



Machine learning based approach for multimedia surveillance during fire emergencies

Faisal Saeed¹ · Anand Paul¹  · Won Hwa Hong² · Hyuncheol Seo²

Received: 13 July 2018 / Revised: 28 December 2018 / Accepted: 25 March 2019 /
Published online: 6 June 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

Video based surveillance of manmade disasters such as fire has become very hot topic in research and it is playing an important role in the development of smart environment. The disasters like fire cause many economic and social damages. We can prevent these damages by early detection of the fire. The current advancement in embedded processing have permitted the detection of fire using vision-based i.e. Convolutional Neural Networks (CNNs) for the surveillance. Therefore, we proposed a method using machine learning techniques for Multimedia Surveillance during fire emergencies. Our proposed model has two main deep neural networks models. Firstly, we used a hybrid model made of Adaboost and many Multilayer perceptron (MLP) neural networks. The purpose of hybrid Adaboost-MLP model is to predict fire efficiently. This model used different sensors data like smoke, heat, and gas for training. After predicting the fire, we proposed a CNN model to detect the fire immediately. These results show that our trained model has near 91% fire detection accuracy. We can the false positive results are quite low. These results can be improved more by further training.

Keywords Machine learning · Multimedia surveillance · CNN · Fire detection

✉ Anand Paul
paul.editor@gmail.com; a.rehman.iiui@gmail.com

Faisal Saeed
bscsfaisal821@gmail.com

Won Hwa Hong
hongwh@knu.ac.kr

Hyuncheol Seo
notsools@gmail.com

¹ Department of Computer Science and Engineering, Kyungpook National University, Daegu 702-701, Korea

² School of Architectural, Civil, Environmental and Energy Engineering, Kyungpook National University, Daegu 702-701, Korea

1 Introduction

These days Multimedia based surveillance becomes more vital as there are lots of incidents become unnoticed even with CCTV. The November 2015 terror attacks in Paris threw the social media community into a frenzy [26]. A fire hose of information, opinions, and news streamed through Twitter, Facebook, and other networks. Some of it real, some of it false, and a lot of miscommunication. In fact, social media has made disaster response more complicated, yet often more effective.

Well, in any disaster, especially one producing casualties, there are two things that people want to know: what is going on, and are my friends and family safe? Before the internet was widespread, those questions were answered by legacy news sources, like television and radio. The whereabouts of loved ones would likely be dependent on a series of phone calls. But today, the game has completely changed. As soon as a major event occurs, a flood of information from hundreds, if not thousands of sources springs up, readily available to anyone with internet access. It has even transformed the nature of journalism, with citizen journalists sometimes providing the most up to date and accurate information, while traditional reporters scramble for sources. But with this overwhelming amount of information, there is a glaring problem i.e. which person we can trusted.

This is where another major factor comes in i.e. crowdsourcing. While people post information, there are educated groups fact checking, referencing, and supporting those ideas. Companies like Grasswire can collect huge numbers of tweets and photos, while following up on stories and checking with experts, all in real time. The faked photo of the Sikh man was almost immediately flagged as “fake” by Grasswire, preventing it from gaining as much traction.

The internet’s ability to regulate itself can even lead to even stronger journalistic standards than in staffed legacy news networks. But at the same time, crowdsourcing information isn’t always the best idea. During the hunt for the 2013 Boston Bombers, the crowdsourcing media website, Reddit, attempted to figure out the bombers’ identity. A sort of witch-hunt based on weak evidence led to multiple innocent people being targeted in the media and online, despite having nothing to do with the case at all.

Clearly, is it important that there be some editorial standard and organization, or a crowd can run amok with the wrong information. But social media has also made it fast and easy to reassure everybody – right – Social Media [13]. Instead of having to create a phone tree, we can reassure anyone interested in our wellbeing, and deal with the disaster at hand. Public officials are also able to reach as many people as possible by providing directions and updates online, thereby potentially saving lives. Despite some drawbacks to having a barrage of information, social media has made disaster response more effective, and under the right organization, more accurate. As technology changes the way we interact with the world, it is important to continue regulating and adapting to new tools about the power of one social media platform [16].

Apparently, people in the neighbourhood took out their phones, took pictures of the fire, and posted them on Facebook, rather than calling the fire department. It wasn’t until 15 min after the fire started that the fire department was alerted. Now the question of whether to post it first on the social media or to alert the relevant authority becomes a matter of personal choice.

In 2014 gas explosion occurred in two apartment building in East Harlem neighbourhood of Manhattan in NYC Fig. 1. People were taking selfies with the burning building as their background [34]. NYC fire department responded to the scene after the explosion within two minutes. This was a heart of the city scenario, what if such tragedy happens in a remote area, the authorities will not come to know such things unless someone alert them right away.



Fig. 1 Aerial view of the explosion

It's hardly news that the Internet has turned much of our world upside down, it is true for how we do communication when an event occurs, is like we book our travel, social media is creating new crises new ways to respond new channels to exploit but while most everyone in public relations and crisis communications understands that these changes are big.

We would like to suggest that they are much bigger than most realize in fact event communications in the social media era is the same as it was before, let's look at how communication used to be done in the old days. So to speak an event occurs we picture it as an explosion because crisis by nature tend to be sudden explosive events with major consequences when an event like this happens the response team is activated if we use the incident command system the various sections are organized and the public information officers (PIO) identified and gets to work as head of the communications function he or she will refer to him her as the PIO gathers up information from the responders drafts the press release gets the information approved by command, answers media questions hands out the release of the waiting media then the PIO sets up the press conference which is the main way to get response information to the media.

Recently, the various technologically advanced ways of fire detection have been introduced, as most of the time the fire breakout has been found to cause severe economic damage as well as loss of lives. Fire is easily spreadable to its surrounding carrying strong heat and stopping the people from escape. When the fire breakout spreads to the surrounding it can becomes more destructive resulting in higher casualties and death. Fire accounts for huge economic lose and death. However, early fire detection is more effective to astound the economic lose and to save the human lives. The fire detection at very initial stage is necessary to escape the fireplace and to drench the fire source properly.

The most suitable way to detect fire at early stage and to avoid losses is fire alarm system installation. Fire alarm system is equipped with various interconnected devices which are working together to detect fire and give fire alerts to the surrounding people via connected video and audio appliances regarding fire breakout and emergency exit. The various detection devices such as heat, smoke and gas detectors initially detect their respective happenings followed by auto-alarm activation. The bells, mountable sounders and horns are fitted in the alarm system to give proper alert.

Fire carries various features such as light, heat, smoke and many others which are detectable by different sensors. These days large amount of fire data exist on the cloud in the form of smoke and heat sensors data as well fire images data (smoke images, fire images). We can use these datasets in fire detection as well as we can use data uploaded on social media like fire images etc. This is the era of artificial intelligence and by using machine learning methods we can predict the fire as well detect it. To provide the efficient solution for above mentioned problem, we proposed a strategy by using two machine learning methods. Our first proposed model is hybrid neural network model which is made up of Adaboost and many MLP (Multi-layer Perceptron) neural networks. The main aim of using hybrid Adaboost-MLP model is to predict fire proficiently. This model utilized distinctive sensors information like smoke, heat, and gas for training. After predicting, we used a deep neural network for detection of fire in the videos or images. We used Convolutional Neural network for detection. Proposed CNN model is trained on available data on cloud as well as data gathered from social media.

Rest of the paper is organized as follows. Section 2 is describing our motivation. Section 3 is overview of proposed work and its subparts are describing the detail of each module. Section 4 is detail description of Experiments and results. Finally, at the end we concluded our work.

2 Motivation

Since much of the initial steps related to the fire lies with the social media, we would like to collect information from the social media, and the sensor data (that are available) in case of a wild fire, sensors inside the forest (if any), and the satellite images. Satellite image may not clearly reveal exact details of the fire, but it can easily identify the existence of fire. Once this is established then we can gather other data both from the social media and from sensor data from the sensors placed inside the forest or in the research stations near by the fire. And use ML technique to find required solution.

3 Proposed work

Early fire detection for surveillance of industries, homes, forests, nuclear power plants and other different public areas can be fruitful regarding ecological, economic, and social damage. However, plenty of work done on early fire detection, but still it is a challenging problem to come with a more efficient method. Thus, the problem demands to come up with an algorithm that can achieve better accuracy and can minimize the false alarms [34]. A fire has features like light, smoke, heat, gas and many others. As this is the era of artificial intelligence and many others machine-learning based algorithms, so, to provide a better solution for the problem mentioned above, we proposed a method using machine learning techniques. Our proposed model has two main deep neural networks models. Firstly, we used a hybrid model made of Adaboost and many MLP neural networks. The purpose of hybrid Adaboost-MLP model is to predict fire efficiently. This model used different sensors data like smoke, heat, and gas for training. After predicting the fire, we proposed a CNN model to detect the fire immediately. Our CNN model used two types of inputs, i.e. (i) the input pictures and videos by nearest selected surveillance camera node (ii) or the data posted on social media by humans of that region. The detailed description of each part is as follows subparts.

3.1 Adaboost-MLP hybrid model

Our proposed hybrid model for fire prediction is demonstrated in Fig. 2. It shows, first there is Adaboost algorithm to make better predictions with MLP neural networks, many MLP neural networks as weak predictors for prediction of fire and final all the forecasted results from MLP neural networks are summarized using weight provided by AdaBoost. A detailed description of these parts is in following subparts.

3.1.1 Adaboost algorithm

Adaboost abbreviated as Adaptive Boosting is a powerful machine learning algorithm. It is formulated by Scientist Freund [14]. Commonly, it is used as a conjunction of many other weak artificial neural networks. It combines the output of other classifiers ('weak learners') into a weighted sum, and then it characterized as the final output (boosting). Firstly, it was designed for feature classification and regression, but because of its great efficiency regarding classification, it has been used in many image processing applications [17, 30, 32]. The AdaBoost algorithm has sensitivity against noisy data and outliers. In our proposed scheme, we used Adaboost algorithm to boost the output of many MLP neural networks for better and efficient fire prediction.

3.1.2 MLP neural networks

MLP, short for multilayer perceptron, basically it is a feed-forward artificial neural network [31]. It has excellent performance for nonlinear matching and has great generalization abilities, for that reason, it has been applied in many signal processing works. A simple MLP neural network consists of three nodes, i.e. one input node, one output node and one hidden node. Except for input layer, every neuron (each node is a neuron) uses nonlinear activation function (commonly hyperbolic tangent or logistic function activation functions). For training, it utilizes backpropagation technique. We used different MLP neural networks as a weak prediction model for fire.

3.2 Convolutional neural network architecture

Deep CNN's have presented the powerful applications in computer vision projects like classification, renewal of images [5, 8, 20, 38], object detection image segmentation and, object detection [35, 36], Image segmentation and localization [21, 23]. The CNN model was

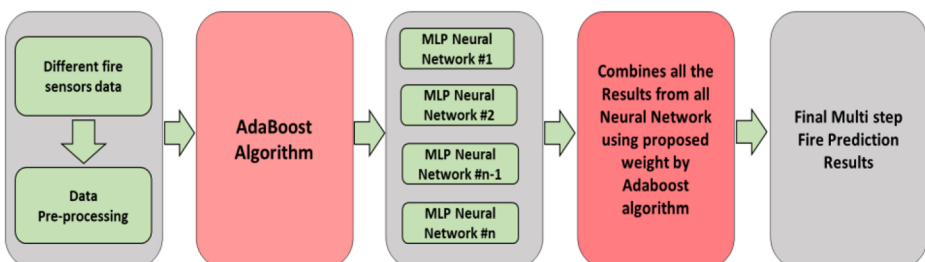


Fig. 2 Proposed AdaBoost-MLP model

firstly derived by Fukushima [15], he devised an architecture of hierarchical neural network after the marvellous research of Hubel [19]. After that, a CNN model was designed by Lecun [12] for the classification of digits. D. Ciresan [11] proposed a CNN model for image classification.. The popularity of CNN's among various computer applications is associated with their hierarchical architecture. The basic idea behind the deep convolutional neural network is the partition of the problem into subparts, continuously, until the final solution [1–3, 7, 27]. Other benefits of using CNN are, networks can be modified very easily, training datasets and other parameters can be updated without any complexity.

Normally, CNN comprises with different types of processing layers including Convolutional, Max-polling and many FC layers. These layers are connected in such manner that output of one layer is turned into the input of next connected layer. Convolutional layers are the core blocks of the CNN model. These layers are made of a rectangular frame of neurons. These neurons have a small amenable field which can be extended through the input volume. So, the convolutional layer is the convolution of the preceding layer. But sometimes more convolutional layers make the network more complex especially for recognition systems. After convolutional layers, sometimes there may be max-pooling layers. The primary function of the pooling layers is subsampling of their inputs. After several convolutional and max-pooling layers, fully connected layers come for high-level reasoning in the neural network. For real-time applications, these layers can be connected and deformed together many times for expected results.

In our proposed CNN model, we made three convolutional layers with three max-pooling layers. Our proposed model is illustrated in the Fig. 3. The input for CNN model is taken from the stored images or images that posted by humans on social media using the internet. Then these images are resized to $224 \times 224 \times 3$. Resized images then passed through the first convolutional layer. On that layer, 96 kernels of size 11×11 with a stride of 4 are applied and produced 96 features. These features are then sent to max-pooling layer for reducing complexity, and the size of the feature maps is reduced by the factor of 2. After that, second convolutional layer came with 256 kernels of size 5×5 came followed by similar max-pooling layer. Then the last convolutional layer having the same 256 kernels and last pooling layer like previous. At the end of the model, there exists three final and fully connected layers which classify the output of previous layers into results. In the final layers we used SoftMax

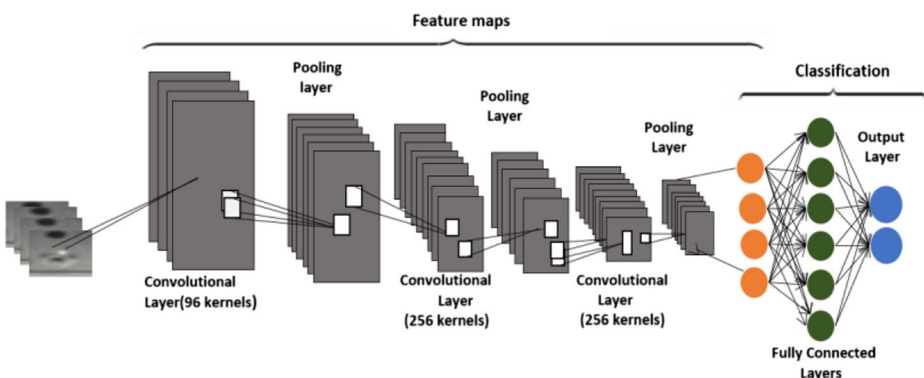


Fig. 3 Proposed CNN mode

classifier to classify the final output. The final output will be classified as “Fire” and “Normal”. SoftMax discriminant classifier do it by providing weighted distance between training and testing samples from that particular data class. We used ReLu activation function with coefficient a = 13 for convolutional and fully connected layers similarly as [4, 37] shown in Fig. 4.

For calculating precision p, recall r, F_{measure}, and accuracy, we used the following equations. Precision defines a relevant object as fractions of the retrieved objects, but recall is a representation of relevant objects of the complete set. We used F_{measure} to measure the efficiency of precision and recall. We calculated these by using following equations.

$$\begin{aligned}
 p &= \frac{T^p}{T^p + F^p} \\
 r &= \frac{T^p}{T^p + F^n} \\
 F_Measure &= 2 \times \left(\frac{p \times r}{p + r} \right) \\
 Accuracy &= \frac{T^p + T^n}{T^p + T^n + F^p + F^n}
 \end{aligned}$$

Where:

T^p (True positive) = (Fire images predicted as fire images).

T^n (True negative) = (Normal images predicted as Normal images).

F^p (False positive) = (Normal images predicted as fire images).

F^n (False negative) = (Fire images predicted as Normal images).

To measure the true positive rate and false positive rate following are used.

$$\begin{aligned}
 T^p \text{ rate} &= T^p \times (T^p + F^n) \\
 F^p \text{ rate} &= 1 - T^n \times (T^n + F^p)
 \end{aligned}$$

4 Experiments and results

The purpose of our work is to predict fire before happening and classification (if a fire happens) whether the input image contains the fire. For that, the both of our ANN models are trained using a collection of labelled data (sensors data and images).

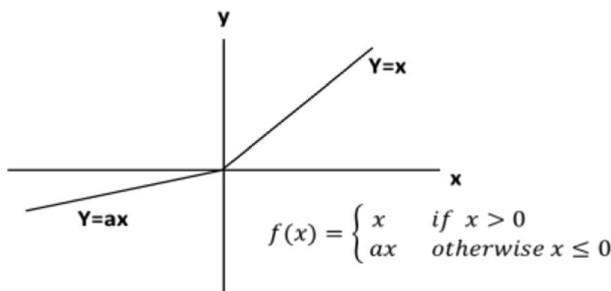


Fig. 4 Leaky Relu

Furthermore, after prediction, we want to locate the position of the fire. The training of each model is described below.

4.1 Training of Adaboost-MLP model

The outcomes of the fires are many like light smoke, many gases, heat and many other things. These parameters can detect it. We used CO₂ gas data and heat data as a training data for both Adaboost and MLP models. The complete process of MLP network training is illustrated in Fig. 5. From the figure, we can see that: we chose 900 samples (each sample is a vector which contain many sensors values) of original fire sensor pre-processed data and then we divided this data into three parts, i.e. a set of training data, validation data and a set of testing data. Each sample is in the form of vector and each vector contains many sensors values. Total data that we used is almost 3.45 GB. The first set, i.e. 1st to 700th samples are used to train the model, 2nd dataset of hundred samples (701–800) is used to validate the model and rest 801st to 900th are used to validate the data.

In training set, we divided more into eight different input streams. After training the model, to validate our MLP model we used 100 samples (100 vectors), i.e. 801st to 900th. In our model, the number of input neurons are 8; hidden neurons are 16 and only one output neuron. The learning rate was 0.01 and the training was done using 100 iterations.

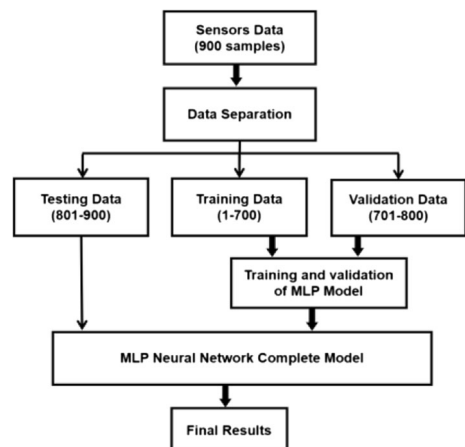
The basic operation of an Adaboost algorithm that is commonly used [22] are as follow:

To calculate weights $\{D_t(i)\}$ of the of fire data series the following equation can be used:

$$D_t(i) = \frac{1}{n} \text{ Where } i = 1, 2, 3, \dots, n-1, n \quad t = 1, 2, 3, \dots, T$$

Where n is the sample size of fire data series and T belongs to our MLP based forecasting model. After selecting the weights, now to calculate the forecasting error ε_i (the error occurs during MLP forecasting for every input channel) our model used equation (1) and then by using equation (2) to calculate the sum of all error ε_i and to calculate the weights for MLP predictor we used equation (3). We used equation (4) to update the sampling weight of input

Fig. 5 Flow chart of MLP neural network training process



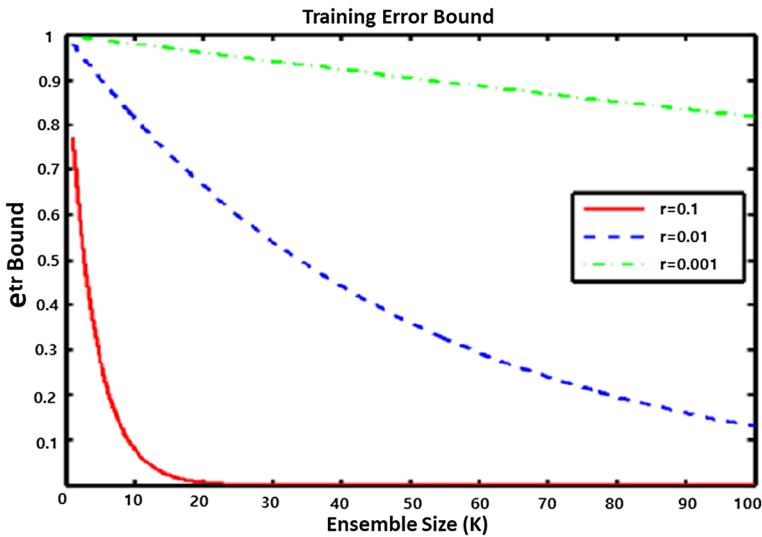


Fig. 6 Training error of adaboost MLP model

streams. In equation (4) z_t is the normalize impact. This procedure repeated many times until all MLP models gives output. Finally, summarized all the MLP predictions using Adaboost framework to make final and efficient prediction. Figure 6 is the training error of our model. The formulae are as follow:

$$\varepsilon_i = \frac{|x_i - \bar{x}_i|}{x_i} \text{ where } i = 1, 2, 3, \dots, n \tag{1}$$

$$\varepsilon_t = \frac{1}{n} \sum_{i=1}^n \varepsilon_i \tag{2}$$

$$W_t = \frac{1}{2} \ln \left[\frac{1 - \varepsilon_t}{\varepsilon_t} \right] \tag{3}$$

$$\begin{cases} Dt(i) = \frac{D_{t-1}(i) \beta^{-\varepsilon_i}}{z_t} \\ \beta_t = \frac{\varepsilon_t}{1 - \varepsilon_t} \end{cases} \tag{4}$$

Table 1 Dataset description

Sr.	Name	Videos	Images
1	Foggia fire dataset	31	40,456
2	Chino Smoke dataset	13	18,213
3	Social Media dataset	06	6521
4	Real time Collected Videos	07	9504

Table 2 Training and testing dataset of smoke and fire

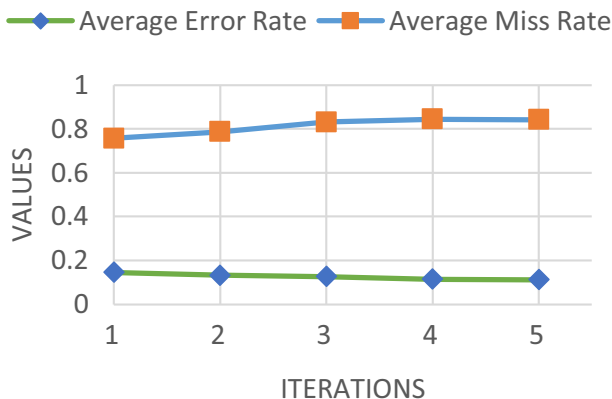
	Images	Training Images	Testing Images	Validation Images
Smoke	18,213	7624	9335	1254
Fire	27,456	16,532	6353	4571
Normal	29,050	18,214	5815	5021

4.2 Training of CNN model

As usually, CNN's work with two types of neuron layers, i.e. a first convolutional layer which is built up with non-linear stimulation function and second layer consist of non-linear subsampling algorithm. Mixing of these variants of layers can be fruitful for building applications like object detection. As mentioned in the proposed scheme we used ReLu activation function. The reason behind chosen of ReLu as activation function over Sigmoid and Hyper tangent is high accuracy and quick training. The basic formula of Leaky Rectified Linear Unit is following [37]:

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ ax & \text{otherwise } x \leq 0 \end{cases}$$

The goal behind classification using CNN model is to classify the image content into fire or normal. To achieve this goal, we used 74,719 RGB labelled images from different datasets Foggia's video dataset [25], Shino's dataset [10] and some other datasets [24] for the training of our model. The detailed description is shown in Table 1. We used 45% data for training and rest 55% for testing purpose. Table 2 is showing the data d partitioned for training and testing. We trained our model on a computer having specification Intel(R) Core(TM) i5-3570 CPU @ 3.40GHz and 3.80GHz RAM 32GB and graphics card Nvidia Titan XP. We used Stochastic gradient descent and backpropagation algorithm (to set the best weight as first we chose weight randomly). We calculated average learning rate. Initially, it was 0.01, but after every five epochs, it decreases by factor 0.95. We trained our model in 6 iterations. We calculated average

**Fig. 7** Overview of miss rate and error rate during training of CNN Model

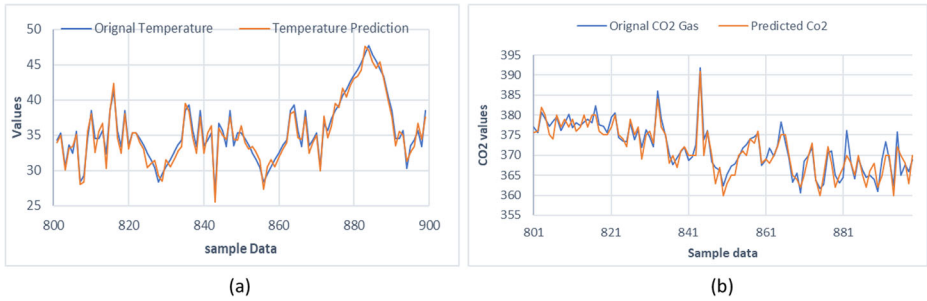


Fig. 8 Predicted result proposed Adaboost-MLP model

error and miss rate for each iteration. The graph in Fig. 7 is showing variation in error rate and miss rate. After each iteration error rate was decreasing and miss rate was increasing. Miss rate was in 80% to 90%, and it showed the accuracy with each iteration and error rate decreased with factor 0.1. To avoid overfitting in fully connected layers, we used dropout of 0.5.

4.3 Results of proposed model

The results for both deep learning models are in the following subsection.

4.3.1 Results of Adaboost-MLP model

For testing the performance of our Adaboost-MLP model, we used a sample data of size 801–900 for temperature and CO2 gas. Fig. 8 (a) and (b) shows the forecasted result for temperature and CO2 gas respectively. From figures, we can see our hybrid Adaboost-MLP model have better predictions of temperature and gas.

4.3.2 Results of CNN based fire detection

We did our experiments on two datasets: Foggia video fire dataset [25] and Chino smoke dataset [10]. The fire dataset has 31 fire videos with both indoor and outdoor

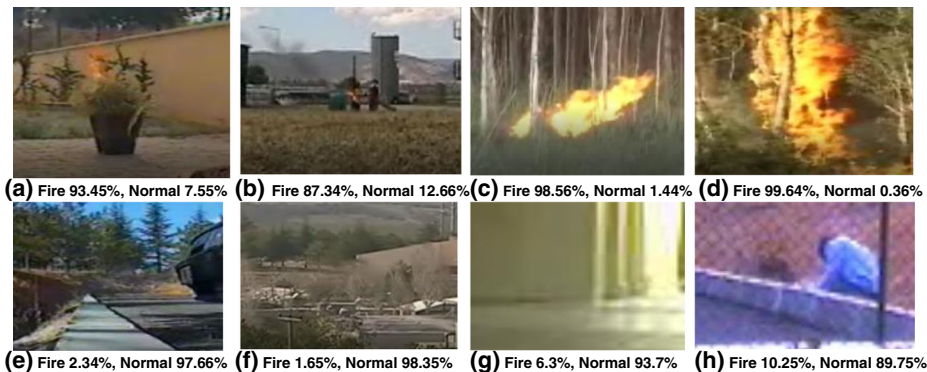


Fig. 9 CNN Model results for different images



Fig. 10 Feature map of original image

environments, of which 17 videos are of fire, and the rest videos consist no fire. The second data set is of smoke in which some videos are of smoke, and others are without smoke. We select these both datasets because these have videos captured in different scenes in indoor and outdoor environments. Among these videos, some videos contain fire-like objects or situation like fire but not actual fire, and it is very difficult to predict or classify. Similarly, the techniques which based on motion detection may fail to classify a scene like fire or smoke on mountains, cloud, or fog. Due to these reasons, the selected datasets are more challenging. Some of the results for these datasets are shown in Fig. 9.

We divided our CNN model into parts. First part is for mapping the features and second one for high reasoning classification. Our main aim is to characterize a fire in video frame. Other models use sliding window for the classification of images. We used sliding window technique as well a different method instead of sliding a window of size 64×64 for to detect the fire. We tried to detect the fire window with last feature map. First part of our CNN model which includes three convolutional and three max-polling layers can evaluate last feature map. We evaluate last feature map

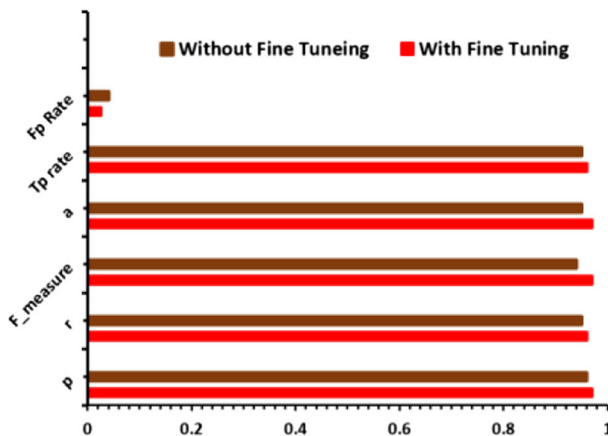


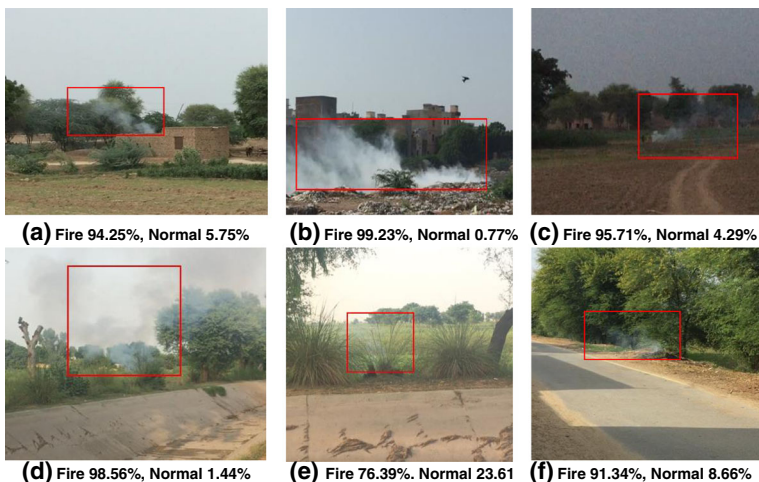
Fig. 11 Graph for calculated result of P, r, F-measure, T_p rate and F_p rate

Table 3 Comparison of proposed technique with existing methods

	Proposed Work	[33]	[18]	[6]	[9]
Accuracy(%)	91.45	92.86	90.32	83.87	87.10
Prediction	Yes	No	No	No	No
Real time Testing	Yes	No	No	No	No
Social Media dataset	Yes	No	No	No	No
False Positive(%)	9.05	13.33	5.88	29.41	11.76
False Negative(%)	2.05	0	14.29	0	14.29

by using first part of CNN model. To detect the smoke and fire in video, a sliding window having a size of 12×12 which is created from 64×64 RGB image is applied on last feature map in Fig. 10. We used GPU and a tensor $12 \times 12 \times 1 \times N$ (where is the number of windows) of last feature map to speed up our predictions.

We calculated results for precision p , recall r , F -measure, accuracy a , true positive T_p and false positive F_p are shown in Fig. 11. These results show that our trained model has near 91% fire detection accuracy. We can the false positive results are quite low. These results can be improved more by further training. The overall performance of with our both model is very efficient. In the future this system can be extended to reflect more smarter system down to individuals who can contribute in detecting fire and sending intimation to the fire-station [28, 29], an efficient multimedia system for retrieving fire detection scheme also be employed in the fire-station as they can directly detect fire from the CCTV cameras installed in the forest vicinity. Table 3 is the comparison of our proposed model with other existing models. We also tested our model for real time dataset. We made 8 different video which contain smoke and fires. We tested these videos using our own model. Figure 12 is the tested result of these dataset.

**Fig. 12** Real time testing result

Acknowledgements This study was also supported by the National Research Foundation of Korea (NRF) grant funded by the Korean Government (NRF-2017R1C1B5017464). This study was also supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIP) (NRF-2016R1A2A1A05005459).

References

1. Abdel-Basset M, Manogaran G, Abdel-Fatah L, Mirjalili S (2018) An improved nature inspired meta-heuristic algorithm for 1-D bin packing problems. *Pers Ubiquitous Comput* 1117–1132. doi:<https://doi.org/10.1007/s00779-018-1132-7>
2. Abdel-Basset M, Manogaran G, Mohamed M, Rushdy E (2018) Internet of things in smart education environment: supportive framework in the decision-making process. *Concurr Comput*. <https://doi.org/10.1002/cpe.4515>
3. Abdul R, Paul A, Gul MJ et al (2018) Exploiting small world problems in a SIoT environment. *Energies* 11: 2089. <https://doi.org/10.3390/en11082089>
4. Andrew LM, Awni YH, Andrew YN (2013) Rectifier nonlinearities improve neural network acoustic models. *Proc 30 th Int Conf Mach Learn Atlanta, Georgia*: 1–6
5. Anwar S, Hwang K, Sung W (2015) Fixed point optimization of deep convolutional neural networks for object recognition. *ICASSP, IEEE international conference on acoustics, speech and signal processing - proceedings*. IEEE, Brisbane, QLD, Australia: 1131–1135
6. Çelik T, Demirel H (2009) Fire detection in video sequences using a generic color model. *Fire Saf J* 44:147–158. <https://doi.org/10.1016/j.firesaf.2008.05.005>
7. Çelik T, Özkarmanlı H, Demirel H (2007) Fire and smoke detection without sensors: image processing based approach. *15th European Signal Processing Conference (EUSIPCO)*. IEEE, Poznan, Poland: 1794–1798
8. Chan TH, Jia K, Gao S et al (2015) PCANet: a simple deep learning baseline for image classification? *IEEE Trans Image Process* 24:5017–5032. <https://doi.org/10.1109/TIP.2015.2475625>
9. Chen TC, Wu P, Chiou Y (2004) An early fire-detection method based on image processing: 1707–1710
10. Chino DYT, Avalhais LPS, Rodrigues Jr. JF, Traina AJM (2015) BoWFire: detection of fire in still images by integrating pixel color and texture analysis. *28th SIBGRAPI conference on graphics, patterns and images*. IEEE, Salvador, Brazil: 95–102
11. Ciresan DC, Meier U, Masci J, et al (2011) Flexible, high performance convolutional neural networks for image classification. In: *proceedings of the twenty-second international joint conference on artificial intelligence*. IJCAI, Barcelona, Catalonia, Spain: 1237–1242
12. Cun YL, Boser B, Denker JS, et al (1989) Handwritten digit recognition with a Back-propagation network. In: David T (ed) *Advances in neural information processing systems 2*. Morgan Kaufmann: 396–404
13. Fraustino JD, Liu B, Yan J (2012) Social media use during disasters: a review of the Knowledge Base and gaps. *Dep Homel Secur Sci Technol Cent Excell* 1–39
14. Freund Y (1995) Boosting a weak learning algorithm by majority. *Inf Comput*. <https://doi.org/10.1006/inco.1995.1136>
15. Fukushima K (1980) Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol Cybern* 36:193–202. <https://doi.org/10.1007/BF00344251>
16. (2018) Gas explosion in harlem blows up 2 buildings, 2 women dead & 17 injured. In: WIKIPEDIA. <http://wordondastreet.com/gas-explosion-harlem-blows-2-buildings-2-women-dead-17-injured/>. Accessed 16 Dec 2018
17. Guo L, Ge PS, Zhang MH et al (2012) Pedestrian detection for intelligent transportation systems combining AdaBoost algorithm and support vector machine. *Expert Syst Appl* 39:4274–4286. <https://doi.org/10.1016/j.eswa.2011.09.106>
18. Habiboğlu YH, Günay O, Çetin AE (2012) Covariance matrix-based fire and flame detection method in video. *Mach Vis Appl* 23:1103–1113. <https://doi.org/10.1007/s00138-011-0369-1>
19. Hubel DH, Wiesel TN (1977) Functional architecture of macaque monkey visual cortex. *Proc R Soc B Biol Sci* 198:1–59. <https://doi.org/10.1098/rspb.1977.0085>
20. Jiang B, Yang J, Lv Z et al (2017) Internet cross-media retrieval based on deep learning. *J Vis Commun Image Represent* 48:356–366. <https://doi.org/10.1016/j.jvcir.2017.02.011>
21. Li W, Zhao R, Xiao T, Wang X (2014) DeepReID: Deep Filter Pairing Neural Network for Person Re-Identification. *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, Columbus, OH, USA: 152–159

22. Liu H, Tian HQ, Li YF, Zhang L (2015) Comparison of four Adaboost algorithm based artificial neural networks in wind speed predictions. *Energy Convers Manag* 92:67–81. <https://doi.org/10.1016/j.enconman.2014.12.053>
23. Luo P, Tian Y, Wang X, Tang X (2014) Switchable Deep Network for Pedestrian Detection. *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, Columbus, OH, USA: 899–905
24. Mueller M, Karasev P, Kolesov I, Tannenbaum A (2013) Optical flow estimation for flame detection in videos. *IEEE Trans Image Process* 22:2786–2797. <https://doi.org/10.1109/TIP.2013.2258353>
25. Narwani S, Selva Kumar A (2016) Real-time fire detection for video surveillance applications using a combination of experts based on color, shape and motion. *IEEE Trans Circ Syst Video Technol* 25:1545–1556. <https://doi.org/10.1109/TCSVT.2015.2392531>
26. (2018) November 2015 Paris attacks. In: WIKIPEDIA. https://en.wikipedia.org/wiki/November_2015_Paris_attacks. Accessed 16 Dec 2018
27. Paul A, Jiang YC, Wang JF, Yang JF (2012) Parallel reconfigurable computing-based mapping algorithm for motion estimation in advanced video coding. *ACM Transactions on Embedded Computing Systems (TECS)* 11(S2):40
28. Paul A, Chen B-W, Bharanitharan K, Wang J-F (2013) Video search and indexing with reinforcement agent for interactive multimedia services. *ACM Trans Embed Comput Syst* 12:1–16. <https://doi.org/10.1145/2423636.2423643>
29. Paul A, Ahmad A, Rathore MM, Jabbar S (2016) Smartbuddy: defining human behaviors using big data analytics in social internet of things. *IEEE Wirel Commun* 23:68–74. <https://doi.org/10.1109/MWC.2016.7721744>
30. Paul A, Daniel A, Ahmad A, Rho S (2017) Cooperative cognitive intelligence for internet of vehicles. *IEEE Syst J* 11(3):1249–1258. <https://doi.org/10.1109/JSYST.2015.2411856>
31. Rafei M, Sorkhabi SE, Mosavi MR (2014) Multi-objective optimization by means of multi-dimensional MLP neural networks. *Neural Netw World* 24:31–56. <https://doi.org/10.14311/NNW.2014.24.002>
32. Rehman A, Rathore MM, Paul A et al (2018) Vehicular traffic optimisation and even distribution using ant colony in smart city environment. *IET Intell Transp Syst*:1–8. <https://doi.org/10.1049/iet-its.2017.0308>
33. Rosario DL, Antonio G, Alessia S, Mario V (2014) Improving fire detection reliability by a combination of video analytics. *International conference image analysis and recognition*. Springer, Cham: 477–478
34. Saeed F, Paul A, Rehman A et al (2018) IoT-based intelligent modeling of smart home environment for fire prevention and safety. *J Sens Actuator Netw* 7:11. <https://doi.org/10.3390/jsan7010011>
35. Vadim K, Maxime O, Minsu C, Laptev I (2016) ContextLocNet: context-aware deep network models for weakly supervised localization. *European conference on computer vision (ECCV)*: 350–365
36. Wenlu Z, Rongjian L, Houtao D et al (2015) Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. *Neuroimage* 108:214–224. <https://doi.org/10.1016/j.neuroimage.2014.12.061>
37. Xu B, Wang N, Chen T, Li M (2015) Empirical evaluation of rectified activations in convolutional network. doi:[https://doi.org/10.1016/S0266-8920\(98\)00020-4](https://doi.org/10.1016/S0266-8920(98)00020-4)
38. Yang J, Jiang B, Li B et al (2017) A fast image retrieval method designed for network big data. *IEEE Trans Ind Inform* 13:2350–2359. <https://doi.org/10.1109/TII.2017.2657545>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Faisal Saeed is currently working towards his Master-combined PhD program, his research interest includes, Fire-detection, Machine learning methods and fake-event detection using ML systems. He has already published his initial work in Sensors and Actuator journal, he is active member of CCMP lab and he has been merited as one of few selected student for full scholarship and fund for his PhD program.



Anand Paul, (SM)** received the Ph.D. degree in Electrical Engineering from the National Cheng Kung University, Tainan, Taiwan, in 2010. He is currently working as an Associate Professor in the School of Computer Science and Engineering, Kyungpook National University, South Korea. He is a delegate representing South Korea for M2M focus group and for MPEG. His research interests include Algorithm and Architecture Reconfigurable Embedded Computing. He is IEEE Senior member and has guest edited various international journals and he is also part of Editorial Team for Journal of Platform Technology, ACM Applied Computing review and Cyber-Physical Systems. He serves as a reviewer for various IEEE /IET/Springer and Elsevier journals. He is the track chair for Smart human computer interaction in ACM SAC 2015, 2014. He was the recipient of the Outstanding International Student Scholarship award in 2004–2010, the Best Paper Award in National Computer Symposium and in 2009, and International Conference on Soft computing and Network Security, India in 2015.



Won-Hwa Hong has more than 20 years experience in Architectural Engineering and Disaster Management more than 250 scientific publications and 8 patents. He received the Ph.D. degree in Architectural Engineering from Waseda University, Tokyo, Japan in 1994. He has taught at University of Waseda from 1995 to 1997, and is a professor in Department of Architecture and Architectural Engineering of Kyungpook National University since 1999. Won-Hwa Hong is currently working as a head of Disaster and Safety Management Institute funded by Ministry of Public Safety and Security, South Korea (MPSS) His research interests are energy demand and consumption in existing and future buildings, resilience technologies for the built environment, IoT Based Technologies for Smart City He serves as a reviewer for Various Springer and Elsevier Journals. Won-Hwa Hong is the Director at the Asia Institute of Urban Environment



Hyuncheol Seo is currently a fourth year PhD candidate under the supervision of professor Won-Hwa Hong in the School of Architectural, Civil, Environmental and Energy Engineering, Kyungpook National University. His main research interests are energy demand and consumption in existing and future buildings, especially effect of socio-cultural changes on energy demand of buildings and cities. He was the recipient of the Global Ph. D Fellowship Program in 2013–2016 biggest and most competitive nationwide education program for graduate students in South Korea, jointly sponsored by the Korean Ministry of Education and the Korean National Research Foundation (NRF). Hyuncheol Seo received the Contribution Award (Mahoroba Prize) from Asia Institute of Urban Environment Japan in 2012, the Best Paper Award in 11th International Conference of Asia Institute of Urban Environment in 2014, and the 9th International Workshop Series by HSST in 2016.