm that can achieve better accuracy in the aforementioned scenarios while minimizing the number of false alarms. To achieve this goal, we explored deep CNNs and devised a fine-tuned architecture for early fire detection during surveillance for effective disaster management systems. After successful fire detection, another desirable requirement is to send an immediate alert to the disaster management system along with the representative keyframes. To this end, we devised an adaptive prioritization scheme for the camera nodes of the surveillance system, considering the contents they perceive. Finally, the data of high-priority nodes is transmitted using a reliable channel selected through our reliable channel selection algorithm. Our system is overviewed in Fig. 2.

3.1. Convolutional neural network architecture

Convolutional neural networks have exhibited state-of-the-art performance in a variety of computer vision tasks including image classification and retrieval [28–30], object detection [31,32], localization [33], and image segmentation [34]. Their success in such a wide variety of applications is attributed to their hierarchical architecture, where they learn discriminative features from raw data automatically. A typical CNN consists of different types of processing layers including convolution, pooling, and fully connected. These layers are arranged in such a way that the output of one layer becomes the input of the next layer. At each convolution layer, a number of kernels are applied on the input data to generate feature maps. Pooling layers select maximum activations within small neighbourhoods of these features maps to reduce their dimensionality and introduce translation invariance. Fully connected layers followed by stacks of convolutional and pooling layers model high-level abstractions in the data and serve as high-level representations of the input. The weights of all the convolutional kernels and neurons in the fully connected layers are learnt during the training process and correspond to essential characteristics of the training data, useful for performing the intended classification [35].

The model we used had a similar architecture to the AlexNet model [36], with modifications according to our problem of interest. It had a total of five convolution layers, three pooling layers, and three fully connected layers as, shown in Fig. 3. As input, the model receives color images of size $224 \times 224 \times 3$. In the first convolution layer, 96 kernels of size 11×11 are applied with a stride of 4 on the input image to generate 96 feature maps. The first pooling layer selects maximum activations from these feature maps in small neighborhoods of 3×3 with a stride of 2 pixels. Consequently, the size of the feature maps is reduced by a factor of 2. The second convolution layer consists of 256 kernels of size 5×5, followed by a max pooling layer similar to the first one. It is followed by a stack of 3 consecutive convolution layers having 384, 384, and 256 kernels, respectively, with uniform kernel sizes of

a. Fire: 100%, Normal: 0.0%　b. Fire: 23.96%, Normal: 76.04%　c. Fire: 87.52%, Normal: 12.48%

d. Fire: 16.87%, Normal: 83.13%　e. Fire: 5.19%, Normal: 94.81%　f. Fire: 94.78%, Normal: 5.22%

Fig. 4. Sample query images along with their probabilities for CNN-based fire detection.

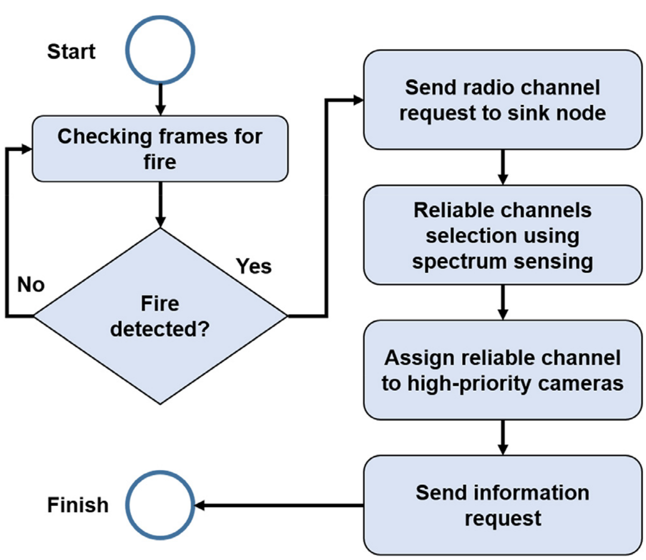

Fig. 5. Flow diagram of dynamic channel selection algorithm and reliable communication.

3×3. The last pooling layer is similar in operation to the first two pooling layers. At the end, there exist three fully connected layers each having 4096, 4096, and 2 neurons (corresponding to the number of classes). The output of the last layer is fed into the Softmax classifier which computes probabilities for the two classes.

Convolutional neural networks usually require a lot of data for training due to the large number of parameters needed to properly tune these networks. Especially the last fully connected layers are very prone to overfitting due to the large number of parameters [37]. To avoid the risk of overfitting in these layers, they are followed by dropout layers, having a dropout ratio of 50%. Several models were trained using the collected training data and their classification performance was assessed using a variety of benchmark datasets. We also evaluated the transfer learning approach to attempt improving classification accuracy, and we observed that it helped us improve classification accuracy by 4–5% on the test set. Transfer learning works on the principle of reusing previously learned knowledge to solve problems more effectively and efficiently [38]. Humans have a natural tendency to apply knowledge across different domains. In the area of deep learning, it has exhibited promise in a wide range of areas. In the current context, we used a pre-trained AlexNet model (trained on ImageNet [39]) and fine-tuned it with our dataset by modifying the last fully connected layer and keeping a slower learning rate (0.001). The slow learning rate allows the previously learned parameters to be minimally adjusted in order to perform the intended classification task. Model fine-tuning was performed for 10 epochs, achieving an improvement of about 5% in classification accuracy compared to the freshly trained model.

3.2. CNN-based fire-detection

After the training and fine-tuning process, a target model is achieved which can be used for prediction of fire at early stages. Unlike conventional fire detection methods, where a lot of effort is required for pre-processing as well as feature engineering, our proposed CNN-based method does not require any pre-processing. Further, it avoids the conventional time-consuming and tedious approaches of extracting hand-crafted features as it learns very powerful features automatically from the provided data in raw form. In addition to this, the proposed CNN-based model learns details at small scales, enabling it to detect fire even at small scale, i.e., in the early stages. For testing, the query image is passed through the proposed model, which results in probabilities for both classes of fire and normal. Based on the higher probability, the image is assigned to its appropriate class. An example of query images along with their probabilities is shown in Fig. 4.

3.3. Dynamic channel selection using cognitive radio networks

Due to congestion, dedicated spectrum allocation is not a feasible solution for multimedia surveillance systems. Therefore, it is

Table 1
Details of dataset 1.

Video name	Resolution	Frames	Frame rate	Modality	Description
Fire1	320 × 240	705	15	Fire	Fire in bucket with person walking around it.
Fire2	320 × 240	116	29	Fire	Fire at comparatively long distance from the camera in a bucket
Fire3	400 × 256	255	15	Fire	Big forest fire
Fire4	400 × 256	240	15	Fire	Same as description of Fire3
Fire5	400 × 256	195	15	Fire	Same as description of Fire3
Fire6	320 × 240	1200	10	Fire	Fire on ground with red colour
Fire7	400 × 256	195	15	Fire	Same as description of Fire3
Fire8	400 × 256	240	15	Fire	Same as description of Fire3
Fire9	400 × 256	240	15	Fire	Same as description of Fire3
Fire10	400 × 256	210	15	Fire	Same as description of Fire3
Fire11	400 × 256	210	15	Fire	Same as description of Fire3
Fire12	400 × 256	210	15	Fire	Same as description of Fire3
Fire13	320 × 240	1650	25	Fire	An indoor environment with fire in bucket
Fire14	320 × 240	5535	15	Fire	Paper box inside which fire is produced.
Fire15	320 × 240	240	15	Normal	Smoke visible from closed window with appearance of red reflection of sun on the glass.
Fire16	320 × 240	900	10	Normal	Smoke pot near red-coloured dust bin.
Fire17	320 × 240	1725	25	Normal	Smoke on ground with nearby moving vehicles and trees
Fire18	352 × 288	600	10	Normal	Smoke far away from camera on hills
Fire19	320 × 240	630	10	Normal	Smoke on red-coloured ground
Fire20	320 × 240	5958	9	Normal	Smoke on hills with nearby red-coloured buildings
Fire21	720 × 480	80	10	Normal	Smoke at a larger distance behind moving trees
Fire22	480 × 272	22,500	25	Normal	Smoke behind hills in front of UOS.
Fire23	720 × 576	6097	7	Normal	Smoke above hills
Fire24	320 × 240	342	10	Normal	Smoke in room
Fire25	352 × 288	140	10	Normal	Smoke at a larger distance from camera in a city
Fire26	720 × 576	847	7	Normal	Same as description of Fire24
Fire27	320 × 240	1400	10	Normal	Same as description of Fire19
Fire28	352 × 288	6025	25	Normal	Same as description of Fire18
Fire29	720 × 576	600	10	Normal	Red-coloured buildings covered by smoke
Fire30	800 × 600	1920	15	Normal	A lab with red-coloured front wall where a person moves holding a red ball
Fire31	800 × 600	1485	15	Normal	A lab with red-coloured tables and person moving with red-coloured bag and ball.

Table 2
Comparison with different fire detection methods on dataset 1.

Technique	False Positives (%)	False-Negatives (%)	Accuracy (%)
Proposed after fine tuning	**9.07**	2.13	**94.39**
Proposed without fine tuning	9.22	10.65	90.06
[13]	11.67	**0**	93.55
[49]	13.33	0	92.86
[50]	5.88	14.29	90.32
RGB [51]	41.18	7.14	74.20
YUV [51]	17.65	7.14	87.10
[23]	29.41	0	83.87
[6]	11.76	14.29	87.10

Table 3
Comparison with different fire detection methods on dataset 2.

Technique	Precision	Recall	F-Measure
Proposed after fine tuning	**0.82**	**0.98**	**0.89**
Proposed without Fine Tuning	0.85	0.92	0.88
[46]	0.4–0.6	0.6–0.8	0.6–0.7
[23]	0.4–0.6	0.5–0.6	0.5–0.6
[54]	0.3–0.4	0.2–0.3	0.2–0.3
[55]	0.6–0.7	0.4–0.5	0.5–0.6

imperative to use a reliable communication mechanism, which can be provided by cognitive radio (CR) by preserving the limited resources of surveillance systems and improving their spectrum utilization [40,41]. The accessing mechanism for the spectrum in CR-assisted sensor networks is of an opportunistic nature and hence additional hardware of reasonable cost with cognitive properties is required to enable the delay-sensitive applications to work properly [42]. It is evident from recent studies that spectrum-related problems such as scarcity, bandwidth, long-range communication, and cost can be resolved by the incorporation of CR in surveillance networks [43,44].

In spectrum sensing (SS), detection of a licensed node's activity is important and to ensure the needs of CR networks, a low probability of false alarms with a higher probability of detection is required. To increase the performance of SS, co-operation among the cognitive nodes is desirable. With this motivation, we employed a co-operative SS algorithm for multi-visual sensors surveillance systems, whose main flow diagram is given in Fig. 5.

Avoiding the conventional packet transmission of the first-in-first-out mechanism, the presented framework assigns more privileges to high priority cameras, allowing them to send their packets reliably to the sink node. Our co-operative SS algorithm consists of n CR-assisted cameras, each with a mechanism of energy detection for local SS [45]. The decision of each individual camera is sent to the sink node, which measures the channel conditions and ranks them based on the results of SS and channel quality parameters. At the end, the most reliable channels are assigned to high priority cameras by the sink node, which can be used for the dissemination of important fire frames to disaster management systems.

4. Results and discussion

This section explains in detail the experiments conducted to evaluate the performance of the proposed framework. Firstly, we provide the experimental setup along with its details. Next, we explain different experiments performed on various datasets from the literature and compare our work with the state-of-the-art methods. Finally, we present the strengths of our method against different attacks.

4.1. Experimental details

All the experiments were performed using a dataset of 68,457 images, collected from different fire datasets of both images and

Fig. 6. Sample images from the dataset 1 videos, with and without fires.

videos, such as Foggia's video dataset [13] containing 62,690 frames, Chino's dataset [46] of 226 images, and other datasets [12, 47]. Following the experimental setup of [13], we used 20% of the data from this dataset for training and 80% for testing. To this end, we trained our model with 10,319 images, of which 5258 images contain fire and 5061 are normal images without fire. The proposed model was trained by the system with specifications as follows: Intel Core i5 CPU equipped with 64 GB RAM with Ubuntu OS, NVidia GeForce GTX TITAN X (Pascal) having 12 GB onboard memory, and Caffe deep learning framework [48]. The rest of the experiments were conducted using MATLAB R2015a with a Core i5 system containing 8 GB RAM.

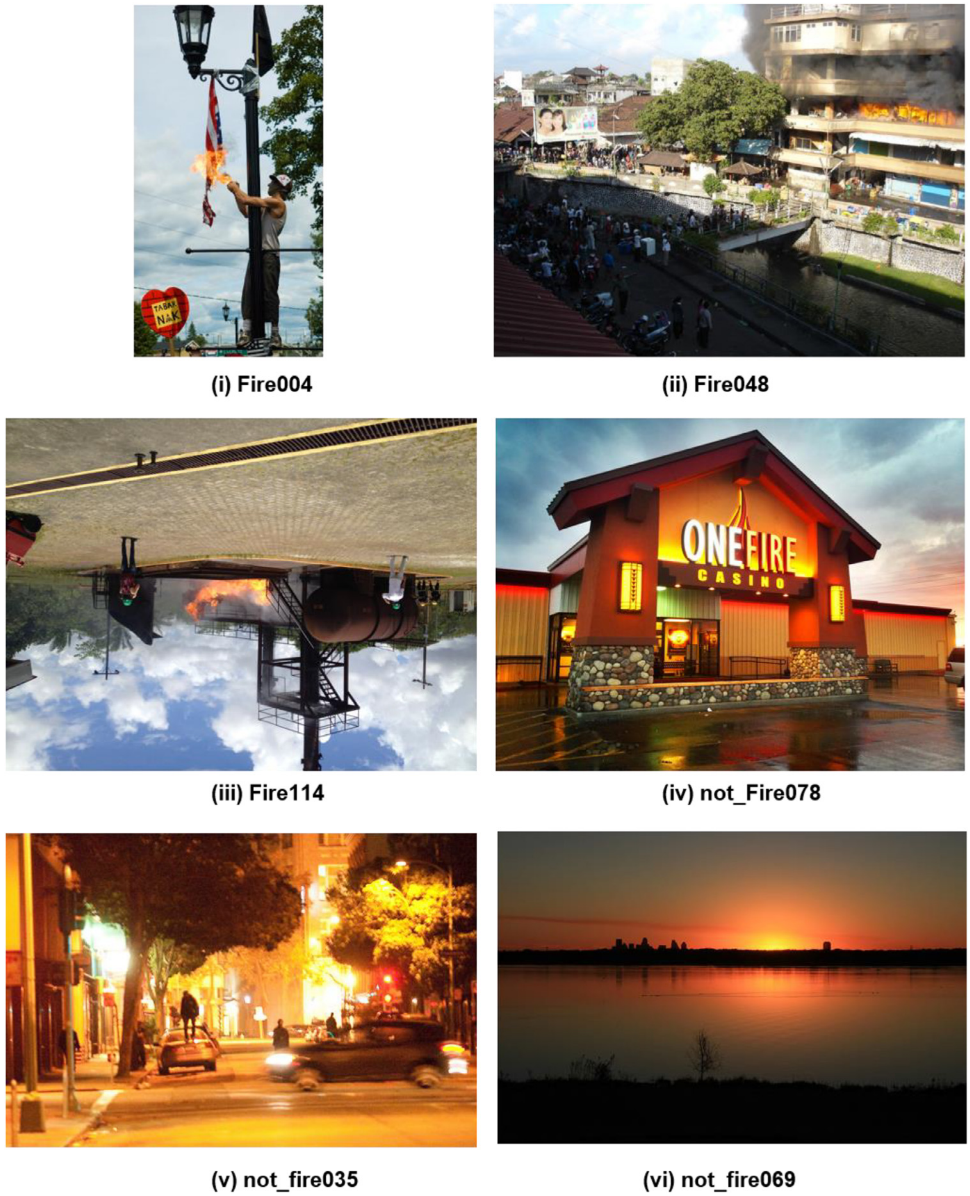

Fig. 7. Sample images from the dataset 2.

4.2. Experiments on different datasets

We mainly focused our experiments on two datasets: Foggia's video dataset [13] and Chino's dataset [46]. The first dataset consists of 31 videos with both indoor and outdoor environments, of which 14 videos contain fire and the remaining 17 videos are without fire. The reasons for selecting this dataset include its large number of videos captured in different scenes in indoor and outdoor environments as well as its challenges. For instance, the last 17 videos contain fire-like objects and situations, which can be predicted as fire, making the classification more difficult. To this end, color-based methods may fail to differentiate between real fire and scenes with red color objects. Similarly, motion-based techniques may wrongly classify a scene with mountains containing smoke, cloud, or fog. These compositions make the dataset more challenging, enabling us to stress our framework and investigate its performance in various situations of the real environment. The detailed information about this dataset and a set of images from it are shown in Table 1 and Fig. 6. The results achieved based on this dataset and its comparison with state-of-the-art fire detection methods are shown in Table 2.

Fig. 6 shows sample images from selected videos of the first dataset of 31 videos. Fig. 6(i)–6(vi) represent the frames containing fire, while Fig. 6(vi)–6(xii) show images containing no fire. It can be noted from the given images that the dataset contains frames belonging to both indoor (Fire13) and outdoor (Fire1, Fire2 etc.) environments. The images also illustrate that some videos contain a large amount of fire, such as Fire3, and some have very little fire, such as Fire13 and Fire14. Another challenge introduced in the dataset is the distance of the fire from the camera. For instance, Fire2 video contains a very small fire at a larger distance. On the other hand, Fire13 video indicates a small fire but at a comparatively small distance. Besides this, there are red-colored objects and grounds such as a signboard (Fire14) and reddish grass (Fire6) in many videos, making the dataset very challenging for fire detection methods.

The proposed work is compared with other related methods in Table 2. The existing methods for comparison are selected carefully

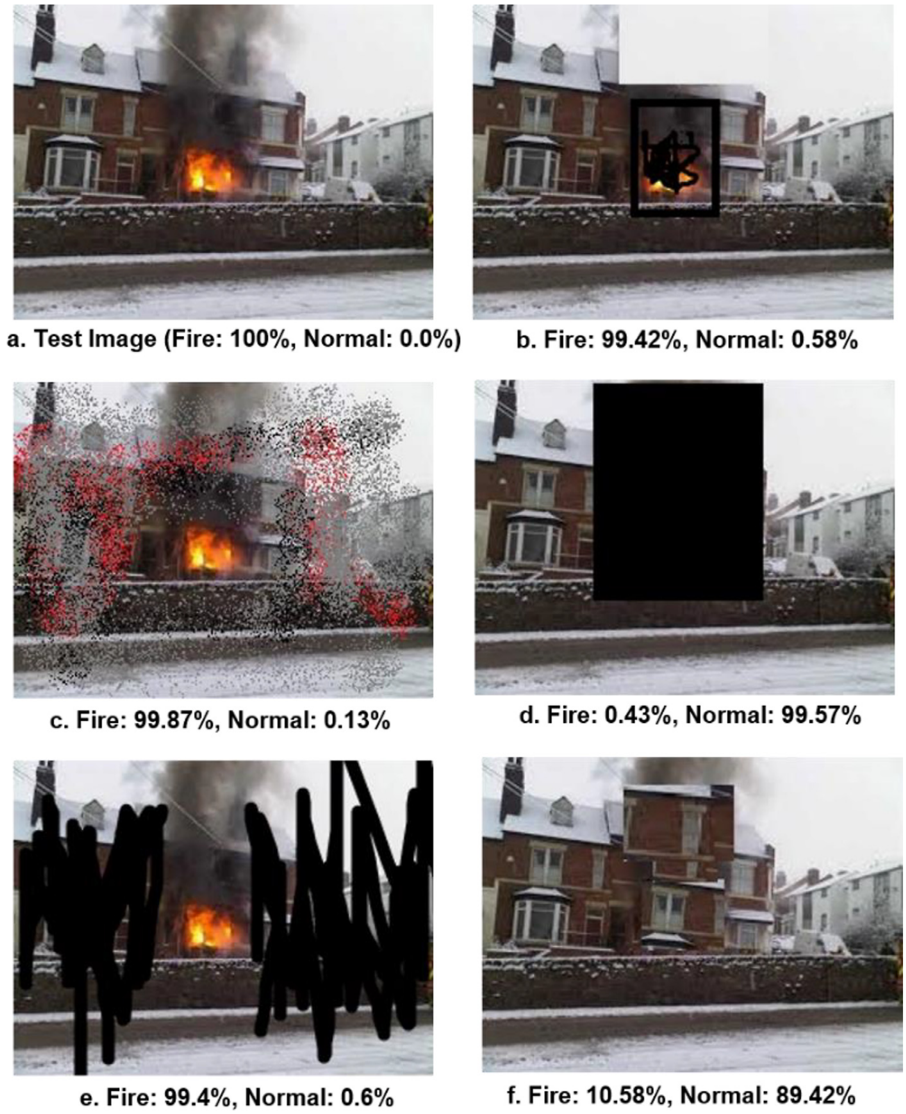

Fig. 8. Illustration of the effect of noise on the performance of our fire detection scheme. Images (a, b, c, and e) are predicted as fire while images (d and f) are predicted as normal.

considering the underlying dataset, year of publication, and features used. For instance, the selected existing methods are based on different features such as shape, color, and motion [51] and [50] with the range of publication of 2004–2015. From the results given in Table 2, the best method in terms of false positives is [50], but its false negatives are greater than other methods except for [6]. In addition, its accuracy is 90.32%, which is less than two existing methods [13] and [49] as well as the proposed work. The work of [13] is good compared to other methods, but the false positives are still 11.67% and there is still room for improvement in both accuracy and false positives. The proposed work, inspired from deep features, reported further improvement by increasing the accuracy from 93.55% to 94.39% and reducing the false positives from 11.67% to 9.07%. Although, our work also resulted in false negatives of 2.13%, it still maintained a better balance between accuracy, false positives, and false negatives, making our method more suitable for early fire detection, which is of paramount interest to disaster management systems.

The second dataset [46] is comparatively small but very challenging. The total number of images in this dataset is 226, out of which 119 images contain fire while the remaining 107 are fire-like images containing sunsets, fire-like lights, and sunlight coming through windows etc. A set of selected images from this dataset are shown in Fig. 7. For better evaluation of the performance, the results for this dataset are collected using another set of metrics including precision [52], recall, and F-measure [53]. The results achieved by our method using this dataset are reported and compared with existing methods in Table 3. By using deep features and fine tuning our fire detection model, we successfully outperformed the state-of-the-art methods by achieving the highest score of precision 0.82, recall 0.98, and F-measure 0.89, validating the effectiveness of our early fire detection method.

4.3. Robustness evaluation

In this section, we investigate the robustness of our fire detection method using different tests such as noise attacks, cropping, and rotations. Fig. 8(a) shows a test image containing fire, which is predicted as fire by our method with accuracy of 100%. In Fig. 8(b), we blocked the major part of the fire and passed the image through our framework. The image is still predicted as fire with accuracy of 99.42%. In Fig 8(c) and 8(e), we attacked the image with noise, yet our method successfully predicted the resulting images as fire with accuracy around 99%. Finally, in Fig. 8(d)

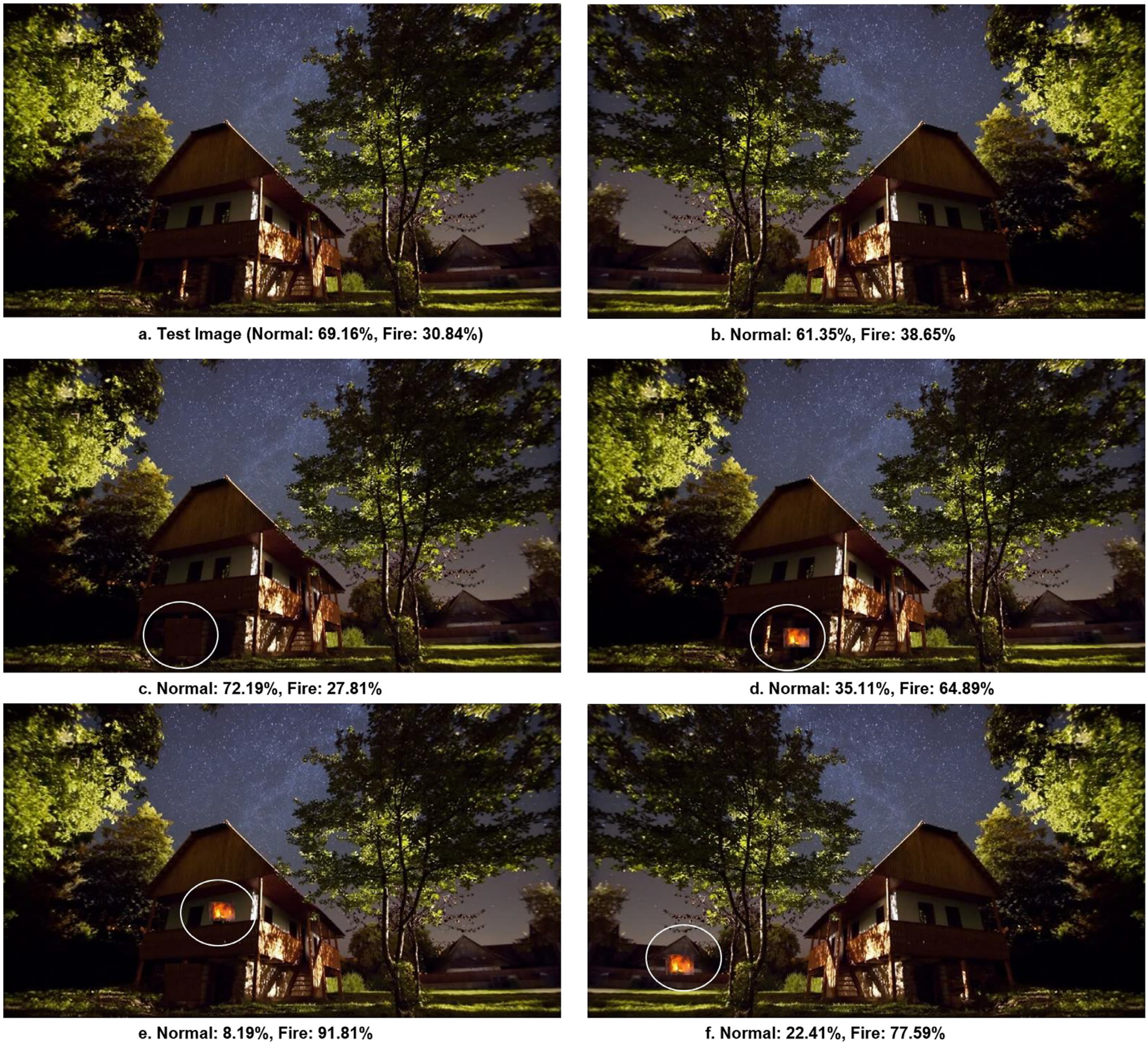

Fig. 9. Illustration of fire detection using a challenging test image. The white circles show the modified regions of the input image for different tests.

and 8(f), we tested how accurately our method has modelled the fire. To this end, we blocked the fire part of the images and passed them through our framework. It can be seen that our technique successfully predicted them as "normal" with accuracy 99.57% (d) and 89.42%, respectively. Considering the results of the various tests, it is evident that our method can detect fire at early stages under varying conditions in spite of noisy images, which can occur in the real world during surveillance.

In Figs. 9 and 10, we investigated the performance of our method against other tests using a normal test image in which some parts look like fire, making it challenging for fire detection methods. Fig. 9(a) is the input normal test image while Fig. 9(b) is its flipped version. Both of them are predicted as normal by our method with accuracy of 69.16% and 61.35%, respectively. In Fig. 9(c), a fire-like portion of the image has been blocked, showing an increase in accuracy from 69.16% to 72.19%. In Fig. 9 (d–f), a small portion of real fire is placed on different regions of the normal image and is passed through our model. It can be seen that our method has predicted them as fire despite the small size of the fire, showing the effectiveness of this fire detection method at early stages. Fig. 10 illustrates the effects on the performance of our approach against different rotations. Fig. 10(a) is the input normal image while Fig. 10(b) is the rotated image at 180%. Fig. 10(c) is the rotated version of Fig. 10(a) where the rotation is 90°. In Fig. 10(d), a small amount of real fire is placed on Fig. 10(b) and is tested with our method. It is evident from all cases that our method can successfully differentiate between fire and normal images and can detect fire at early stages, which is helpful to disaster management systems.

Apart from the above mentioned evaluation from different aspects, it is important to discuss the computational complexity of a fire detection algorithm. Our proposed algorithm can process

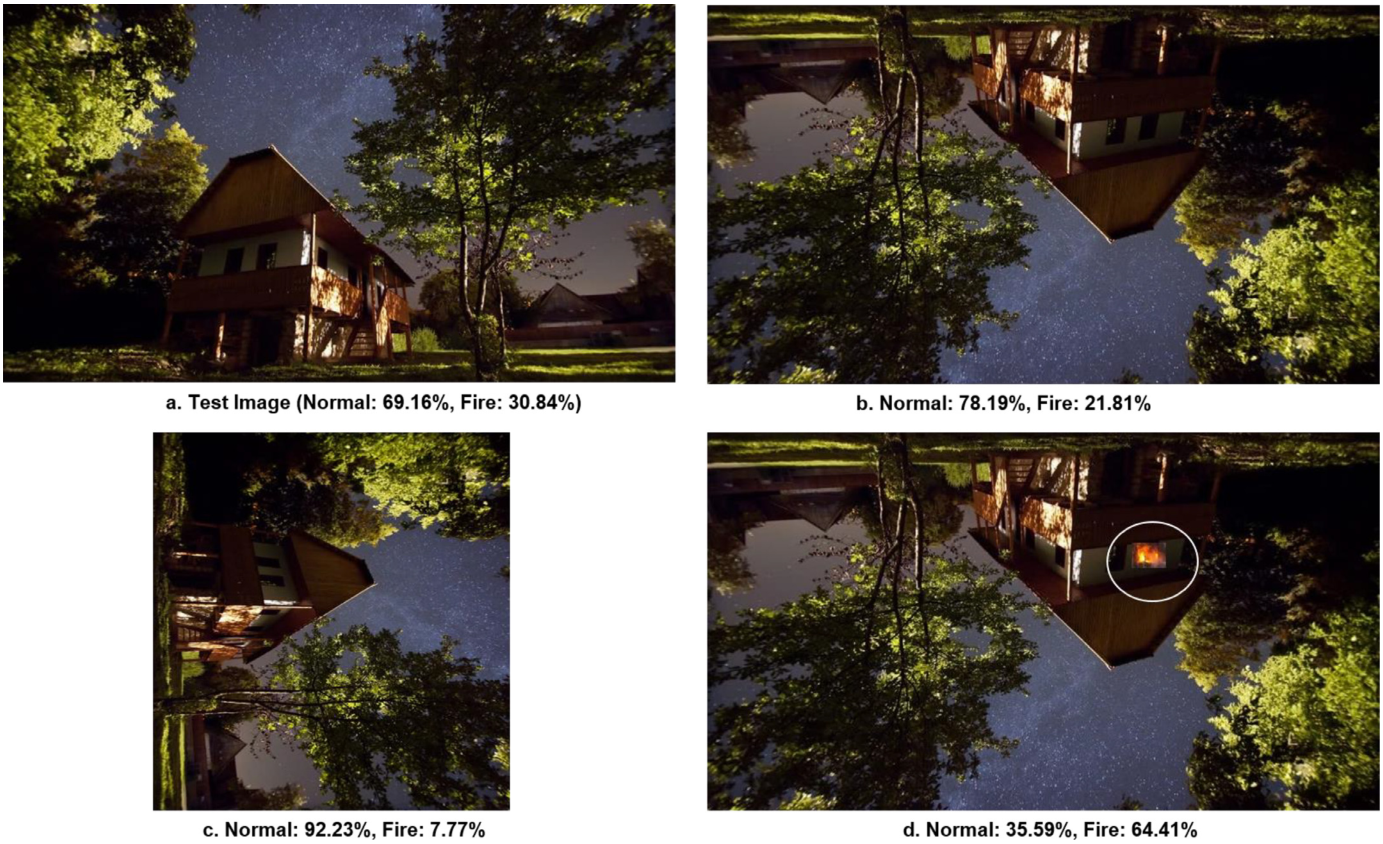

Fig. 10. Performance evaluation of our method against different rotations. The regions enclosed in white circles represent the modified parts of images for different tests.

17 frames/s using the specification mentioned in Section 4.1, which is sufficient to detect fire at early stages using cameras working on 25–30 frames/s.

5. Conclusions

Due to recent advances, CCTV cameras are able to perform different types of processing such as object and motion detection and tracking. Considering these processing capabilities, it is possible to detect fire at its early stage during surveillance, which can be helpful to disaster management systems, avoiding huge ecological and economic losses, as well as saving a large number of human lives. With this motivation, we proposed an early fire detection method based on fine-tuned CNNs during CCTV surveillance. Incorporating deep features in our framework, we showed that fire can be detected at early stages with higher accuracy in varying indoor and outdoor environments while minimizing the false fire alarms. Another desirable aspect of disaster management is autonomous response and reliable communication, for which we proposed a prioritization mechanism that can adaptively change the priority of camera nodes based on the importance of the contents it perceives. The reliability of the important frames and early response to the disaster management system is ensured by a dynamic channel selection scheme using cognitive radio networks. Through experiments on videos containing fire-like moving objects and real fire in indoor and outdoor environments, we confirmed that our framework can detect fire at early stages with good accuracy and minimum false fire alarms, as well as ensuring an autonomous response and reliable transmission of representative contents under surveillance, which can greatly facilitate disaster management systems.

The proposed system improved the fire detection accuracy, with minimum false alarms, but the model size is comparatively heavy, i.e., 238 MB. In future work, we plan to explore light-weight CNNs for reducing the model size while keeping a balance between accuracy and false alarms. Besides this, the proposed framework disseminates the important frames with no authentication mechanism at the disaster management system. In this context, data hiding approaches such as steganography [56,57] and watermarking [58] can be used for embedding some information inside keyframes for authentication purposes, as reported in recent works for social networks [59,60].

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2016R1A2B4011712).

References

[1] V. Hristidis, S.-C. Chen, T. Li, S. Luis, Y. Deng, Survey of data management and analysis in disaster situations, J. Syst. Softw. 83 (2010) 1701–1714.
[2] N.R. Adam, B. Shafiq, R. Staffin, Spatial computing and social media in the context of disaster management, IEEE Intell. Syst. 27 (2012) 90–96.
[3] X. Song, L. Sun, J. Lei, D. Tao, G. Yuan, M. Song, Event-based large scale surveillance video summarization, Neurocomputing 187 (2016) 66–74.
[4] R. Chi, Z.-M. Lu, Q.-G. Ji, Real-time multi-feature based fire flame detection in video, IET Image Process. 11 (2016) 31–37.
[5] S.-W. Lo, J.-H. Wu, F.-P. Lin, C.-H. Hsu, Cyber surveillance for flood disasters, Sensors 15 (2015) 2369–2387.
[6] T.-H. Chen, P.-H. Wu, Y.-C. Chiou, An early fire-detection method based on image processing, in: Proceedings of International Conference on Image Processing, ICIP'04, 2004, 2004, pp. 1707–1710.
[7] D. Guha-Sapir, F. Vos, R. Below, S. Penserre, Annual disaster statistical review 2015: the numbers and trends, 2015. http://www.cred.be/sites/default/files/ADSR_2015.pdf.

[8] T. Toulouse, L. Rossi, M. Akhloufi, T. Celik, X. Maldague, Benchmarking of wildland fire colour segmentation algorithms, IET Image Process. 9 (2015) 1064–1072.
[9] T. Qiu, Y. Yan, G. Lu, An autoadaptive edge-detection algorithm for flame and fire image processing, IEEE Trans. Instrum. Meas. 61 (2012) 1486–1493.
[10] C.-B. Liu, N. Ahuja, Vision based fire detection, in: Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004, 2004, pp. 134–137.
[11] T. Celik, H. Demirel, H. Ozkaramanli, M. Uyguroglu, Fire detection using statistical color model in video sequences, J. Vis. Commun. Image Represent. 18 (2007) 176–185.
[12] B.C. Ko, S.J. Ham, J.Y. Nam, Modeling and formalization of fuzzy finite automata for detection of irregular fire flames, IEEE Trans. Circuits Syst. Video Technol. 21 (2011) 1903–1912.
[13] 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. Circuits Syst. Video Technol. 25 (2015) 1545–1556.
[14] M. Mueller, P. Karasev, I. Kolesov, A. Tannenbaum, Optical flow estimation for flame detection in videos, IEEE Trans. Image Process. 22 (2013) 2786–2797.
[15] 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 (2006) 49–58.
[16] R.C. Luo, K.L. Su, Autonomous fire-detection system using adaptive sensory fusion for intelligent security robot, IEEE/ASME Trans. Mechatron. 12 (2007) 274–281.
[17] P.V.K. Borges, E. Izquierdo, A probabilistic approach for vision-based fire detection in videos, IEEE Trans. Circuits Syst. Video Technol. 20 (2010) 721–731.
[18] I. Mehmood, M. Sajjad, S.W. Baik, Mobile-cloud assisted video summarization framework for efficient management of remote sensing data generated by wireless capsule sensors, Sensors 14 (2014) 17112–17145.
[19] A. Sorbara, E. Zereik, M. Bibuli, G. Bruzzone, M. Caccia, Low cost optronic obstacle detection sensor for unmanned surface vehicles, in: Proceedings of IEEE Sensors Applications Symposium, SAS, 2015, pp. 1–6.
[20] B.C. Ko, K.-H. Cheong, J.-Y. Nam, Fire detection based on vision sensor and support vector machines, Fire Saf. J. 44 (2009) 322–329.
[21] K. Dimitropoulos, P. Barmpoutis, N. Grammalidis, Spatio-temporal flame modeling and dynamic texture analysis for automatic video-based fire detection, IEEE Trans. Circuits Syst. Video Technol. 25 (2015) 339–351.
[22] Z. Zhang, J. Zhao, D. Zhang, C. Qu, Y. Ke, B. Cai, Contour based forest fire detection using FFT and wavelet, in: Proceedings of International Conference on Computer Science and Software Engineering, 2008, pp. 760–763.
[23] T. Celik, H. Demirel, Fire detection in video sequences using a generic color model, Fire Saf. J. 44 (2009) 147–158.
[24] G. Marbach, M. Loepfe, T. Brupbacher, An image processing technique for fire detection in video images, Fire Saf. J. 41 (2006) 285–289.
[25] W. Phillips Iii, M. Shah, N.da Vitoria Lobo, Flame recognition in video, Pattern Recogn. Lett. 23 (2002) 319–327.
[26] D. Han, B. Lee, Development of early tunnel fire detection algorithm using the image processing, in: Proceedings of International Symposium on Visual Computing, 2006, pp. 39–48.
[27] A. Rahman, M. Murshed, Detection of multiple dynamic textures using feature space mapping, IEEE Trans. Circuits Syst. Video Technol. 19 (2009) 766–771.
[28] 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 (2015) 5017–5032.
[29] 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.
[30] J. Yang, B. Jiang, B. Li, K. Tian, Z. Lv, A fast image retrieval method designed for network big data, IEEE Trans. Ind. Inform. 13 (5) (2017) 2350–2359.
[31] 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 (2016) 142–158.
[32] S. Anwar, K. Hwang, W. Sung, Fixed point optimization of deep convolutional neural networks for object recognition, in: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP , 2015, pp. 1131–1135.
[33] V. Kantorov, M. Oquab, M. Cho, I. Laptev, ContextLocNet: context-aware deep network models for weakly supervised localization, in: Proceedings of European Conference on Computer Vision, 2016, pp. 350–365.
[34] W. Zhang, R. Li, H. Deng, L. Wang, W. Lin, S. Ji, et al., Deep convolutional neural networks for multi-modality isointense infant brain image segmentation, NeuroImage 108 (2015) 214–224.
[35] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, F.E. Alsaadi, A survey of deep neural network architectures and their applications, Neurocomputing 234 (2017) 11–26.
[36] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Proceedings of Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.
[37] M. Sun, Z. Song, X. Jiang, J. Pan, Y. Pang, Learning pooling for convolutional neural network, Neurocomputing 224 (2017) 96–104.
[38] S.J. Pan, Q. Yang, A survey on transfer learning, IEEE Trans. knowl. Data Eng. 22 (2010) 1345–1359.
[39] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, Imagenet: a large-scale hierarchical image database, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2009, pp. 248–255.
[40] A.O. Bicen, V.C. Gungor, O.B. Akan, Delay-sensitive and multimedia communication in cognitive radio sensor networks, Ad Hoc Netw. 10 (2012) 816–830.
[41] D. Jiang, X. Ying, Y. Han, Z. Lv, Collaborative multi-hop routing in cognitive wireless networks, Wirel. Pers. Commun. 86 (2016) 901–923.
[42] I.F. Akyildiz, T. Melodia, K.R. Chowdhury, A survey on wireless multimedia sensor networks, Comput. Netw. 51 (2007) 921–960.
[43] R. Umar, A.U. Sheikh, A comparative study of spectrum awareness techniques for cognitive radio oriented wireless networks, Phys. Commun. 9 (2013) 148–170.
[44] I. Mehmood, M. Sajjad, W. Ejaz, S.W. Baik, Saliency-directed prioritization of visual data in wireless surveillance networks, Inf. Fusion 24 (2015) 16–30.
[45] W. Zhang, R.K. Mallik, K.B. Letaief, Optimization of cooperative spectrum sensing with energy detection in cognitive radio networks, IEEE Trans. Wirel. Commun. 8 (2009).
[46] 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: Proceedings of the 28th 2015 Conference on Graphics, Patterns and Images, SIBGRAPI , 2015, pp. 95–102.
[47] S. Verstockt, T. Beji, P. De Potter, S. Van Hoecke, B. Sette, B. Merci, et al., Video driven fire spread forecasting (f) using multi-modal LWIR and visual flame and smoke data, Pattern Recogn. Lett. 34 (2013) 62–69.
[48] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, et al., Caffe: convolutional architecture for fast feature embedding, in: Proceedings of the 22nd ACM International Conference on Multimedia, 2014, pp. 675–678.
[49] R. Di Lascio, A. Greco, A. Saggese, M. Vento, Improving fire detection reliability by a combination of videoanalytics, in: Proceedings of International Conference on Image Analysis and Recognition, 2014, pp. 477–484.
[50] Y.H. Habiboğlu, O. Günay, A.E. Çetin, Covariance matrix-based fire and flame detection method in video, Mach. Vis. Appl. 23 (2012) 1103–1113.
[51] A. Rafiee, R. Dianat, M. Jamshidi, R. Tavakoli, S. Abbaspour, Fire and smoke detection using wavelet analysis and disorder characteristics, in: Proceedings of the 3rd International Conference on Computer Research and Development ICCRD , 2011, pp. 262–265.
[52] K. Muhammad, M. Sajjad, M.Y. Lee, S.W. Baik, Efficient visual attention driven framework for key frames extraction from hysteroscopy videos, Biomed. Signal Process. Control 33 (2017) 161–168.
[53] K. Muhammad, J. Ahmad, M. Sajjad, S.W. Baik, Visual saliency models for summarization of diagnostic hysteroscopy videos in healthcare systems, SpringerPlus 5 (2016) 1495.
[54] L. Rossi, M. Akhloufi, Y. Tison, On the use of stereovision to develop a novel instrumentation system to extract geometric fire fronts characteristics, Fire Saf. J. 46 (2011) 9–20.
[55] S. Rudz, K. Chetehouna, A. Hafiane, H. Laurent, O. Séro-Guillaume, Investigation of a novel image segmentation method dedicated to forest fire applications, Meas. Sci. Technol. 24 (2013) 075403.
[56] K. Muhammad, M. Sajjad, I. Mehmood, S. Rho, S.W. Baik, A novel magic LSB substitution method (M-LSB-SM) using multi-level encryption and achromatic component of an image, Multimed.Tools Appl. 75 (2016) 14867–14893.
[57] R.J. Mstafa, K.M. Elleithy, A video steganography algorithm based on Kanade–Lucas–Tomasi tracking algorithm and error correcting codes, Multimed. Tools Appl. 75 (2016) 10311–10333.
[58] X.-L. Liu, C.-C. Lin, S.-M. Yuan, Blind dual watermarking for color images' authentication and copyright protection, IEEE Trans. Circuit Syst. Video Technol. (2016), doi:10.1109/TCSVT.2016.2633878.
[59] K. Muhammad, M. Sajjad, I. Mehmood, S. Rho, S.W. Baik, Image steganography using uncorrelated color space and its application for security of visual contents in online social networks, Future Gener. Comput. Syst. (2016), doi:10.1016/j.future.2016.11.029.
[60] K. Muhammad, J. Ahmad, S. Rho, S.W. Baik, Image steganography for authenticity of visual contents in social networks, Multimed. Tools Appl. 76 (18) (2017) 18985–19004.

Khan Muhammad (S'16) received the bachelors degree in computer science from Islamia College Peshawar, Pakistan with research in information security. He is currently pursuing M.S. leading to Ph.D. degree in digitals contents from College of Software and Convergence Technology, Sejong University, Seoul, Republic of Korea. He has been a Research Associate with the Intelligent Media Laboratory (IM Lab) since 2015. His research interests include image and video processing, wireless networks, information security, image and video steganography, video summarization, diagnostic hysteroscopy, wireless capsule endoscopy, deep learning, computer vision, and CCTV video analysis. He has authored over 40 papers in peer-reviewed international journals and conferences, such as IEEE Transactions on Industrial Informatics, Neurocomputing, Future Generation Computer Systems, the IEEE Access, the Journal of Medical Systems, Biomedical Signal Processing and Control, Multimedia Tools and Applications, Pervasive and Mobile Computing, SpringerPlus, the KSII Transactions on Internet and Information Systems, MITA 2015, PlatCon 2016, FIT 2016, and ICNGC 2017.

Jamil Ahmad received his BCS degree in Computer Science from the University of Peshawar, Pakistan in 2008 with distinction. He received his Masters degree in 2014 with specialization in Image Processing from Islamia College, Peshawar, Pakistan. He is also a regular faculty member in the Department of Computer Science, Islamia College Peshawar. Currently, he is pursuing Ph.D. degree in Sejong University, Seoul, Korea. His research interests include deep learning, medical image analysis, content-based multimedia retrieval, and computer vision. He has published several journal articles in these areas in reputed journals including Journal of Real-Time Image Processing, Multimedia Tools and Applications, Journal of Visual Communication and Image Representation, PLOS One, Journal of Medical Systems, Computers and Electrical Engineering, SpringerPlus, Journal of Sensors, and KSII Transactions on Internet and Information Systems. He is also an active reviewer for IET Image Processing, Engineering Applications of Artificial Intelligence, KSII Transactions on Internet and Information Systems, Multimedia Tools and Applications, IEEE Transactions on Image Processing, and IEEE Transactions on Cybernetics. He is a student member of the IEEE.

Sung Wook Baik received the B.S degree in computer science from Seoul National University, Seoul, Korea, in 1987, the M.S. degree in computer science from Northern Illinois University, Dekalb, in 1992, and the Ph.D. degree in information technology engineering from George Mason University, Fairfax, VA, in 1999. He worked at Datamat Systems Research Inc. as a senior scientist of the Intelligent Systems Group from 1997 to 2002. In 2002, he joined the faculty of the College of Electronics and Information Engineering, Sejong University, Seoul, Korea, where he is currently a Full Professor and Dean of Digital Contents. He is also the head of Intelligent Media Laboratory (IM Lab) at Sejong University. His research interests include computer vision, multimedia, pattern recognition, machine learning, data mining, virtual reality, and computer games. He is a professional member of the IEEE.