Development of an Early Fire Detection Technique Using a Passive Infrared Sensor and Deep Neural Networks

Karish Leo Britto Leo Xavier and Visakha K. Nanayakkara*, Department of Mechanical Engineering, Kingston University London, Friars Ave, London SW15 3DW, UK

Abstract. Early detection of fire is key to mitigate fire related damages. This paper presents a differential pyro-electric infrared (PIR) sensor and deep neural networks (DNNs) based method to detect fire in real-time. Since the PIR sensor is sensitive to sudden body motions and emits a continuous time-varying signal, experiments are carried out to collect human and fire motions using a PIR sensor. These signals are processed using one-dimensional continuous wavelet transform to perform feature extraction. The corresponding wavelet coefficients are converted into RGB spectrum images that are then used as inputs for a deep convolutional neural network. Various pre-trained DNN architectures are adopted to train and identify the collected data for background (no motion), human motion, and fire categories: small quasi-static and spreading fires. Experimental results show that the ShuffleNet architecture yields the highest prediction accuracy of 87.8%. Experimental results for the real-time strategy which works at a speed of 12 frames-per-second show 95.34% and 92.39% fire and human motion detection accuracy levels respectively.

1. Introduction

Early fire detection in residential or industrial buildings is very important to take appropriate actions to prevent human and property losses, physical and mental trauma. Yet, fires can be avoided or controlled if identified at the incipient stage. Fires happen due to various reasons such as negligence, electronic device malfunctioning, cooking, smoking, flammable liquid and gas leakages, arson, etc. The stages of a fire can be described as incipient (ignition), growth, fully developed, and decay [1]. Fire growth and the maximum heat release rates mainly depend on the building design features including flammable cladding and the types of combustibles (fire loads) such as interior decorations, furniture, etc. [2, 3]. According to one of the common building communication standards BACnet, fire safety is

---

*Correspondence should be addressed to: Visakha K. Nanayakkara, E-mail: V.Nanayakkara@kingston.ac.uk

given the highest priority in building automation [4]. With the increased complexity in building operations, it is important to improve sensitivity and accuracy of fire detection [5].

There are various fire sensing technologies developed to reduce fire accidents [1]. The commonly used point sensors like smoke, heat, and gas sensors have limited capabilities [6]. A fire has to burn for a while to reach the sensor and to exceed the set threshold point to initiate a warning alarm [7]. This gives plenty of room for a normal fire to spread rapidly. Due to the limited coverage area, conventional sensors are not efficient in large rooms and open spaces like backyards, car parks, forests, etc. In addition, their accuracy is affected by the external disturbances such as steam, smoking, light smoke from cooking, causing to trigger false alarms. According to fire and rescue services (FRSs) in England for the year ending December 2020, the false fire alarms amount to 42% of the incidents attended [8].

In order to address issues in conventional fire sensors, advanced smoke and fire detection systems use video fire detection techniques to recognise fires at their early stages [9]. Various colour analysis and machine learning techniques are used to extract and identify fire in recorded images and videos. Even though vision-based techniques are more accurate and capture a detailed view of the environment, complexity in fire spread and other interfering background events in practical settings affect their feature extraction capability [10]. It is also noted that fixing cameras in workplaces may affect employee performance by creating an environment of personal privacy issues [11]. Infrared (IR) imaging systems are effective to address issues like poor visible light or image blending with the background [1]. These systems are highly reliable and can act as surveillance videos to detect fires. Though they have many advantages, they also have certain disadvantages. Cameras and onboard computers pose cost barriers for a scalable solution. They give false alarms depending on the proximity to the sensor or any other interference such as lamp lights, hot air flow, workshop or industrial activities, that affect the sensitivity [12]. In order to address data loss and data distortion issues in IR flame detection, a robust fusion algorithm based on Radial Basis Function neural network with Takagi–Sugeno fuzzy model is proposed in [13]. Image processing using deep learning (DL) technology gives significantly high performance in fire detection [14]. However, there are challenges such as the need for a large number of training data collected at various settings and fire types, identification of simultaneous fires, illumination changes, and quantification of fire flame and smoke intensity [9].

This paper proposes a new technique to detect fires in enclosed spaces using a differential PIR sensor and DNNs. PIR sensors detect IR radiation emitted by objects. They are sophisticated, yet low cost and highly capable of detecting any abrupt changes in IR radiation from object movements within its viewing range [15–17]. Hence, PIR sensors are versatile to detect a flame spread. Moreover, differential detection in PIR sensors is useful to avoid false alarms caused by small but uniform IR changes in its viewing range such as indoor temperature or sunlight variations. This is due to the sensor arrangement in which the sensor is split into two halves and they are wired up together. Unless there is not sufficient increase in IR radiation to capture from one half, the output signals produced

from the two halves will cancel each other. Due to the relatively high viewing range of 7 m (when used with Fresnel lens) in an angle covering 110 degrees, it is suitable to monitor large rooms with high ceilings [18]. In addition, the PIR sensor approach requires low processing power compared with computer vision (surveillance) smoke or fire detection systems [11]. Unlike in security cameras, human privacy intrusion issues will not arise with PIR sensors because the sensor output is merely a digital or an analog signal [19].

Different types of real fire spreads and human motion data in an indoor setting are collected in this paper using a differential PIR sensor. These signals are pre-processed using wavelet transform [20]. The wavelet coefficients that correspond to the extracted fire features are fed as inputs to train a selected class of DNNs to classify fire and non-fire events i.e., fire categories and human motion, within the viewing range of the PIR sensor. Image pre-processing increases the accuracy of flame recognition rate in DNNs [21, 22]. Conventional machine learning approaches need a considerable amount of expertise to extract features, classify them, and predict a fire spread. Due to complex image feature extraction capability and availability of pre-trained CNNs like AlexNet, ResNet, GoogLeNet, etc., DL approach can be used to detect fire flames robustly [10, 12, 14, 21, 23]. In addition, compact variants of CNN architectures such as NasNet-A-Mobile and ShuffleNetV2 are used in [24] with modified layers to improve non-temporal fire and no-fire classification. The proposed method in [24] shows 95% full frame and 94.4% superpixel based fire detection. In our work, the adopted pre-trained CNN models are also modified to improve accuracy at the cost of training time. The CNN network architecture with the highest validation accuracy is used to demonstrate the real-time fire prediction.

In Sect. 2, the proposed methodology is presented under event categorisation, data acquisition, signal processing, deep learning frameworks, DNN architecture, and the real-time working strategy. Fire prediction results along with real time performance are presented in Sect. 3. Discussion and conclusion are given in Sects. 4 and 5.

2. Proposed Methodology

In this work, changes in IR radiation for fire and non-fire events are recorded using an HC-SR501 PIR motion sensor [25]. The experimental setup to collect data is shown in Fig. 1. All experiments are carried out following health and safety guidelines and regulations for fire experiments.

In the proposed technique, the differential PIR sensor is configured with an Arduino Uno micro-controller (ATmega328P). The data from the differential PIR sensor is a continuous time-varying analog signal. Its amplitude is proportional to the hot body size, speed of the motion, and distance from the sensor. The signal is processed using wavelet transform. The wavelet coefficients are converted to RGB (Red-Green-Blue coloured) spectra. Due to this colour segmentation process in RGB space, most likely flame regions can be identified and fed as compatible input images to pre-trained DNNs to classify the events happening within the
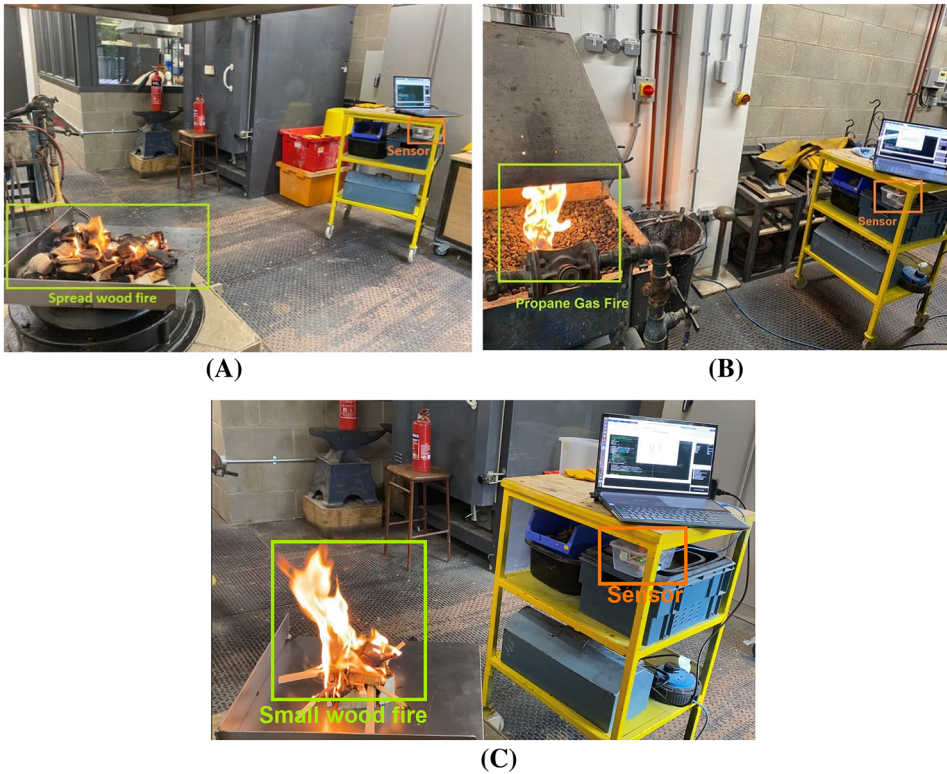
Figure 1. Experimental setup during data collection. Sub figures (A), (B), and (C) are the photos captured during the events of spreading wood fire, propane gas fire, and small wood fire respectively. The PIR sensor and the Arduino Uno is packed inside a transparent plastic box.

viewing range of the sensor. Figure 2 represents the proposed fire detection technique.

The rest of the section explains the key areas of this work under event categorisation, data acquisition, digital signal processing, data pre-processing, DNN architecture, and real-time performance strategy.

2.1. Events Categorisation

Since the proposed fire detection technique is intended for indoor settings and the differential PIR sensor is sensitive to movements, variables such as distance from the sensor, different human motion patterns, and moving speeds are also considered in this work. Therefore, PIR sensor data are collected for the following five different motion events to record their changes in IR radiation.
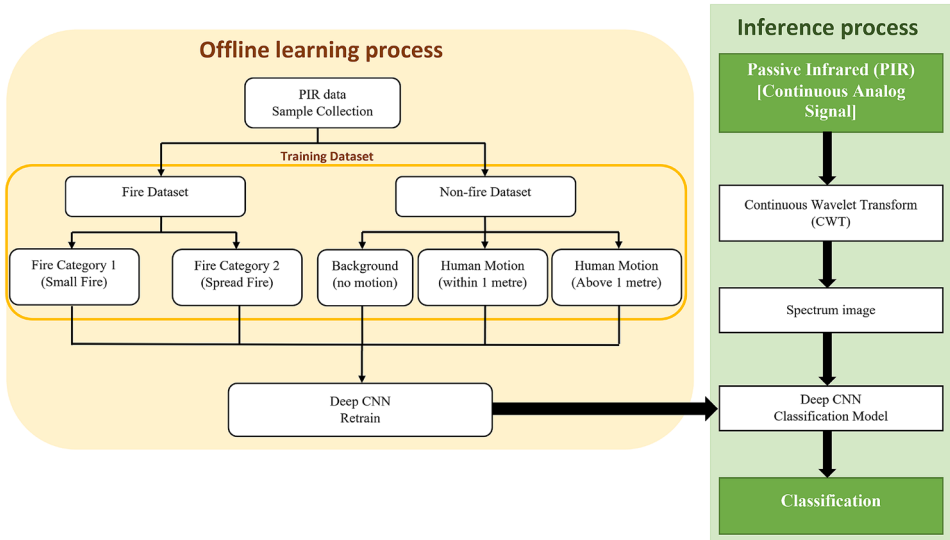
Figure 2. The fire detection technique adopted in this paper. The left side of the chart denotes offline learning including data collection, event categorisation, and the DNN training. The right side illustrates the inference process. The methodology of the inference process during real-time is shown in Fig. 6.

1. Background (no motion)
2. Human motion (within 1 m from the sensor)
3. Human motion (more than 1 m from the sensor)
4. Fire category 1 (small wood fire)
5. Fire category 2 (both spreading wood and propane gas fires)

When there is no motion in front of the sensor, it is categorized as the background event. The human motions within 1 m and beyond contain sensor signals for human motion within the ranges 0.2 m to 1 m and 1.1 m to 3 m respectively. Since the PIR sensor is sensitive to motion, the fluctuations of the output voltage are high or low for fast or slow motions respectively. Thus, both human slow walking motion and human fast walking motions are recorded. In addition, human movements on a swivel chair are also recorded. Finally, two fire category events are recorded (see Fig. 1). The fire category 1 has the recording of small wood fire between 0.5 m and 3 m from the sensor. The fire category 2 consists of larger or uncontrollable fire. Spreading wood fire and propane gas fires are recorded for this event. The area coverage of wood during the event of small wood fire and spreading wood fire are 10 cm^2–12 cm^2 and more than 12 cm^2 respectively.

2.2. Data Acquisition

The two pyro-electric sensing elements inside the PIR sensor generate positive or negative differential changes if a warm body enters or leaves the sensor detecting area which is bounded by two slots on the sensor. Hence the sensor produces high or low digital voltage signals depending on variations in IR radiation levels. Since the amplitude of these signals are very small, the amplified analog signals are extracted from the PIR sensor by tapping directly from its sensor circuit amplifier chip BISS0001. Thus, the amplified analog signals are recorded during the five events explained in Sect. 2.1. The output signals are shown in Figs. 3 and 4.

Since the focus is on fire, the flame flickering frequency should be considered for signal sampling. The flame flicker behaviour is a wideband frequency range covering 1 Hz–13 Hz [17]. Considering the sampling rule, the sampling rate is selected as 50 Hz, which is more than twice the flame flickering frequency. The analog signal from the PIR sensor is transferred to Arduino Uno which is then transferred to the computer for further signal processing. The Arduino Uno is configured using MATLAB [26] to perform signal processing using the trained DL network.

Figures 3 and 4 show that the samples of amplified PIR signals have different characteristics for the events considered in time domain. During the background event where no motion occurs, the signal amplitude varies from 0.58 V to 0.59 V. This range acts as the centre line around which the signal fluctuates when a motion is detected. When the intensity of the infrared heat increases, the amplitude increases (above or below the centre line). Since the PIR is sensitive to infrared heat motion, the amplitude fluctuation decreases when the distance of the event increases and vice versa. It can be seen from Fig. 3 that the fire categories 1 (small wood fire) and 2 (spreading wood fire and propane gas fire) signals have almost similar trends with varying distances. For instance, the spreading wood fire and propane gas fire signals within and at 1 m have amplitude variations between 0 V and 1.09 V. However, when the distance increases (at 2 and 3 m), their fire signal amplitudes narrow down to a range between 0.5 V and 0.8 V. Since the spreading fire (category 2) occupies a wider space within the viewing range of the sensor, its motion is detected better than the small fire (category 1). For example, at 1 m distance, small fire signal amplitude drops to a range between 0.35 V and 0.94 V, whereas spreading fire signal amplitude stays between 0 V and 0.96 V.

In contrast, human movement signals within 1 m show an amplitude range between 0.08 V and 0.83 V (see Fig. 4). Similar to fire signals, this range narrows down further with the distance. It is observed that there is no noticeable difference in the amplitude for human slow and fast walking motions beyond 1 m. In addition, it can be seen that the human motion and the fire beyond 2 m have similar amplitude ranges. This makes it challenging to differentiate daily human activities from a fire using amplitude-based event categorisation.

2.3. Signal Processing

Signal processing focuses on analysing, modifying, and synthesising a signal based on the user requirements [27]. Due to the dynamic nature of fire flame radiation
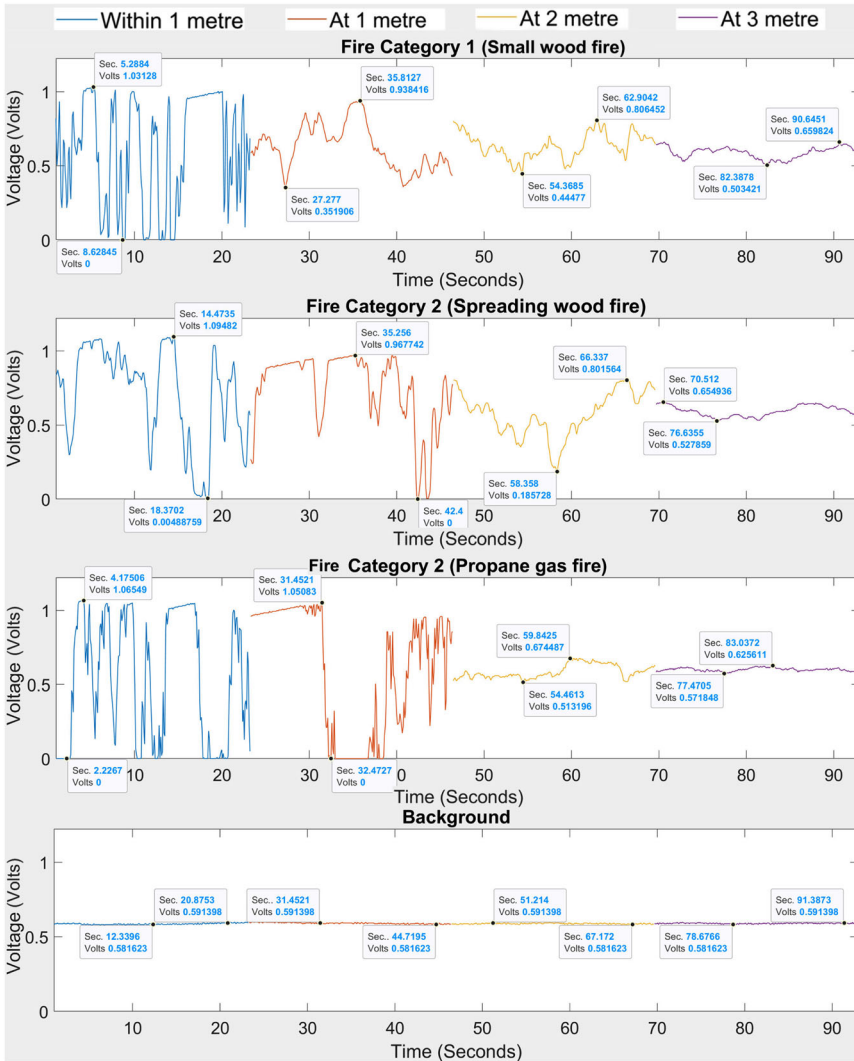
Figure 3. The PIR sensor signals of fire category 1 (small wood fire), fire category 2 (spreading wood fire), fire category 2 (propane gas fire) and background (no motion) from the distance of within 1 m, at 1 m, 2 m, and 3 m. The x-axis consists of 1000 data points captured for 92 s. The y-axis represents the amplitude of the sensor. This is not a continuous recording of data from 0 s to 92 s. During data collection, each event (at all distances) is continuously recorded for 4 mins. For the demonstration purpose of this paper, all the data captured at varying distances are trimmed to 20 s or 25 s portions, merged as a continuous flow of data, and sorted in colours. The highest and lowest values of each event are marked.
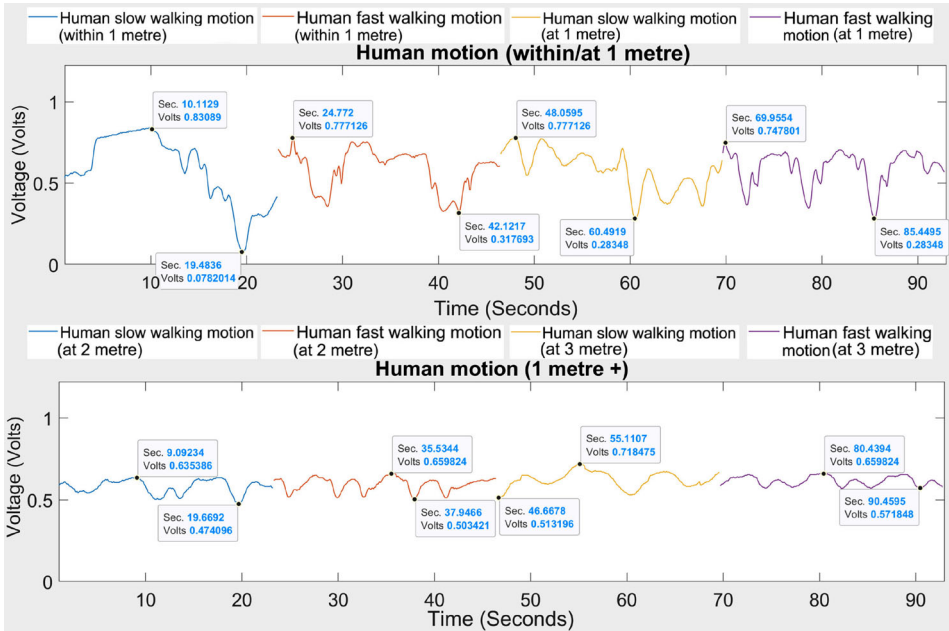
Figure 4.   The PIR sensor signals of human motion (within/at 1 m) and human motion (more than 1 m). The x-axis consists of 1000 data points captured for 92 s. The y-axis represents the amplitude of the sensor. This is not a continuous recording of data from 0 s to 92 s. During data collection, each event (at all distances) is continuously recorded for 4 mins. For the demonstration purpose of this paper, all the data captured at varying distances are trimmed to 20 s or 25 s portions, merged as a continuous flow of data, and sorted in colours. The highest and lowest values of each event are noted.

levels, the fire signals are non-stationary. Widely used signal processing techniques such as Fourier transform assumes the signal being processed is stationary. Hence, in this work, the PIR sensor signals are analysed using wavelets. Unlike Fourier transform, the wavelet analysis helps to build a time-frequency relationship using both time and frequency domains [28]. Wavelets also have inbuilt filters which eliminate noises from the signal and have less computational cost than Fourier transform.

Wavelets are asymmetric and irregular waveforms that last for a short period of time with an average value of zero. Wavelet analysis produces constituent wavelets of different scales and positions of the mother wavelet [20]. Therefore, in continuous wavelet transform (CWT), the input signal is multiplied by delayed and stretched versions of the chosen wavelet and then integrated over the time duration of the signal. This process generates wavelet coefficients C. The process can be shown as:

$$C(scale, position) = \int_{-\infty}^{\infty} f(t)\psi(scale, position, t)dt \qquad (1)$$

where $f(t)$ is the original signal (the PIR signal in this case) in time domain and $\psi(t)$ is the mother wavelet function. There are different types of mother wavelet functions such as Haar, Daubechies, Coiflets, Symlets, Biorthogonal, etc. The mother wavelet needs to be appropriately selected to match the transient spikes of the original PIR sensor signal. Then, better noise cancellations without distortions can be achieved [28]. The wavelet coefficients C indicate how closely the chosen wavelet is correlated across sections of the original signal.

In this paper, the Morse wavelet [29] in the MATLAB Wavelet toolbox [30] is used. The amplified analog output from the PIR sensor is affected by noise and signal drift due to atmospheric temperature variations. The CWT filter bank is used to filter this noise and yield the signal with respect to frequencies. In general, the PIR signal oscillation frequency for a flickering flame is higher than that of PIR signals of a moving object [17]. Various parameters of filter bank are changed in order to visualise distinct features for each event.

2.4. Deep Learning Frameworks

DL frameworks provide building blocks for designing and training deep neural networks using their algorithms, pre-trained models, and applications as well as visualization of the model training progress. Widely used DL frameworks are TensorFlow, Keras, PyTorch, Theano, Caffe, etc. In this paper, the DL toolbox in MATLAB is used [31].

2.5. Data Pre-processing and Dataset Construction

Since the input data structure for DNN should be in image format (3-D array), data pre-processing is carried out. The data collected from the PIR sensor is in a continuous time series format (refer Figs. 3, 4) which is then segmented in lengths of 50 data points in steps of 5 for training the DNN. These 50 data points are assigned as data sets that are processed through CWT.

Then, the size of the resulting 2-D matrix with absolute coefficients is 61×50 in which each row and column corresponds to a scale and a data point respectively. These absolute coefficients are re-scaled and converted into a spectrum of size $61 \times 50 \times 3$ (scale $\times$ datapoints $\times$ RGB). The selected DNN architecture (refer Table 4) requires an array of size $224 \times 224 \times 3$ as the input data. Thus, scalograms from CWT are transformed into RGB image format and are resized as required using data augmentation techniques.

Figure 5 shows examples of RGB image format of the scalograms with respect to the five events generated by CWT. From the $224 \times 224 \times 3$ matrix (spectrum image size), the DNN extracts the distinct features for each class/event and learns to recognise the event happening in front of the sensor. To avoid overfitting problems, the same number of data should be provided during the training progress. Thus, each class/event (refer Sect. 2.1) consists of 5915 RGB spectra for training, 2535 for validation, and 1690 for testing.
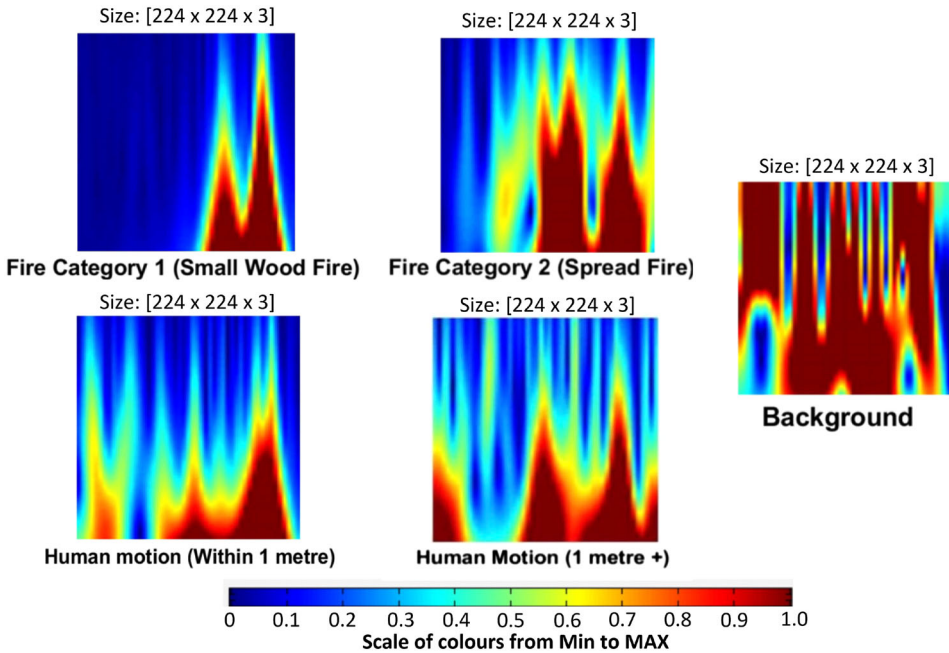
Figure 5. Examples of scalograms with respect to the five events. The colour of the spectrum varies from min to max colours specified in the scale. Each event consists of a single set of samples (50 data points) which are processed using CWT and converted into RGB spectrum images of size 224 × 224 × 3 as required input dimensions for DNN architectures specified in Table 1. The x and y axes are reshaped data points and the scaling factor respectively.

2.6. DNN Architecture

CNN is the most established DL algorithm to process data patterns, extract, and classify features [32]. The DNN architecture has three key layers: convolution, pooling, and fully connected. The convolutional layer plays an important role in DNNs. In order to perform feature extraction, convolution requires a specific size kernel or filter to slide over the input image based on the stride given. The stride which is the distance between two consecutive kernels, determines the columns in which the filter should move on the input image. During the process, the element-wise matrix multiplication is achieved at every location between the kernel and the input to extract the feature map from the image. Likewise, multiple kernels act as different feature extractors. During the training process, the best kernels in the convolution layers and weights in the fully connected layers are identified for the training data set. The convolutional layer proves effective because the layers closer to the input learn low level features such as lines, curves, etc. and as the layers go deeper, the network learns high level features such as shapes, specific objects, etc.

Table 1
DNN Architectures Used in This Work

DNN architecture	No. of layers	No. of parameters (M)
1. VGG-16	16	138
2. Modified GoogLeNet (Model 1)	144	7
3. Model 2	130	6.04
4. Model 3	103	4.15
5. Model 4	88	3.88
6. MobileNet-V2	154	3.4
7. ShuffleNet-V2	50	1.4

They are sorted according to the number of parameters, from highest to the lowest. The models 1–4 are custom-designed DNNs which are trimmed versions of GoogLeNet [33]. Their architectures are illustrated in Table 2

Table 2
Modified GoogleNet Architectures for Models 1–4

GoogleNet [33]	Model 1	Model 2	Model 3	Model 4
convolution	1	1	1	1
max pool	1	1	1	1
convolution	1	1	1	1
max pool	1	1	1	1
inception (3a)	1	1	1	1
inception (3b)	1	1	1	1
max pool	1	1	1	1
inception (4a)	1	1	1	1
–	–	–	dropout (60%)	dropout (60%)
inception (4b)	1	1	–	–
inception (4c)	1	1	1	1
inception (4d)	1	–	–	–
inception (4e)	1	1	1	–
max pool	1	1	1	1
inception (5a)	1	1	1	1
inception (5b)	1	1	–	–
avg pool	1	1	1	1
dropout(40%)	dropout (60%)	dropout (60%)	dropout (60%)	dropout (60%)
linear	1	1	1	1
softmax	1	1	1	1

In here, "1" and "–" denote inclusion and elimination of a particular layer respectively

In this work, MobileNet-v2, ShuffleNet, VGG-16, GoogLeNet, and custom-designed DNNs are chosen to train data (see Table 1).

During training, the selection of hyper-parameters such as the stride, number, and the size of kernels play a vital role in optimising the model. All NNs are trained using the Stochastic Gradient Descent with Momentum (SGDM) optimiser. The initial learning rate of $1.0 \ e^{-4}$ is assigned with a mini-batch size of 84.
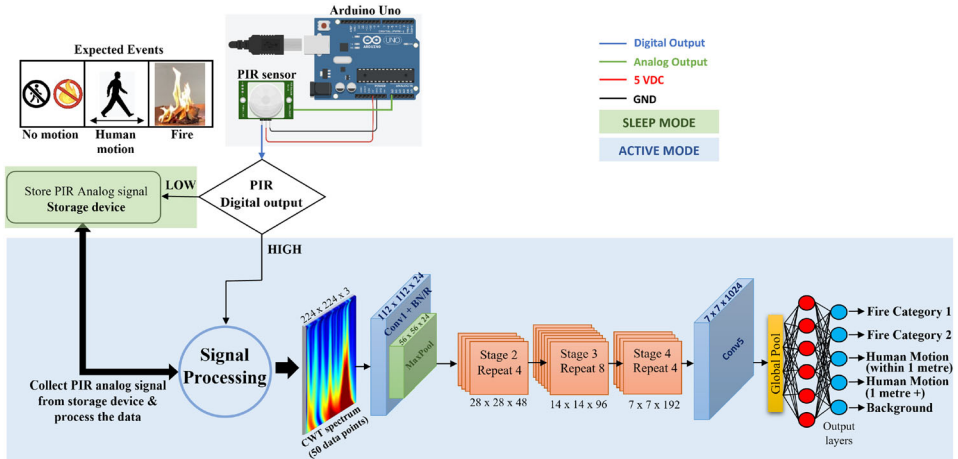
Figure 6. The ShuffleNet-V2 architecture functioning in either ''SLEEP MODE'' or ''ACTIVE MODE'' depending on the PIR sensor output being either LOW or HIGH.

In addition, a dropout regularisation technique (Table 2) is utilised to reduce the error by fitting a function appropriately for the given training dataset and to avoid overfitting [32]. An attempt to increase the dropout layers in the custom-designed DNN is also achieved to evaluate the performance. Throughout the training process, the model is saved in every 5 epochs. Each training process is set to 10 epochs. After completing 10 epochs, the training data is validated. If the loss is not decreased in nearly 3 epochs, the training model is stopped to avoid overfitting.

2.7. Real-Time Working Strategy

The schematic diagram of the real-time working strategy is shown in Fig. 6. The two stages are explained below.

2.7.1. Sleep Mode When no motion is detected, it is known as the normal state or sleep mode. The digital output of the PIR sensor acts as a trigger. In the normal state, the digital output of PIR will be LOW which shuts down the signal processing unit and the DNN. Meanwhile, the analog output from the PIR sensor is stored in the storage device.

2.7.2. Active Mode When a motion is detected, the digital output from the PIR will be HIGH. This initiates the trigger which turns on the signal processing unit and the DNN to classify the event happening inside the viewing range of the sensor. The system works until the DNN classifies the background (no motion). Then, it goes back to the SLEEP MODE (shutting down the signal processing unit and the DNN model).

Table 3
DNN Training Results

DNN architecure	Validation accuracy (%)
1. VGG-16	80.02
2. Modified GoogLeNet (model 1)	82.50
3. Model 2	81.47
4. Model 3	78.33
5. Model 4	78.11
6. MobileNet-V2	78.65
7. ShuffleNet-V2[a]	82.64
8. ShuffleNet-V2[b]	83.6

[a]Without pre-trained weights
[b]With pre-trained weights. Its validation accuracy and training loss are shown in Fig. 7

Table 4
ShuffleNet-V2 Architecture [36]

Layer	Output size	Ksize	Stride	Repeat	Output channels 0.5×
Image	224×224				3
Conv1	112×112	3×3	2	1	24
MaxPool	56×56	3×3	2		24
Stage2	28×28		2	1	48
	28×28		1	3	
Stage3	14×14		2	1	96
	14×14		1	7	
Stage4	7×7		2	1	192
	7×7		1	3	
Conv5	7×7	1×1	1	1	1024
GlobalPool	1×1	7×7			
FC					1000
FLOPs					41M
Weights					1.4M

Thus, the algorithm works when a motion is detected. This makes the model energy-efficient by resisting the signal processing unit and deep CNN model at the normal state. During active mode, the last 50 data points are processed by the signal processing unit and DL network to classify the event. As a result, the entire real time working strategy is similar to a video stream that refreshes every time when a new data point is collected from the PIR sensor. The refresh rate is discussed in Sect. 3.3.
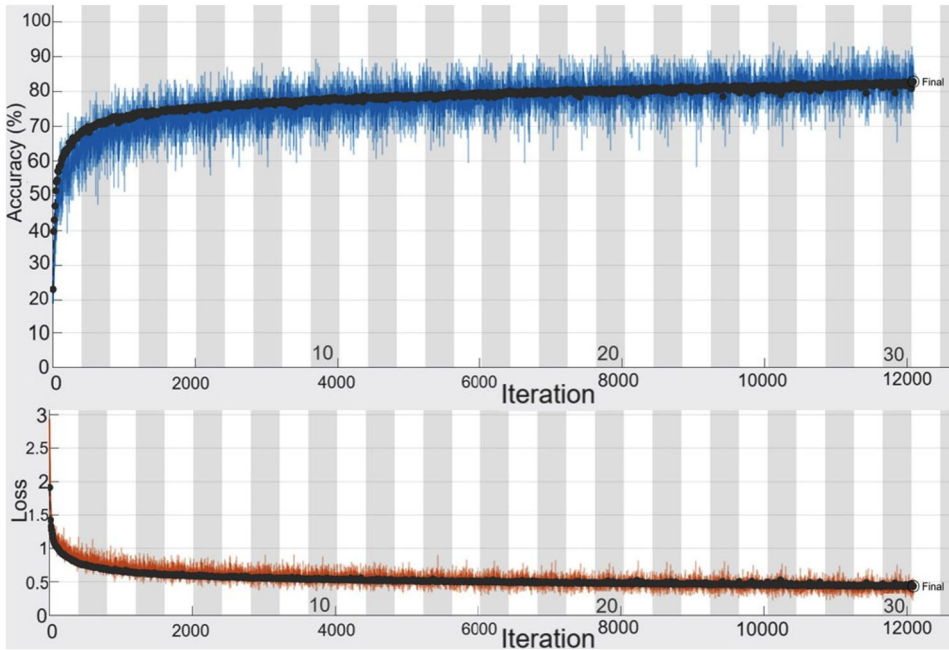
Figure 7.   ShuffleNet-V2 validation accuracy and validation loss with respect to epochs.

3. Experimental Results

This section describes the training and CNN validation results. Then the real time performance observation is presented for the five events selected in this work.

3.1. Training Results

All DNNs listed in Table 1 are trained using the experimental data for the five events collected in a controlled environment (see Fig. 1). The corresponding training results are given in Table 3. When GoogLeNet with the default dropout layer is used, the validation accuracy is 79.91%, whereas GoogLeNet without any dropout layers gives a validation accuracy of 78.1%. However, adding more dropout layers (refer Table 2) gradually brings down the accuracy of the GoogLeNet model as shown in Table 3. A batch normalization (BN) layer is added to the GoogLeNet architecture to improve the training speed and the performance. Following the guidelines from [34], the BN layer is added inside the last 4 inception modules out of 9 modules stacked linearly in GoogLeNet architecture [35]. However, when the dropout layer of GoogLeNet is removed, the accuracy is 76.25% which is less accurate than using the default GoogLeNet.

Table 3 shows that the ShuffleNet with pre-trained weights gives 83.6% validation accuracy. Since it has a lesser number of layers and parameters (refer Table 1), its computational cost is also low. Thus, the ShuffleNet-v2 is chosen as

Confusion Matrix

	Background	Fire category 1 (Small Wood Fire)	Fire category 2 (Spread Fire)	Human Motion (1 metre +)	Human Motion (Within 1 metre)	
Background	**1689** 20.0%	**0** 0.0%	**7** 0.1%	**2** 0.0%	**0** 0.0%	99.5% 0.5%
Fire category 1 (Small Wood Fire)	**0** 0.0%	**1401** 16.6%	**211** 2.5%	**36** 0.4%	**50** 0.6%	82.5% 17.5%
Fire category 2 (Spread Fire)	**1** 0.0%	**154** 1.8%	**1359** 16.1%	**45** 0.5%	**25** 0.3%	85.8% 14.2%
Human Motion (1 metre +)	**0** 0.0%	**74** 0.9%	**52** 0.6%	**1517** 18.0%	**162** 1.9%	84.0% 16.0%
Human Motion (Within 1 metre)	**0** 0.0%	**61** 0.7%	**61** 0.7%	**90** 1.1%	**1453** 17.2%	87.3% 12.7%
	99.9% 0.1%	82.9% 17.1%	80.4% 19.6%	89.8% 10.2%	86.0% 14.0%	**87.8% 12.2%**

Output Class (vertical axis label)

Target Class (horizontal axis label, with categories: Background, Fire category 1 (Small Wood Fire), Fire category 2 (Spread Fire), Human Motion (1 metre +), Human Motion (Within 1 metre))

Figure 8. ShuffleNet-V2 confusion matrix. Each event consists of 1690 images.

the DNN model for this work and the architecture is illustrated in Table 4. Here, the DNN architecture inputs are the wavelet coefficients of PIR sensor signals converted to RGB spectra. When the ShuffleNet-V2 architecture is trained using 1-channel PIR sensor data, the validation accuracy drops to 80.86% compared to 83.6% using RGB spectra. The training progress of ShuffleNet-V2 is shown in Fig. 7.

3.2. Test Results

Figure 8 shows that the true positives (predictions) are higher than false predictions. Each event has 1690 images, in which the background has the highest true positives of 99.9%. Human motion predictions achieved 86% accuracy within 1 m and 89.8% for more than 1 m. The fire category 1 and fire category 2 score an
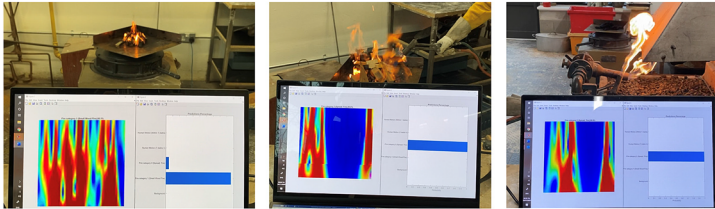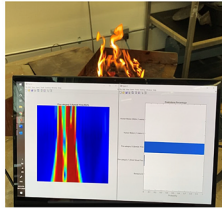
(A) - Fire Category 1 (small wood fire) at 1m

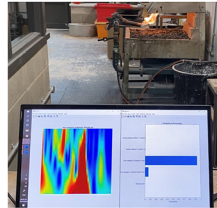(B) – Fire Category 2 (spread fire) within 0.6-0.8m

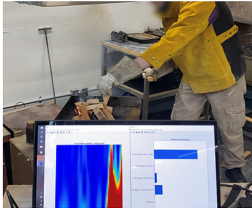(C) – Fire Category 2 (propane gas fire) at 1m
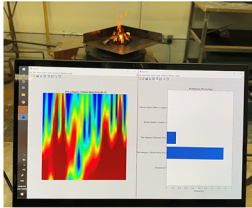
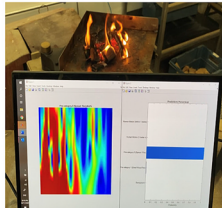(D) - Fire Category 1 (small wood fire) within 1.5-2m

(E) - Fire Category 2 (spread wood fire) within 0.6-0.8m
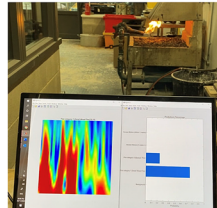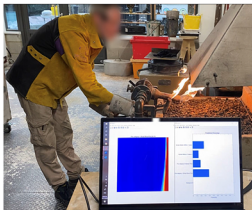
(F) - Fire Category 2 (propane gas fire) at 2m

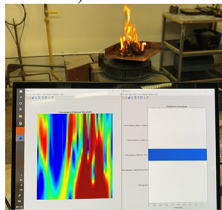(G) - Fire Category 1 (small wood fire) + human motion within 1m

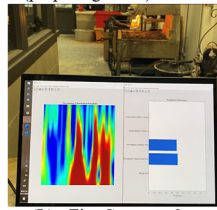(H) - Fire Category 2 (spread wood fire) at within 0.6-0.8m

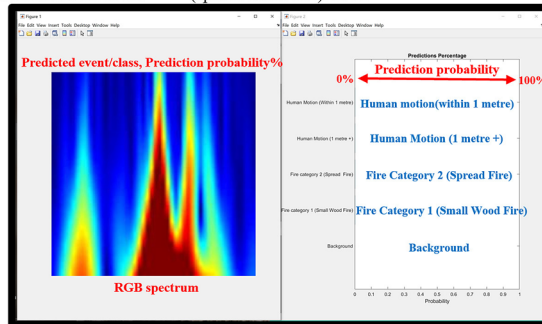(I) - Fire Category 2 (propane gas fire) at 2m

(J) - Fire Category 2 (propane gas fire) + human motion at 1m

(K) - Fire Category 2 (spread wood fire) at 2m

(L) - Fire Category 2 (propane gas fire) within 2.5-3m

(M) – Laptop Screen for prediction probability reference

accuracy of 82.9% and 80.4% respectively. Since both fire categories have similar features such as flame length, flame tilt angle, etc. at some instances, there are false positives for the two fire categories. False positive rate for fire category 1 is 2.5%, while that for fire category 2 is 1.8%. Similar predictions occur in human motion events as well. However, false predictions occur only within the main category events: fire and human motions. Overall, the accuracy rate of true prediction and false prediction are 87.8% and 12.2% respectively. When a counter system is introduced, the final fire detection is 95.34% and human motion detection is 92.39% as given in Table 6. In this system, an alarm is triggered only when a fire is predicted more than three times, without any oscillations to human motion or background.

3.3. Real-Time Performance Observation

The real-time performance evaluation is conducted on Intel Core i9-9980HK 2.4GHz CPU with 32 GB RAM and Nvidia GeForce RTX 2060. The real-time performance is carried out at the speed of 12 frames per second. Observations for the selected real-time fire events are discussed below. Some samples of model predictions captured during the real-time performance are shown in Fig. 9. The key to refer prediction probabilities is given in Fig. 9(M).

3.3.1. Background As shown in Fig. 8, the real-time performance for background prediction is almost 100% true positive prediction.

3.3.2. Human Motion (Within and Beyond 1 m) Human slow and fast walking motions, and human rotating on a swivel chair are observed in this category.

- *Within 1 m*: The true positive is higher than false positive. In case of mixed events (see Fig. 9G) with the small wood fire and human motion (within 1 m), the model predicted 81.8% as human motion (within 1 m) and fire category 1 as 16%.
- *From 1 m–3 m*: In this occasion, the probability of true positives is almost 87%. False predictions occur in between true predictions within the 12 predictions per second. Most of the false predictions are human motion (within 1 m).

3.3.3. Fire Category 1 (Small Wood Fire)

- *Within/at 1 m*: The algorithm yields high true predictions of nearly 95% (see Fig. 9A). However, a small percentage (1.8%) gets misclassified to fire category 2 if the small wood fire moves faster due to wind.

- *From 1 m–2 m*: It predicts high true positives (86%) with less false positives (14%) as shown in Fig. 9D. During the event, 0.9% of the predictions oscillate to human motion (1 m +) in between the true positive predictions.
- *From 2 m–3 m*: Within 2 m–3 m distance from the fire, high false positives are noted mostly due to human motion (1 m +).

3.3.4. Fire Category 2 (Spreading Fire)

- *Within/at 1 m*: In both spreading wood and propane gas fires, 100% true predictions are achieved without any false prediction oscillations during the real time performance (see Fig. 9B, C, E, H). The model also performs well in mixed events of propane gas fire (category 2) and human motion (at 1 m) as shown in Fig. 9J. Since the propane gas fire is smaller at that instance, the model predicts fire category 1, human motion (within 1 m), and fire category 2 as 37%, 25%, and 18% respectively.
- *From 1 m–2 m*: In this occasion (spreading wood fire and propane fire), almost 90% true positives are observed with less false predictions of small wood fire (see Fig. 9F, I, K).
- *From 2 m–3 m*: During the spreading wood fire, high true positive predictions are noted. However, high oscillations of false predictions are observed during propane fire. Due to the distant view (see Fig. 9L), the propane gas fire is predicted as almost 50% for both fire categories 1 and 2. There is a 50% chance to predict a propane gas spread fire as a small fire at a distance of 2 m–3 m.

4. Discussion

In this paper, experiments are carried out using a PIR sensor to collect human, fire, and background data to train DNNs and then the fire detection is tested in real-time. Since, the ShuffleNet-V2 yields the highest validation accuracy of 87.8% at the speed of 12 frames per second, it is used along with the PIR sensor to perform real-time fire detection experiments. PIR sensor based fire detection is almost equivalent in performance to a video-based fire detection system. However, video-based systems are relatively more expensive. Unlike in vision-based fire detection systems, the PIR sensors can detect fire using one-dimensional signals with less computational cost. In addition, the real-time working strategy is modified to reduce false alarms and hence to use computational power effectively. As discussed in Sect. 3.3, real-time predictions can fluctuate within the 12 predictions per second period. Hence, a counter is built to initiate an alarm if a fire is predicted in more than 3 predictions, i.e., for 0.25 s, without any oscillations to human motion or background. This counter system effectively reduces the false alarm rate as shown in Table 5.

In literature, PIR sensors are used mainly to detect either fire/flame or human motions [17, 37, 38]. In our proposed methodology, a PIR sensor is used to pre-

Table 5
Event Detection Percentages of the Proposed Model with Counter System Enabled

Events	Number of cases	Number of true predicted cases	Percentage of detection (%)
Fire test	322	307	95.34
Human motion test	841	777	92.39
Background test	349	348	99.71

Each sequence consists of three sets of samples (50×3 data points) to make three predictions to trigger an alarm. In this case, random data sets are tested from the five categories noted in Sect. 2.1

Table 6
Comparison of Results from the Proposed Method (After Enabling Counter System) and Other Similar Models

Model references	True positive event	True negative event
[17]—Fire detection	Fire test: 98%	Non-fire: 100%
[37]—Fire detection	Fire test: 98%	Non-fire: 100%
[38]—Human detection	Human motion: 96%	Background: 100%
Proposed model for both fire and human detection[a]		
- Fire detection	Fire test: 95.34%	Non-fire: 99.69%
- Human detection	Human motion: 92.39%	Background: 99.71%

[a]Fire, human motion, and background detection are done simultaneously even though they are shown separately for comparison purpose. All three categories of data sets are used together in training and testing as shown in sample Fig. 9G and J

dict both fire and human motion in real-time when happening within the viewing range of the sensor. Table 6 illustrates some fire and human motion detection predictions found in literature compared with the proposed model predictions. The proposed model shows promising results in detecting both fire and human motion. The referenced models intended either to detect fire or human motion patterns; not both. According to [17, 37], the events are categorised only as fire and non-fire events, in which the authors consider paper and alcohol fire as the fire category. Work in [38] focuses only human motion event detection. In our proposed method, the intention is to detect the spread of fire using two different fire sources and human motion patterns. In addition, the proposed model is capable of differentiating the fire as small or spreading fire within 2 m (see Fig. 9).

Table 5 illustrates the background (no motion), fire and human detection rates as 99.71%, 95.34% and 92.39% respectively. As shown in Fig. 9, our proposed method tests both fire and human motion categories: small and spreading fire, and human motion within 1 m and beyond 1 m from the PIR sensor. Figures 3 and 4 show that the small wood fire at 3 m is similar to a human motion at a distance of more than 1 m; specifically at 3 m. This shows that different human motion patterns result in different patterns of PIR data which are similar to fire patterns

at some instances. As noted in Table 5, this issue is addressed in our model using the counter system to stop confirming any false fire predictions or detection.

The DNN is trained with limited data for human motion because the human motion has various patterns of movement. The PIR sensor produces different output data for different human movements/actions. For instance, the human walking motion is different from human running motion; human hand waving above the head is different from human dancing movements. Hence, it is emphasised that the DNN is trained with limited data from human slow walking, fast walking, and rotating on the swivel chair. Thus, during real-time prediction, if a human does some irregular motions (excluding the trained human motion events) such as dancing in front of the sensor, there is a higher possibility to predict it as fire. But, a high true positive prediction is obtained when testing with three simple patterns of human motion mentioned above.

The accuracy and robustness of the algorithm can be improved by training it with different human motion patterns and fire categories. Moreover, this work can be extended to use alternative DL networks such as the variations of Long short-term memory (LSTM) as presented in [5, 39] to improve fire detection reliability.

In this study, the PIR sensor did not experience occlusions. The literature on visual recognition of objects with occlusions use extensive training of DNNs with labelled images of objects with occlusions [40]. A similar approach can be used for fire categorization with occlusions. The accuracy can be further increased by using additional sensors such as smoke sensors.

In advanced high-tech buildings, the motion sensors are used as automatic on/off light switches to reduce power consumption. Our fire and human motion detection results indicate that the proposed technique can be extended to incorporate those light on/off switch systems to serve multiple purposes such as fire and human action detection. Human detection can also work like a burglar alarm which can be turned on when the user is outside the office or residential building.

5. Conclusion

In this paper, a fire detection algorithm based on a differential PIR sensor and DNNs is proposed. Experiments are carried out using a PIR sensor to collect human motion, small and spreading wood fire, and background data which are then used to train DNNs. Comparative training results with various DNN architectures are presented. Since, the ShuffleNet-V2 yields the highest validation accuracy of 87.8% at the speed of 12 frames per second, it is used to carry out real-time fire detection with the PIR sensor. The system predicts two categories of fire and three patterns of human motion successfully during real-time experiments. The detection rates for background (no motion), fire types, and human motions are 99.71%, 95.34%, and 92.39% respectively. Moreover, enabling the counter technique in the system reduces the false alarm rate by 65.66%. The overall false prediction is 12.2% before the counter technique, and it drops to 4.19% after the counter technique. This study indicates that the proposed system can be used to

differentiate varieties of fire types and human motion patterns if sufficiently trained.

6. Supplementary information

A video showing the experiments conducted to collect fire data and for real-time fire predictions, is submitted along with the manuscript.

Acknowledgements

Author Contributions

KLBLX—conducted the experiments, analysed data, and wrote the paper. VKN—supervised, contributed to write and edit the paper.

Funding

Data Availability

Not applicable.

Code Availability

Not applicable.

Declarations

Competing interests   All the authors confirm that there are no financial and personal relationships with other people or organisations that could inappropriately influence or bias this work.

Ethical Approval   Not applicable.

Consent to Participate   Not applicable.

Consent for Publication   The authors declare that all the authors made significant intellectual contribution in interpreting and revising the work and

approved for publication.

Open Access

SUPPLEMENTARY INFORMATION

The online version contains supplementary material available at https://doi.org/10.1007/s10694-022-01319-x.

References

1. Gaur A, Singh A, Kumar A, Kulkarni KS, Lala S, Kapoor K, Srivastava V, Kumar A, Mukhopadhyay SC (2019) Fire sensing technologies: a review. IEEE Sens J 19(9):3191–3202
2. Zhang G, Li H, Zhu G, Li J (2020) Temperature fields for fire resistance analysis of structures exposed to natural fires in large space buildings. Struct Des Tall Spec Build 29(4):1708. https://doi.org/10.1002/tal.1708
3. Oswald D (2021) Homeowner vulnerability in residential buildings with flammable cladding. Saf Sci 136:105185. https://doi.org/10.1016/j.ssci.2021.105185
4. Snoonian D (2003) Smart buildings. IEEE Spectr 40(8):18–23. https://doi.org/10.1109/MSPEC.2003.1222043
5. Xu Z, Guo Y, Saleh JH (2021) Advances toward the next generation fire detection: deep LSTM variational autoencoder for improved sensitivity and reliability. IEEE Access 9:30636–30653
6. Liu Z, Kim AK (2003) Review of recent developments in fire detection technologies. J Fire Prot Eng 13(2):129–151. https://doi.org/10.1177/1042391503013002003
7. Jadon A, Omama M, Varshney A, Ansari MS, Sharma R (2019) Firenet: a specialized lightweight fire & smoke detection model for real-time iot applications. arXiv:1905.11922
8. Fire and rescue incident statistics in England (2020) https://www.gov.uk/government/statistics/fire-and-rescue-incident-statistics-england-year-ending-december-2020. Accessed July 2021
9. Çetin AE, Dimitropoulos K, Gouverneur B, Grammalidis N, Günay O, Habiboglu YH, Töreyin BU, Verstockt S (2013) Video fire detection-review. Dig Signal Process 23(6):1827–1843

10. Li P, Zhao W (2020) Image fire detection algorithms based on convolutional neural networks. Case Stud Therm Eng 19:100625

11. Ayyubi S, Miao Y, Shi H (2014) Automating standalone smoke alarms for early remote notifications. In: 2014 13th international conference on control automation robotics & vision (ICARCV). IEEE, pp 675–680

12. Mao W, Wang W, Dou Z, Li Y (2018) Fire recognition based on multi-channel convolutional neural network. Fire Technol 54(2):531–554. https://doi.org/10.1007/s10694-017-0695-6

13. Wen Z, Xie L, Feng H, Tan Y (2019) Robust fusion algorithm based on RBF neural network with TS fuzzy model and its application to infrared flame detection problem. Appl Soft Comput 76:251–264. https://doi.org/10.1016/j.asoc.2018.12.019

14. Geetha S, Abhishek C, Akshayanat C (2021) Machine vision based fire detection techniques: a survey. Fire Technol 57(2):591–623. https://doi.org/10.1007/s10694-020-01064-z

15. Soyer EB (2009) Pyroelectric infrared (PIR) sensor based event detection

16. De P, Chatterjee A, Rakshit A (2020) PIR sensor-based AAL tool for human movement detection: modified MCP-based dictionary learning approach. IEEE Trans Instrum Meas 69(10):7377–7385

17. Erden F, Toreyin BU, Soyer EB, Inac I, Gunay O, Kose K, Cetin AE (2012) Wavelet based flickering flame detector using differential PIR sensors. Fire Saf J 53:13–18. https://doi.org/10.1016/j.firesaf.2012.06.006

18. Leavitt A (2018) What is the best motion detector for security systems?

19. Gami H (2017) Movement direction and distance classification using a single pir sensor. IEEE Sens Lett 2(1):1–4

20. Misiti M, Misiti Y, Oppenheim G, Poggi JM (1996) Wavelet toolbox for use with MATLAB. Mathworks, pp 23–37

21. Zhong Z, Wang M, Shi Y, Gao W (2018) A convolutional neural network-based flame detection method in video sequence. SIViP 12(8):1619–1627

22. Ryu J, Kwak D (2021) Flame detection using appearance-based pre-processing and convolutional neural network. Appl Sci 11(11):5138

23. Sharma J, Granmo O-C, Goodwin M, Fidje JT (2017) Deep convolutional neural networks for fire detection in images. In: international conference on engineering applications of neural networks. Springer, pp 183–193

24. Thomson W, Bhowmik N, Breckon TP (2020) Efficient and compact convolutional neural network architectures for non-temporal real-time fire detection. In: 2020 19th IEEE international conference on machine learning and applications (ICMLA). IEEE, pp 136–141

25. (2021) Mpja.com: Hc-sr501 PIR motion detector

26. mathworks.com: Matlab arduino toolbox (2021)

27. Mneney S (2008) An introduction to digital signal processing. River Publishers, Gistrup

28. Byeon Y-H, Pan S-B, Kwak K-C (2019) Intelligent deep models based on scalograms of electrocardiogram signals for biometrics. Sensors 19(4):935

29. Lilly JM, Olhede SC (2012) Generalized morse wavelets as a superfamily of analytic wavelets. IEEE Trans Signal Process 60(11):6036–6041

30. (2021) mathworks.com: Matlab wavelet toolbox

31. (2021) mathworks.com: Matlab deep learning toolbox

32. Yamashita R, Nishio M, Do RKG, Togashi K (2018) Convolutional neural networks: an overview and application in radiology. Insights Imaging 9(4):611–629

33. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)
34. (2022) mathworks.com: batch normalization layer
35. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1–9
36. Ma N, Zhang X, Zheng H, Sun J (2018) Shufflenet V2: practical guidelines for efficient CNN architecture design. arXiv:1807.11164
37. Ugur Toreyin B, Birey Soyer E, Urfalioglu O, Enis Cetin A (2008) Flame detection system based on wavelet analysis of pir sensor signals with an hmm decision mechanism. In: 2008 16th European signal processing conference, pp 1–5
38. Urfaliglu O, Soyer EB, Toreyin BU, Cetin AE (2008) PIR-sensor based human motion event classification. In: 2008 IEEE 16th signal processing, communication and applications conference, pp 1–4. https://doi.org/10.1109/SIU.2008.4632611
39. Jain P, Coogan SC, Subramanian SG, Crowley M, Taylor S, Flannigan MD (2020) A review of machine learning applications in wildfire science and management. Environ Rev 28(4):478–505
40. Kortylewski A, Liu Q, Wang A, Sun Y, Yuille A (2021) Compositional convolutional neural networks: A robust and interpretable model for object recognition under occlusion. Int J Comput Vis 129(3):736–760