Fire Technology, 57, 591–623, 2021

Manufactured in The United States
https://doi.org/10.1007/s10694-020-01064-z

Machine Vision Based Fire Detection Techniques: A Survey

S. Geetha, C. S. Abhishek and C. S. Akshayanat, School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu 600127, India*

Received: 16 April 2020/**Accepted:** 11 November 2020

Abstract. The risk of fires is ever increasing along with the boom of urban buildings. The current methods of detecting fire with the use of smoke sensors with large areas, however poses an issue. The introduction of video surveillance systems has given a great opportunity for identifying smoke and flame from faraway locations and tackles this risk. Processing this huge amount of data is a problem with using these video and image data. In recent times, a number of methods have been proposed to deal with this challenge and identify fire and smoke. Image processing algorithms for detecting flame and smoke, motion-based estimation of smoke, etc are some of the methods that are proposed earlier. Recently, there has been an array of methods proposed using Deep Learning, Convolutional Neural Networks (CNNs) to automatically detect and predict flame and smoke in videos and images. In this paper, we present a complete survey and analysis of these machine vision based fire/smoke detection methods and their performance. Firstly, we introduce the fundamentals of image processing methods, CNNs and their application prospect in video smoke and fire detection. Next, the existing datasets and summary of the recent methodologies used in this field are discussed. Finally, the challenges and suggested improvements to further the development of the application of CNNs in this field are discussed. CNNs are shown to have a great potential for smoke and fire detection and better development can help prepare a robust system that would greatly save human lives and monetary wealth from getting destroyed from fires. Finally, research guidelines are presented to fellow researchers regarding data augmentation, fire and smoke detection models which need to be investigated in the future to make progress in this crucial area of research.

Keywords: Fire smoke detection, Fire, Smoke, Convolutional neural networks, Image processing, Deep learning, Machine vision

1. Introduction

Fire is one of the major disasters that human-kind faces due to the increasing number of cases each year. The development of more buildings and the crowding of spaces has spiked the problem to highly dangerous levels. For instance, accord-

* Correspondence should be addressed to: S. Geetha, E-mail: geetha.s@vit.ac.in, geethabaalan@gmail.com

ing to the National Fire Protection Association (NFPA) [9], a highly developed country such as the United States responded to an estimated 1,318,500 fires in the year 2018. Of the total 3,655 civilian fire fatalities and 15,200 civilian fire injuries, home fires caused 74 % or 2720 civilian fire deaths. There was a total of $25.6 billion direct property loss as a result of fires.

Due to the fast spreading characteristics of fire, the major challenge is to put out the fire at the initial stages where it can be done, rather than spending hours and hours with increasing risk posed to firefighters and human lives trying to extinguish the fire that has now spread everywhere in the vicinity. Fire detection is an area, where existing and newer technology can be used to immediately tackle the impending crisis.

Today, most of the buildings have built-in smoke sensors and fire alarm systems. This is traditionally the most commonly used fire detection system. The system relies on the fact that smoke as a result of fire, rises up and triggers the smoke sensors installed mostly at the ceilings of the buildings or may use a combination of humidity, temperature and other factors [58]. The sensors then further activate the fire alarm and fire suppression systems. This method, being fairly robust has an inherent delay of smoke rising above and hitting the smoke sensor. This delay in most cases can cause the fire to spread rapidly beyond control and thus, needs to be limited as low as possible. An approach to tackle this is by using common surveillance system's video and image data.

In videos or images, fire can be generally characterized as orange or yellow flames which move from side to side. Smoke can be characterized as a combination of a white, grey and black plumes that contain tiny particles of soot or burnt particles. Smoke arises due to the burning of materials by fire, which can be caused due to various reasons such as arson, electrical sparks, chemical reactions, etc. Smoke is typically heavier than clean air and rises up and moves rapidly propelled by the fire's voracity. Detecting fire or smoke in videos or images poses its own set of challenges. The system that is deployed should be able to distinguish between images that actually have fire in them and images which have the colour of orange or yellow flames but do not have fire in them. The system should also be able to detect smoke, and distinguish between foggy environment and actual smoke. Thus, the system should be robust, accurate and have a very low to nil false detection rate. Furthermore, the usage of videos from surveillance cameras has led to a new type of problem, processing the images. The video cameras produce a stream of videos or images. This large amount of data needs to processed. If done manually, this practically infeasible work would take a huge toll on the workforce. Therefore, with the aim of making the system accurate and as autonomous as possible, numerous methods and systems for fire and smoke detection have been proposed.

Traditionally, the methods explored for detection are flame and smoke detection using basic image processing techniques such as flame detection using colour, motion, etc. The newer technologies, especially in the field of computer vision and supervised learning, such as deep learning, have given a great hope for application in this smoke and fire detection field. Convolutional neural networks (CNNs), a part of deep learning based systems, have been successfully used in various image recognition due to their superior performance compared to traditional image recognition methods.

There are detailed survey papers [10, 15, 39] in the domain of video based fire and smoke detection. The authors of [39] focussed more on hand crafted static and dynamic feature based smoke and fire detection models like features based on colour, texture, and wavelet transform domain which depicted the presence or absence of fire/flame and smoke in the scene. They covered more of the basic layout of traditional smoke detection in video sequences and outlined the common steps like foreground segmentation, feature analysis for ROI identification and classification of smoky/non-smoky regions in the image. Cetin et al. [10] overviewed a collection of papers on short range (less than 100 m) fire detection systems. They presented a detailed investigation of the underlying algorithms—static feature and dynamic extraction and checked the appropriate techniques. They take the survey in the direction of applying Artificial Intelligence to the smoke and fire detection process and bring it down into an engineering problem. The works discussed in our survey are advanced in terms of accuracy and speed as compared to those methods discussed in [10]. Most of the models discussed in [10] are around hand crafted features and are prone to generalisation error of the classifier used. The classifiers by the year of 2013 were not advanced like 2020 models and suffered from errors arising from data-imputation, data-retrieval, variance, biasing, scaling, hyper-parameter and approximation.

The comprehensive survey of Gaur et al. [15] presented deep learning based smoke-fire detection systems and discussed their strengths and weaknesses. They recommended the use of blended features—hand-crafted features and the deep learning model engineered features in the process of smoke and fire detection. This current review is augmenting the survey of [15] and adds value to it by a detailed discussion about the eight benchmark datasets used for evaluation. This aspect will be helpful for the researchers in order to test any model they develop and also to have a reproduce-able research facilitating fair comparison. Further, the deep learning models are evolving at a rapid rate and new variants are discovered frequently, each specialised for specific image characteristics. This survey introduces around nine such CNN variants which have been successfully deployed for smoke and fire detection and thus helps the researchers to try these variants either in isolation or as ensemble models. Approaches employing motion detection, infrared flame detection and contextual object detections are also discussed which augments the survey of [15]. Finally few recommendations for CNN architectures for smoke and fire detection, segmentation and classification into smoky and non-smoky regions are also provided. These guidelines would help researchers in this field to investigate machine vision systems for smoke and fire detection from video sequences.

In this paper we aim to explore and analyse the recent advanced methodologies used in the field of image recognition of fire and smoke and their systems, as well as summarizing these approaches along with recent implementations of CNNs in this field. These approaches use deep learning model engineered features and provide more accuracy and fast detection than the traditional machine learning algorithms. The summary of the methods broadly covers the relevant methodologies that were used from 1996 to 2020.

We intend to review the recent image analysis based methods for autonomous smoke and fire detection and discuss their pros and cons. Moreover, motivated by the success of the recent deep learning based methods for autonomous visual inspection in related applications, this paper provides detailed recommendations/guidelines for the design and development of a machine vision based smoke and fire detection system. We discuss the challenges in the design and development of deep learning methods for smoke and fire detection and provide future research directions. This paper also opens avenues for further research in this domain and poses open research questions for researchers pursuing this area. The following are the contributions of this work:

1. Provide a comprehensive review of image analysis based methods for autonomous smoke and fire detection.
2. Propose a complete smoke and fire detection pipeline using CNNs for event detection, classification and segmentation.
3. Discuss challenges and opportunities in the design and development of deep learning based methods for automatic smoke and fire detection.
4. Highlight open problems and future research directions for research and development of machine vision systems for autonomous smoke and fire detection.

The paper is structured in the following format. Firstly, the basics of CNNs are described along with a brief about the basic image processing techniques. Secondly, their application in the field of flame or fire and smoke detection is discussed. Thirdly, the datasets and the environment setup used in the development is overviewed. Fourthly, analysis of the various traditional image processing methods as well as deep learning methods that are implemented is presented. Finally, the challenges faced through this analysis are explored, followed by some solutions that can be implemented to solve them.

2. Smoke and Fire Detection Methodologies

2.1. Basic Image Processing Techniques

Traditionally, image processing methods use commonly used feature extraction and segmentation techniques to extract the required features from the images. These features are then compared to a set of features that are corresponding to the required object to be detected. If the extracted features from the images match or are similar to the required object's features, then the image can be considered to have the same object. Another method for classifying the extracted features can be done through machine learning methods such as support vector machines (SVM) classifiers.

2.2. CNNs Overview

Convolutional Neural Networks (CNNs) are deep learning networks, an advanced form of machine learning that are primarily used for visual imagery with images and videos being the input data.

Generally, a convolutional neural network consists of an input layer and an output layer, with multiple hidden layers. These hidden layers of a CNN usually consist of a series of convolutional layers that convolve with a dot product or multiplication. Common layers include the activation layer, also known as RELU layer, followed by more convolutional layers such as pooling layers, fully connected layers, normalization layers. They are known as hidden layers due to the masking of the inputs and output by the activation function and the final convolution. A commonly used CNN network's block diagram is shown in Fig. 1. The main use of the layers of the CNN is to extract features and automatically learn useful information from the input data without involving pre-processing and feature selection and extraction techniques. Convolution, the main process behind the CNN feature learning and extraction, is done through applying a sliding dot product mathematical operation of matrices of weights across the entire data of each of the input data, generating of feature maps. This is then usually followed by pooling, a process by which the dimensionality of the feature maps is reduced, resulting in the highlighting of more important and essential features[1, 56].

Thus, the hidden layers of the CNN determine which type of features are being focused on, extracted and learnt by the network. As a result, each variation of the layers of the CNN determine which specific application the network would suit for. Along with these, the input data's focus also determines the level of performance a CNN would provide. A high noise input data would contain high unnecessary details, causing a wrong input of training information to the network. Convolutional neural networks can be classified based on their kernel dimensions where 2D-CNNs consists of kernels of two dimensions whereas 3D-CNNs consists of kernels of three dimensions. The usage of CNNs have been greatly demonstrated with the introduction and development of AlexNet [31], ResNet [18, 19], GoogLeNet [48] and VGG-Net [47] amongst others.

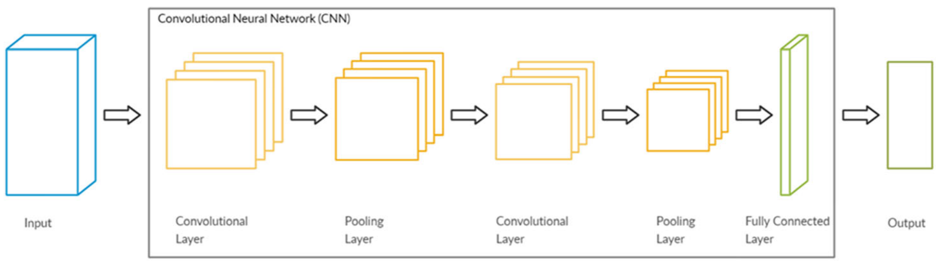

Figure 1. A basic Convolutional Neural Network (CNN) block diagram with two convolutional layers, two pooling layers and a fully connected layer.

A number of methods have been developed in the recent years in the field of fire and smoke detection through traditional image processing methods and through the use of CNNs through deep learning. The traditional image processing methods mainly use the basic image processing techniques with algorithms for identifying fire and smoke, and use machine learning algorithms to classify the results. They involve multi-step feature processing and extraction. A common workflow involved in traditional image processing methods is depicted in Fig. 2a.

CNNs which are deep learning based methods, detect and learn features automatically. Thus, the process applies to all the images and dataset in a unified fashion, thereby eliminating the possibility of errors. A common workflow involved in CNN based methods is depicted in Fig. 2b.

3. Datasets and Environment Overview

Since both the major methodologies, image processing methodologies and CNNs require the use of images for processing, training and testing, some of the commonly used and publicly available datasets are given below. Although, there doesn't exist a standard dataset for the purpose of smoke or flame detection, majority of the below described datasets are frequently used in conjunction with additional images or videos an author may wish to use. The environment setup overview for the two methodologies are also briefed in this section.

3.1. Fire and Smoke Datasets

3.1.1. MIVIA Fire Detection Dataset A fire and smoke detection dataset has been provided by MIVIA Research Lab of University of Salerno in [8, 13], separately. For the fire dataset, a collection of 31 videos are provided. It is composed of, first 14 videos characterized by the presence of fire while the second 17 videos do not contain any fire, but contain objects and events usually detected as fire such as smoke, clouds and red coloured objects moving in the scene. For the smoke dataset, 149 videos, each 15 minutes long is provided. The videos are acquired in real environments and from other sources.

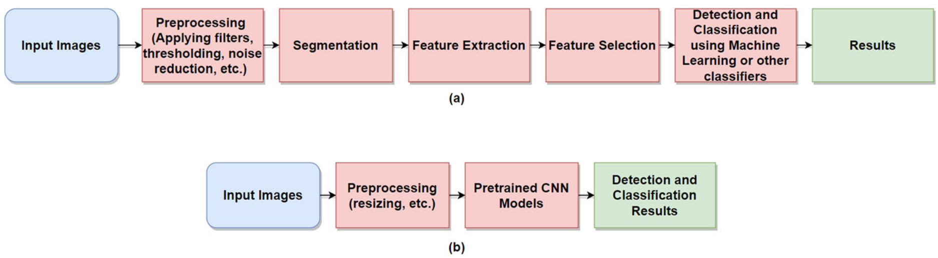

Figure 2. Workflow diagrams of the two categories of flame and smoke detection methods. a Traditional image processing based methods, b Convolutional Neural Network (CNN) based methods.

3.1.2. Firenet Dataset Firenet dataset is a fire dataset that is provided by Arpit et al. [22, 23] from fire and non fire videos that are captured and obtained. The dataset is comprised of 46 videos consisting fire scenes, 16 videos consisting of non fire scenes and an additional 160 images consisting of non fire scenes. In total the dataset consists of 62 videos and 160 images.

3.1.3. Fire Flame Dataset A comprehensive image dataset for fire and smoke detection is given in [7] by Deep Quest AI. The dataset consists of images of 3 classes: Fire, Smoke and Neutral. Each of the classes have 1000 images split as 900 images for training and 100 images for testing. Thus, there are a total of 3000 images in the dataset. Figure 3 shows few sample images from this dataset.

3.1.4. Video Smoke Detection Dataset A video smoke detection dataset consisting of smoke images and videos is released by Dr Feiniu Yuan in [70]. The dataset website consists of 3 smoke videos and 3 non smoke videos. There is a database of smoke and non smoke images with 4 sets of data with Set 1 consisting of 552 smoke and 831 non smoke images, Set 2 consisting of 668 smoke and 817 non smoke images, Set 3 consisting of 2201 smoke and 8511 non smoke images and Set 4 consisting of 2254 smoke and 8363 non smoke images. In addition to these image datasets there are an additional 648 black and white smoke images and two sets of non smoke datasets each consisting of 27707 and 28760 images respectively.

3.1.5. VisiFire Dataset A video dataset for fire and smoke detection is given in VisiFire [3]. A publicly available sample dataset consists of video clips of 4 classes: Fire, Smoke, Other and Forest Smoke. The Fire clips set consists of 13 videos consisting of fire scenes. The Smoke clips set consists of 21 videos consisting of smoke scenes in various environments. The Other clips set consists of 2 videos consisting of cars in one and fire in the other. The Forest Smoke clips set consist of 4 videos consisting of smoke in forest areas. Thus, there a total of 40 video clips in the sample dataset.

3.1.6. State Key Laboratory of Fire Science (SKLFS) Dataset A comprehensive image and video dataset for fire and smoke detection is given in [41] by State Key Laboratory of Fire Science. The dataset consists of 36104 smoke and non-smoke images, with block labels and textures. There are 30000 synthetic image and video datasets and 3578 real image and video datasets on which the deep learning models are trained and developed. The synthetic smoke and non-smoke images have different parameters of rendering, lighting and wind being set randomly in a certain range for diversity. Since different sets of the parameters influence directly the appearance of synthetic smoke images, these images will be realistic or non-realistic.

3.1.7. KMU Fire and Smoke Dataset A video dataset for fire and smoke detection is given in KMU Fire and Smoke Database [29, 30]. A publicly available sample dataset consists of video clips of 4 classes: indoor & outdoor (short distance

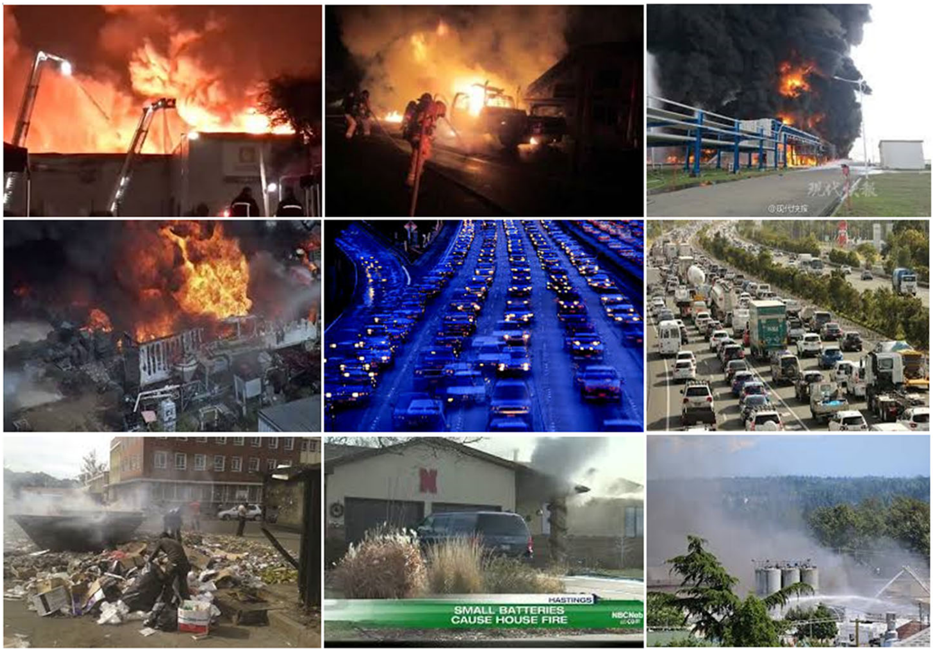

Figure 3. Sample images taken from a fire and smoke image dataset from [7]. Commonly, fire and smoke datasets consists of images or videos mainly in two categories, containing fire and smoke and sometimes others for training and validation purposes.

flame), indoor & outdoor (short distance smoke), wildfire smoke, and smoke or flame-like moving object. Totally there are 308.1 MB video sequences in these four categories.

3.1.8. Corsican Dataset The Corsican Fire Database [51] aims to provide a common dataset of multimodal wildfire images and videos. It provides categories of fire and background properties and is designed to be an evolving database over time. It contains visible spectrum and near infrared (NIR) images in its current form. Additionally, this database contains video sequences captured simultaneously in color and NIR spectrums. These image sequences can serve in the study of multispectral fusion algorithms, the analysis of the performance of fire segmentation in these spectrums, the use of motion for fire segmentation, etc. The users can also contribute to the database by uploading their own images, image sequences (visible, infrared, etc.), corresponding ground-truth, and the image parameters.

3.2. Software and Hardware Environment Overview

Image processing algorithms are usually developed in MATLAB or OpenCV, both of which are high level programming languages and numerical analysis environments that are predominantly used in computer vision based methodology development [43].

In case of deep learning models, the programming language that is mainly used is python. Many frameworks are based on python for deep learning, which include Tensorflow, Keras, Theano, Pytorch, etc. Tensorflow and Keras are the most popularly used deep learning libraries with python. CNNs which are deep learning techniques, greatly depend on these frameworks and thus involves a huge amount of mathematical computation for training, generating and setting weights for the neural network. As a result, there has to be a huge amount of input data to be fed in, and thus a large amount of memory required to hold and process this amount of information. Thus, graphical processing units with multiple computing and superior matrix processing capabilities, greatly aid in the processing of these data.

Nvidia's CUDA (Compute Unified Device Architecture) is a parallel computing platform and application programming interface (API) that supports Nvidia's CUDA-enabled GPUs to use their computing cores to process these data and information [57]. The usage of GPUs for the training and testing of deep learning algorithms has accelerated the development process.

For the development of the basic image processing techniques and superiorly advanced CNNs which are deep learning algorithms, these commonly available programming languages and tools are used.

4. Image Analysis based Fire, Flame and Smoke Detection Systems

Fire flame and smoke detection are the primary indicators for early detection of fire in the surroundings. Therefore, research is primarily focused in this field with the aim of improving the recognition of flame and smoke. The methods and the systems proposed by some of the research works are presented below.

Prior to the introduction of deep learning based image recognition, image processing algorithms were most widely used along with feature extraction and selection techniques. Each of these feature extraction and selection techniques used a unique characteristic of the image such as colour, motion, patterns to help identify the smoke or flame, due to the distinct characteristics of it. These extracted features are then compared with a set of existing standard features through machine learning or other classifiers to classify the image as containing smoke or fire. Some of the image processing algorithms used in the field of smoke and flame detection are given below.

4.1. Hand Crafted Feature based Smoke Detection Systems

In [44] Russo et al. propose a smoke detection method based on Local Binary Pattern (LBP) and Support Vector Machine (SVM). An Approximate Median Filtering Algorithm is applied initially to subtract the background from input frame. Next, shape based filtering method is applied to obtain the region of interest. Then, LBP values and histograms are calculated from the pixels of the region of interest to build a feature vector. Bhattacharyya coefficients are also applied to verify the smoke region for accurate results. In the end, an SVM is used to classify the region of interest as smoke or not. The method achieves a high accuracy rate of 93.27% on a dataset created from videos.

In [49], Tang et al. propose a smoke detection method using local binary pattern, where the motion region is extracted by background subtraction method and every motion area is processed to obtain local information. Each block's texture feature is extracted using local binary model. The texture features of the smoke texture are then used to get the texture feature for smoke image extraction. A support vector machine is then used to classify the extracted features. The method achieves a very good detection rate of 97.238 % and very low 2.176% false alarm rates on a dataset created using images.

Earlier in [69] Yuan et al. propose a novel video-based smoke detection method based on using a histogram sequence of pyramids. Initially, a 3-level image pyramid is constructed by using a multi-scale analysis and local binary patterns (LBP) (used due to rotation and illumination invariance) are extracted at each level of the image pyramid with uniform pattern, rotation invariance pattern and rotation invariance uniform pattern to generate an LBP pyramid. Similarly, local binary patterns based on variance (LBPV) for the same patterns are taken to generate an LBPV pyramid. An enhanced feature vector is constructed by computing the histograms of the LBP and LBPV pyramids and concatenating them. For distinguishing between smoke and non-smoke objects, a neural network classifier is trained and used. The method uses a database created from images for training and testing. It achieves a very high detection rate of above 95.3% and a very low 2.3% false alarm rate.

A smoke-detection framework for high-definition video is proposed by Liu et al. in [34] wherein, small smoke image blocks are used to match the image features of the motion area in the video and further utilizing support vector machine classifier for smoke recognition. For the purpose of extracting the areas for classification, ViBe algorithm and a few other methods are used. This detection framework uses spatial and frequency domain features. In the extraction of local texture features of the spatial domain, compensation of adjacent pixels is added and the gradient of the symmetrical pixels is considered using the center-symmetric local binary pattern features. Local phase quantization (LPQ) features are used in the extraction of frequency domain features. As an improvement over this, the trisection feature fusion scheme for features in the spatial and frequency domains is proposed. A dataset created of images and videos from publicly available sources are used for training and testing. The framework has a very high performance with experimental results giving an average of 97.63% true positive rate (TPR) and an average false positive rate of lesser than 2.5%.

In [14], Gao et al. propose a forest fire smoke detection method based on a diffusion model. The shape of smoke, at the generation stage is identified. The vision and diffusion model is considered and the basic concept of smoke root is considered. While frames are being processed, stable points in dynamic areas as the smoke root candidate points are extracted. All smoke root candidate points information are taken by the model to generate the simulation smoke in the diffusion model simulation stage. The match algorithm based on colour, dynamic areas and simulation smoke is applied to get the result. Contour features are only taken to reduce the complexity of computation. A dataset of videos is created. The method obtains good detection results with accuracy rate of higher than 90% and low false detection rate of 10%, even in cloudy conditions.

Wang et al. in [54] propose a method of video smoke detection using shape, colour and dynamic texture features. The method initially uses an algorithm identifying cone geometry feature is used to differentiate conical region from dynamic regions. Next these conical regions are filtered by using a colour filtering algorithm to test for the candidate smoke region. Then finally by using a texture filtering algorithm, real smoke can be differentiated from candidate smoke regions. A dataset of video clips is used and the method achieves good early detection and low false alarm performances.

Earlier in a method in [76], Zhou et al. propose a video based method for long distance wildfire smoke detection. For the initial smoke region segmentation, Maximally Stable Extremal Region (MSER) detection method is used to extract the local extremal regions of the smoke,reducing motion and colour dependency. Potential smoke regions are then identified from all the possible regions using static visual features of the smoke to eliminate the non-smoke regions. Once a potential smoke region is found, matching extremal regions in the subsequent frames are searched for. The propagating motions of the potential smoke region are checked based on a cumulated region approach so as to identify the distinctive expanding and rising motion of smoke. This approach can also make the smoke motion identification insensitive to image shaking. The proposed method is able to reliably detect long-distance wildfire smoke and produce very less false alarms in real-life applications.

In another method, Wei Ye et. al. in [65] proposed surfacelet transform and hidden Markov tree (HTM) model for smoke detection. Multi-scale decomposition is used through a pyramid model on the image and the signals are decomposed to different directions using 3D directional filter banks. A 3D HMT model is then built for the obtained coefficients from Surfacelet transform with scale continuity model and the Gaussian mixture model . The HMT model parameters are estimated through expectation maximization algorithm. The dynamic texture feature value is taken as the joint probability density. A support vector machine (SVM) classifier is trained with smoke and non-smoke videos and samples. The joint probability density of the divided unit 3D block is input to the SVM to determine the presence of smoke for the input image sequence. The dynamic texture descriptor takes image sequence as a multidimensional volumetric data. This dynamic texture descriptor method achieves high detection accuracy.

In [11], Alexander Filonenko et. al. propose a smoke detection method for cameras. Background subtraction is used to determine moving objects with colour characteristics being utilized to distinguish smoke regions and other scene members. Separate pixels are united into homogenous areas by morphology operations and by connected component labeling methods. The image is further refined through boundary roughness and edge density to reduce false detection rate. The results of the current frame are compared to the previous to check the behavior of objects in the time domain. The method achieves a detection rate of 98% and a dataset of images is created for the purpose of training under different lighting conditions.

A model system proposed by Calderara et al. in [2], uses a stable background supression module joined with a smoke detection module working on segmented objects. This system uses two features for the purpose, energy variation in wavelet model and a colour model of the smoke. A decrease of energy ratio in wavelet domain between background and current image is used to detect smoke representing the variations of texture level. A mixture of Gaussians models this texture ratio for temporal evolution. The colour model is used as a reference to measure the deviation of the current pixel colour from the model. A bayesian classifier is used to detect smoke in the image by combining these two features. A high detection rate of 98.5% and a low false detection rate of 4% is achieved by the system, tested on publicly available datasets and some other source video clips.

Earlier, in [6], Yu Chunyu et. al. propose a novel video smoke detection method using colour and motion features. Candidate smoke regions are identified by using background estimation and colour based decision rule.An approximation of motion field is estimated to be the result of an optical flow for which the calculation is done by a Lucas Kanade optical flow algorithm. Optical flow is calculated for the candidate regions. The optical flow results are used to calculate motion features which are used to differentiate smoke from some other moving objects. Then, a back-propagation neural network is used to distinguish through classification of the smoke features from non-fire smoke features. A dataset of videos was created from [3] and some captured videos. The algorithm is significant for achieving remarkable accuracy of video smoke detection and reducing false alarm rate.

In [5] Chen et al. propose a smoke detection method in videos using a two rule decision strategy of a chromaticity based static decision rule and a diffusion based dynamic characteristic decision rule. The chromatic decision rule is deduced by the grayish colour of smoke and the dynamic decision rule is dependent on the spreading attributes of smoke. Through the analysis of static and dynamic features of smoke, a two stage checking process of smoke detection is taken, wherein the basic strategy is to extract smoke pixels by chromatic checking in the static feature and then these pixels are further verified through the dynamic diffusion checking to check whether it is a smoke as a result of fire or not. The method achieves good performance.

In [53], Wang et al. propose a method based on multi feature fusion for detecting smoke. The suspected region is extracted from the foreground through using a Gaussian mixed model and background subtraction method. Colour features of the suspected region are extracted according to the colour model of early smoke in the RGB and HSI spaces. 2D discrete wavelet transform then is used for background blur feature extraction. The ratio of number of pixels in the suspected region to the number of pixels in corresponding minimum enclosing rectangle is calculated to extract the feature of contour irregularity. Optical flow method is used to the feature of main motion direction. The feature values of colour, background blur, contour irregularity and the main motion direction are used to form a feature vector which is used as the input vector of the support vector machine

(SVM) for classification. The SVM parameters are optimized by Artificial bee colony algorithm. A dataset is created from [3] and other sources while another part is created from capturing videos. Experimental results for the method give a recognition rate for smoke images 99.45% and for non smoke images 100%. Thus, the method achieves very high performance.

4.2. Hand Crafted Feature based Fire Flame Detection Systems

In [35], Liu et al. propose a flame detection algorithm that is based on a saliency detection technique and uniform local binary pattern (ULBP). Utilizing the colour information of flame pixels, probability density function (pdf) of the flame pixel colour is obtained using Parzen window nonparametric estimation. So as to extract the candidate flame area, this a priori probability density function is fused with the saliency detection phase. UBLP is utilized to analyze the image texture of the candidate area and an exponential function with two parameters is utilized to model the texture of the flame area. This is done to reduce false alarms. The method achieved a good accuracy of 89.70%, with true positive rate reaching 99% and true negative rate reaching 75.75% on a dataset created from images and videos.

In [16], Gong et al. proposed a fire detection method based on multifeature fusion of flame. The method involves motion detection and colour detection of the flame initially at the preprocessing stage. Based on the property of similarity of the flame sequence of the image, an algorithm of flame centroid stabilization based on spatiotemporal relation is developed. The centroid of the flame region of each frame of the video is calculated and added to the temporal information so as to obtain the spatiotemporal information of the flame centroid. The next step involves extracting features such as spatial variability, shape variability, and area variability of the flame to improve the accuracy of recognition. Finally, a support vector machine is used for training and analysing the images to detect the presence of fire. The method is trained and tested on a dataset created from videos and achieves a very good accuracy rate of 95.29% and a low false positive rate of 3.09%.

In [52], Vijayalakshmi et al. propose a fire and smoke recognition based on Sensor node and feature of video smoke. To extract the fire and smoke region in the foreground of the video image, Gaussian mixed model, LK optical flow method and background subtraction from foreground methods are used. Multi feature of fire characteristics are used in extracting information. RGB, HSI based respective colour features of the suspected region are extracted. Two-dimensional discrete wavelet transform is used to extract background blur feature because if smoke is present in a scene, the contour edge of the background would become blurry. LK optical flow method and gaussian mixed model is used to extract motion direction feature. Now with the help of DHT 11 digital temperature—humidity sensor in sensor node is used to extract temperature and humidity values. TIMSP430 microcontroller is used for processing the information. Thus the video node and sensor node extracted information are combined to identify the possibility of fire in the scene. Experimental results on a dataset created with images show that the system achieves a good performance with high detection rate of 95% and a low 3% false positive rate.

Liping et al. propose in [33] an approach for the efficient detection of flame in multiple scenes in an image. The approach considers a set of parametric representations called Gradient Features (GF) to use as the features of flame colour changes in the image. The Gradient Features represents the colour changes in RGB channels. A support vector machine is used to generate a set of candidate regions and a decision tree model is used for flame regions judgement based on the Gradient Features. The proposed method was able to obtain a high precision rate of 92.36% from a created image dataset.

By utilizing texture features and optical flow method a fire detection method is proposed by Wang et al. in [55]. Initially, the image of Pyramid is established, and then by utilizing Local Binary Patterns (LBP) and Local Binary Pattern Variance (LBPV), static texture features of various levels can be extracted. Because of the turbulence characteristics of the motion of the smoke, the smoke direction is consistent. An optical flow vector analysis is used to judge the movement directions of the suspicious area contour so as to reduce the computational complexity. The texture features of the smoke image are identified by a Support Vector Machine(SVM). Simulation results show that the proposed algorithm can achieve a high performance with high detection rate of 89.1% and a low false detection rate of 5.4%.

In [26], Xiangang Jiang et. al. propose a static-dynamic fusion feature method that uses a pixel's colour moment and covariance matrix descriptor in both CYMK and YCrCb colour space. The covariance matrix descriptors are used to represent the singular flame's feature in an assemble vector by analyzing rationality of selection and combination of the features. A method of Blending and classification of flame's attributes by sparse dictionary from the covariance matrix descriptor and colour's low order moments is used. Sparse representation classification of flame regions is done by MP and OMP processing. The method achieves a very high accuracy rate of 99.5% with a very low false alarm rate of 0.003%.

A method for real-time video fire flame and smoke detection that is based on foreground image accumulation and optical flow technique is proposed in [68] by Yu et al. The process works by using accumulation images, which are calculated using the foreground images that are extracted using frame differential method. There are two parameters that are used in the foreground image accumulation, to differentiate flame candidate areas from that of smoke. The flame regions are recognized by a statistical model built by foreground accumulation image, while the optical flow is calculated and a motion feature discriminating model is used to recognize smoke regions. The method achieves good performance.

In the case of videos, the methods as proposed in [63] by Xuan Truong et al. involves a multi-stage approach that is comprised of four stages. The first stage is an adaptive Gaussian mixture model used for detecting moving regions and the second, a fuzzy c-means (FCM) algorithm to segment the candidate fire regions from the moving regions on the basis of the colour of fire. The third stage uses special parameters extracted based on the tempo-spatial characteristics of fire regions. The fourth stage implements a support vector machine (SVM) algorithm using these special parameters to distinguish between fire and non-fire. The dataset is created from several video clips. The proposed method has a very high performance with a true positive rate of 94.78% and a false positive rate of 5.22%.

Yadav et al. in [64] propose an optimized flame detection method wherein motion detection is combined with a method used to detect gray cycle pixels nearby the flame in the image which signifies the occurrences of smoke , in the area identified as fire area earlier through colour detection. The method achieves a high detection rate of 92.31% and a low false detection rate 7.69%, thus giving high performance.

A image based fire flame detection method based on colour analysis has been proposed in [20] by Horng et al. in which the fire flame is detected by using a fast colour based analysis technique. The detection system uses a fire flame colour feature model based on the HSI colour space to be built using 70 training flame images. The system then uses this model to seperate regions in the test videos analysed frame by frame, with fire-like colours. Since, background noise regions with similar fire-like colours may also occur, for the removal of these spurious regions,the image difference method and the invented colour masking technique are applied. The fire flame burning degree is then estimated finally, to give an accurate fire alarm. The method achieves a high detection rate of 97%.

In [38], Marbach et al. discuss the use a method of temporal accumulation of time derivate images (videos) to extract the best candidate fire region. Further analysis is performed on these candidate fire regions to extract features that would be used to detect the fire. The features are used to compute a parameter, fire indicator whose pattern describes the occurrence of fire or not. The candidate fire region is taken from the video as a result of the property of fire flames flickering, and the changing intensity of the light emitted. YUV representation is used and a time derivative of the video images are taken to decide this candidate fire region. The candidate region is fixed and characteristic fire features are extracted based on an predetermined threshold based parameter of pixels known as active pixels. There are three main features extracted from the candidate region known as The luminance of the active pixels, the frequency of active pixels, the amplitude of active pixels. The method has good performance results.

4.3. Motion Detection and Contextual Object Detection Based Smoke and Fire Flame Detection Systems

Jian et al. in [25] propose a method to extract smoke suspected regions by combining two steps segmentation and motion characteristics. Initially, regions of interests (ROIs) with smoke are identified and obtained by using two step segmentation methods. This is based on early smoldering smoke appearing as gray white or white region. In the next step, the suspected smoke regions are detected by combining the two step segmentation and motion detection. Morphological processing is used for extracting the smoke regions finally. The Segmentation method uses the Otsu algorithm and the motion detection of smoke is done though using the ViBe algorithm.

In [27] Jinlan et al. present a fire and smoke detection method based on Surendra background and gray bitmap plane algorithm. The dynamic background based on Otsu adaptive threshold and Surendra background is used and the moving target is extracted by continuous cumulative average of moving target area. Along with using mathematical morphology processing method, colour features is used to extract the smoke area and the fire area is extracted by OHTA colour space and gray bitmap plane algorithm. Experimental results show a very effective fire detection performance of the proposed method.

In [24], Jia et al., propose a novel method of segmenting a smoke region in smoke pixel classication based on saliency detection. The model is based on utilizing colour and motion features. Initially, smoke regions are identified by enhancing the smoke colour nonlinearly. This is followed by using enhanced map and motion map to measure saliency. Then finally, the motion energy and saliency map are utilized to estimate the probable smoke regions. The performance of the proposed algorithm is verified on a set of videos containing smoke obtained from [3] and from videos captured. In the experiments, the method achieves average smoke segmentation precision of 93.0%, and the precision is as high as 99.0% for forest fires.

In [36] Luo et al. propose an efficient smoke detection method by condensing video. The method found that smoke trajectories carry special characteristics such as smooth streamline, right-leaning line, fixed source, low-frequency and vertical–horizontal ratio. The effectiveness of the proposed method is evaluated on a dataset of various videos consisting of captured videos, videos from Korea CVPR Lab [29], videos from YouTube and videos from [3]. Experimental results the method differentiates objects which are often mistaken by other algorithms, accurately. The ratio of correct detection is obtained as 83.0%.

In [42], Razmi et al. propose a vision based flame identification system to detect an occurrence of fire in the video. The system uses the methodology of a motion based background estimator, in which the first frames of the video is used to estimate the background image. Then, it is used to subtract the background from each video frame to produce foreground images and only highlights the portion of the background that is shown by the moving objects. After that, Background Subtraction is applied to detect any movement in the video frames. Edge detection is then followed to detect the variance of colour of the pixels. Prewitt edge detector is used for edge detection. The edge detection system compares the colour difference in the image and constructs an edge based on it. The flame's edges would now be shown which is used to analyze the shape and type of fire for alerting fire and smoke detection. The system uses a created image dataset and has a good performance on the dataset.

In [73], Zhaa et al. propose a method that combines context-aware framework with visual smoke detection. Initially, smoke is detected which is followed by contextual objects detection, so as to form consistent goal pairs between smoke and remarkable objectives to estimate the smoke scene. Characteristics of colour histogram and fuzzification is used in smoke detection. The contextual objects are detected using Hoff transform in certain regions where smoke is detected at the center, in order to verify the occurrence of fire. Fusion of these two descriptors is the key in decreasing fire alarm false positives, as implemented through this method. The system attains a fire alarm precision rate of 87.6% on a custom dataset.

4.4. Adaboosted and Infrared based Smoke and Fire Flame Detection Systems

Recently in [59] Wu et al. propose a video based fire smoke detection system using robust AdaBoost. Camera sensors are used to detect fire and smoke videos, which are used to extract static features such as texture, wavelet, colour, edge orientation histogram, irregularity and dynamic features including motion direction, change of motion direction and motion speed. These features are then used to train and test with different combinations. An AdaBoost (RAB) classifier is proposed for improving the training and classification accuracy for detecting smoke and fire. The image datasets are created from pre-created datasets and captured images. The fire smoke detection system gives a high performance based on extensive test experiments with resulting accuracy of 91.25%, detection rate of 99.69% and a very low false alarm rate of 0.31%.

In [32], Lee et al. use the property of smoke and fire having different shapes and colours. They propose a fully connected system which uses two features and an Adaboost algorithm in a linear combination for developing a strong classifier. The local histogram feature by gradient and bin, local binary pattern value and projection vectors for each cell is calculated. According to the histogram magnitude, adapted weighting value is applied to improve recognition rate. To preserve the local region and shape feature which have edge intensity, a normalization sequence is used. For the extracted features, an Adaboost algorithm for strong classification is used. The dataset is created from images and in experiments the model obtains a good recognition rate of 86.19%.

The image flame detection systems use a continuous imaging and pattern recognition of Infrared (IR) images. These images are processed to remove the background noises which may be falsely identified as fire-based flames. For the working of a similar model, the system as patented by Chan et al. [4] uses a Silicon (Si) Charge-Coupled-Device (CCD) array to detect the IR images and filtered by a narrow band IR filter centered at 1140 nm wavelength to remove false alarms in the field of view. The system processes the image in the digital format first by thresholding and then binarizing the images and then performing a pattern recognition by using statistical evaluation methods. The system has a good performance.

5. Deep Learning Model based Fire, Flame and Smoke Detection Systems

5.1. Convolutional Neural Networks (CNNs)

CNNs are being extensively deployed across applications that require image recognition tasks due to their high efficiency and performance with good recognition rates. However, each CNN model varies depending on the implementation technique and as a result, performance varies. CNNs are the most preferred due to the automatic feature learning of smoke and flames specific feature. Figure 4 depicts a general smoke and fire flame detection architecture using CNN. CNNs used in the field of smoke and flame detection are given below. Figure 5 shows a sample detection of fire flame and smoke in the image.

5.1.1. Dual Channel CNN In [74], Zhang et al. propose a Dual-Channel Convolutional Neural Network (DC-CNN) using transfer learning for detecting smoke images. On the first channel of the network, an AlexNet network with transfer learning which is used to extract generalized features, is designed as the main framework of the entire network. The extracting of detailed and specific features is done on the convolutional neural network on the second channel. The two channels of the network are trained separately and their features are fused in the concat layer for robustness detection of smoke in the images. The network achieved a detection of over 99.33% on a publicly available dataset obtained from [13] and other sources.

Earlier in [17], Gu et al. devise a smoke detecting deep dual-channel neural network (DCNN). The proposed model consists of dual channels of deep subnetworks. In the first subnetwork, sequentially connected multiple convolutional layers and max-pooling layers are present. Batch normalization layer is selectively appended to each convolutional layer for reducing overfitting and making training faster. The first subnetwork is shown to be good at extracting the detailed information of smoke, such as texture. In the second subnetwork, in addition to the convolutional, batch normalization, and max-pooling layers, two important components are added, one is the skip connection for avoiding the vanishing gradient and improving the feature propagation, while the other is the global average pooling for reducing the number of parameters and tackling overfitting. The second subnetwork can capture the base information of smoke, such as contours. A concatenation operation is then applied to combine the aforementioned two deep subnetworks to complement each other. Experimental results conducted on the publicly available smoke detection database in [8] verify the proposed DCNN's high detection rate that is above 99.5% on average. Furthermore, this DCNN only employs approximately one-third of the parameters needed by the comparatively tested deep neural network. This model network presented an excellent detection rate of over 99.5%, with lesser parameters required by the network.

In another method of utilizing Dual CNNs, Pundir et al. in [40] propose a robust method for smoke detection which is based on a dual deep learning framework. The proposed architecture makes use of framework based on Deep Convolutional Neural Networks. The first deep learning framework is used for extracting image based features from smoke patches, using the superpixel algorithm. These features consists of smoke texture, smoke colour, sharp edge detection and perimeter disorder analysis. The second deep learning framework,used for extraction of motion based features like smoke moving region, growing region and rising region. Here, an Optical flow method is used to capture the random motion of smoke. These are then input into Deep CNN for extracting motion

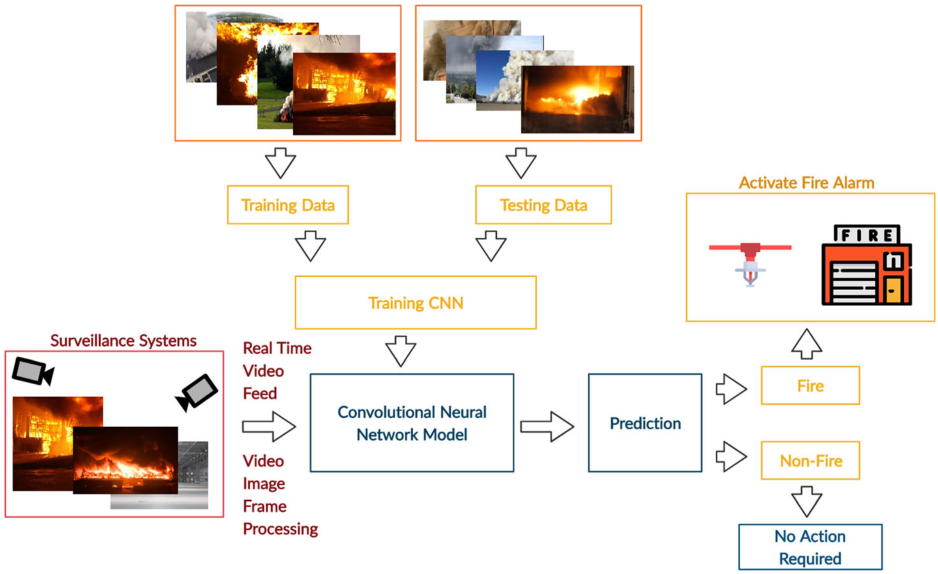

Figure 4. A basic architecture of CNN based fire and smoke detection and alarm systems.

based features. Both the features are used to train a Support Vector Machine and end to end classification, that is the CNN classifier. This powerful network has an detection rate of 97.49%, trained and tested on a dataset that is created from various videos and images. Thus, the network can handle challenging imaging conditions.

Since fire and smoke vary in texture and as a result of the variance of the captured images in terms of colour, angle and lightning conditions, the usage of a Dual Channel CNN is very advantageous for the purpose of detecting and identifying them. The usage of two CNNs allows a more comprehensive feature extraction process, which in turn allows the trained model to detect and identify smoke and fire with a higher accuracy and tolerance to variance as a result of the deeper and detailed feature learning by the model. However, the usage of two CNNs increases the model size and complexity and as a result is compute intensive to train as well as deploy.

5.1.2. Dark Channel CNN In [46], Shi et al. propose to detect smoke by combining dark channel image input and a relative concise convolutional neural network (CNN). To differentiate between the smoke and the background, the dark channel of the image is used. The system has a superior performance with an accuracy rate of 98.41% trained on a dataset created using multiple images.Smoke is traditionally detected through the usage of colour and shape features. However, detection of smoke through using these features is not consistent in multiple scenarios. Here, a Dark Channel transform of the image helps significantly by differentiating

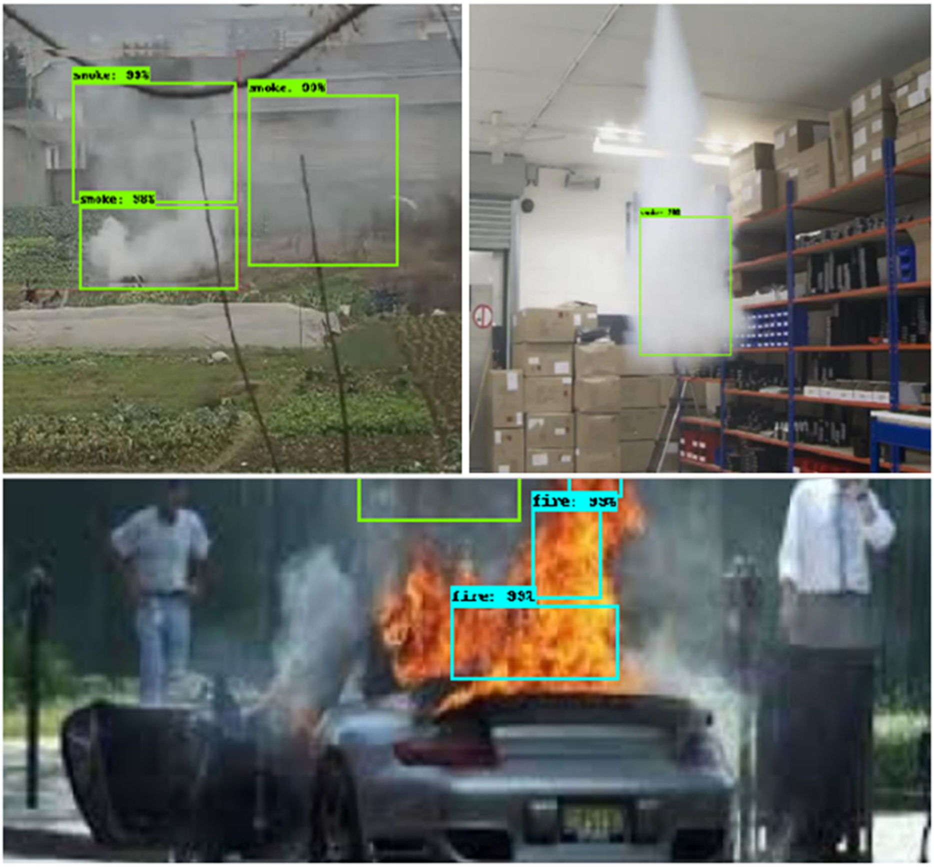

Figure 5. Sample detection of fire flames and smoke in images.

the background and the foggy regions of the images, emphasizing on the smoke, which then can be easily detected by the CNN. Although this method achieves a good performance in terms of detection, Dark Channel transform of the image to be checked is still an overhead for the machine that is performing the detection process.

5.1.3. Spatio-Temporal CNN Hu et al. propose in [21] for a video smoke detection, spatial-temporal based convolutional neural network and a real-time detection architecture utilizing a multitask learning strategy for recognizing smoke and estimating optical flow thereby capturing inter-frame motion features and intra-frame appearance features simultaneously. This overcomes one of the main CNN disadvantages that is avoiding inter-frame motion features and thereby achieves a 97% detection rate with 3.5% false alarm rate trained and tested on a custom created dataset with images frames obtained from videos.

With many methods focusing only on the static features of smoke, one primary disadvantage that is identified is high accuracy rates comes at a cost of significantly high false alarm rates as a result of the high difficulty to distinguish smoke in images. Here, the spatio temporal CNN, effectively, demolishes this disadvantage as it uses both the static motion features of smoke from multiple frames in videos to distinguish it from other similar objects. This ensures a superior accuracy rate while giving a very low false detection rate. Since the spatio temporal CNN is based on using consecutive video frames to detect and distinguish smoke, processing in real time is pretty intensive and thus difficult.

5.1.4. Deep Multi-Scale CNN Yuan et al. propose a design based on a basic block of convolutional neural networks (CNNs) and stack these for a novel deep multi-scale CNN (DMCNN) for smoke recognition in [71]. Several parallel convolutional layers with the same number of filters but different kernel sizes are used as the basic blocks for scale invariance. Every convolutional layer is succeeded by batch normalization for normalizing the output of the convolutional layer. The basic blocks sum up all normalized output from multi-scale parallel layers and then activates the resulting sum as the final output of the block. To fully extract scale invariant features, eleven basic blocks are cascaded and then by a global average pooling and a 2D fully connected layer to construct the DMCNN. Experimental results through training and testing the model on a custom dataset consisting of images conclude that the model achieves more than 97.28% detection rate and 98.63% accuracy rate and lesser than 0.48% false alarm rate.

In yet another approach to tackle translation, scale, rotation and illumination invariance problems that is posed by using of fire and smoke images, the DMCNN method effortlessly solves this by using a multi-structure block based kernel approach. This solves the key issue of scale invariance and ensures high accuracy rates while limiting the false alarm rate to low levels. An inherent disadvantage of using this approach is the complexity of the model as a result of the deep layers and multiple filters that are implemented. As a result, the model might not be able to get deployed or used on low-end machines.

5.1.5. Faster R-CNN In [28], Kim et al. propose a deep learning-based fire detection method using a video sequence imitating the human fire detection process. A Faster Region-based Convolutional Neural Network (R-CNN) is used to detect the suspected regions of fire (SroFs) and of non-fire based utilizing on their spatial features. Then using a Long Short-Term Memory (LSTM), the summarized features within the bounding boxes in successive frames are accumulated to classify whether there is a fire or not in a short-term period. The decisions for successive short-term periods are then taken and combined in the majority voting for the final decision in a long-term period. This is supplemented with the calculation of the areas of both flame and smoke and their temporal changes are taken into account to interpret the dynamic fire behavior with the final fire decision and detection. This system achieves a 97.92% accuracy and a very low 2.47% false detections on a custom dataset created from thousands of images obtained from videos and some available datasets.

In an earlier technique, Junying et al. in [72] propose an object detection method based on deep convolution neural network for smoke detection. Initially, feature extractors such as Inception Net and Resnet are substituted in various neural network object detectors for Faster R-CNN, Single Shot MultiBox Detector (SSD) and Region based Fully Convolutional Networks (R-FCN). Next, the object detection algorithm parameters are optimized on the MSCOCO dataset and further trained conducted on the smoke detection dataset. The model trained on the datasets achieves a good speed and accuracy with mean average precison (mAP) resulting in 56.04%.

The Faster R-CNN based methods achieve excellent accuracy rates and fast performance as a result of the optimization of the feature extraction and region identification techniques through the use of a Region Proposal Network (RPN). However, this process of using a RPN also suffers from the failure to identify the features as a result of invariance created through multiple imaging angles, lighting and scale giving rise to significantly higher false detection rates while requiring significant computation cost and hence cannot perform as effectively in real time.

5.1.6. Recurrent Neural Networks (RNNs) Another convolutional neural network-based algorithm proposed by Yin et al., in [66] is a novel video-based smoke detection method via Recurrent Neural Networks (RNNs). Initially, the algorithm uses deep convolutional motion-space networks to capture the space and motion context information. Next, a temporal pooling layer and RNNs are used to train the smoke model to detect smoke in videos. Dataset is created from videos with variations in illuminations and weather conditions. The network model has a high performance achieving true positive rates of over 95% and true negative rates of over 97%.

In [12], Filonenko et al. propose a combination of a convolutional neural network (CNN) and recurrent neural network (RNN) to detect the smoke in space and time domains. The CNN automatically builds the low level features whereas the RNN finds the relation between the features in different frames of the same event. Dataset is created with image frames from videos and the network achieves an high accuracy of 91.41%.

The Recurrent Neural Network based approaches extract efficient and effective features which form robust smoke features, from multiple domains in adjacent frames, using which smoke is detected with high accuracy and low false detection rates. But, the addition of RNNs to the multiple CNNs present increases the complexity of the network which slows down the frames that can be processed significantly. It also requires intensive computational power to process.

5.1.7. Deep CNNs Zhong et al. in [75] propose a novel flame detection algorithm based on Deep CNN. Initially, a candidate target area extraction algorithm is proposed for the recognition of the suspected flame area. The extracted feature maps of the candidate areas are classified by the deep neural network model. The dataset is created with videos and the method achieves a good performance with accuracy rate in experiments reaching 97.64% and false positive rate reaching 4.9%.

In another method of implementing smoke detection using Deep convolutional neural networks, Yin et al. in [67] propose a novel deep normalization and convolutional neural network with 14 layers to implement automatic feature extraction and classification of smoke. In this type of CNN, the traditional convolutional layers of a CNN are replaced with normalization and convolutional layers to make the training process faster and boost the performance of detection of smoke. The network achieved an excellent performance with detection rate of over 96.37% and false alarm rate below 0.60% using a dataset of images captured and created for training and testing.

In [45], Sharma et al. propose to use deeper Convolutional Neural Networks for fire detection in images and also enhance these networks with fine tuning based on a fully connected layer. Two pretrained state-of-the-art Deep CNNs, VGG16 and Resnet50, are used to develop the fire detection system. The Deep CNNs are tested on a specifically created dataset with the system giving a good accuracy rate of 90% and performance.

Tao et al. [50] propose a deep convolutional neural network wherein the network is trained end to end from raw pixel values to automatically extract features from images and detect smoke. The network achieves a very high 99.56% accuracy, 99.4% detection rate with 0.44% false alarm rate on a dataset obtained from Yuan et al. [70].

In [60], Xu et al. propose a smoke detection method based on two state-of-the-art fast detectors, a single-shot multi-box detector, and a multi-scale deep convolutional neural network, using synthetic smoke image samples. Domain adaptation is incorporated into the fast detectors for training a strong detector with the synthetic smoke images. A series of branches with the same structure as the detection branches are integrated into the fast detectors for domain adaptation. To optimize the model of the adapted detectors and make the model learn a domain-invariant representation smoke detection, an adversarial training strategy is used. Domain discrimination, confusion, and detection are used in the iterative training methodology. The performance of the proposed approach is very high.

Deep CNNs typically employ more layers to deepen the feature extraction process to learn more complex features for the process of identification and distinguishing of smoke and fire. As a result of using more complex features, false detection rates can be contained to very low levels. The addition of more layers, however, adds to the weight of the network and complexity, thereby increasing processing time and limiting deployment use cases.

5.1.8. Basic CNNs In [62], Xu et al. propose a fire detection method based on using colour features, wavelet analysis and convolutional neural networks. Initially, using a colour segmentation method, the candidate region of flame is extracted and the candidate region of smoke is generated by the background fuzzy model based on wavelet analysis. Then this candidate region is filtered by the trained CNN model, and location for the position of flame and smoke in the picture is determined. A dataset was created from images for training and testing. The method has a good performance.

Luo et al. in [37] propose a smoke detection algorithm based on the motion characteristics of smoke and convolutional neural networks (CNN). Initially, to detect the suspected smoke regions, a moving object detection algorithm based on background dynamic update and dark channel priori is used. Then the suspected regions features are extracted automatically by CNN on which smoke identification is performed. The method improves the detection rate to an impressive 99.8%, false alarm rate to 0.31% and accuracy to 99.7% on datasets created from [70] and other images.

With a simplistic structure, basic CNNs based approaches are smaller in model size, give quick results and offer versatile deployment on even low-end machines. The tradeoffs for these benefits are the low accuracy rates and high false alarm rates on real case scenarios due to the shallow feature learning by the CNN.

5.1.9. Deep Saliency Network Another novel video smoke detection method is proposed in [61] by Xu et al. is based on deep saliency network. Using visual saliency detection, most important object regions in an image can be highlighted. To extract the smoke saliency map, the pixel-level and object-level salient convolutional neural networks are combined. The deep feature map and the saliency map are combined to predict the existence of smoke in an image. A complete framework is proposed to detect salient smoke and predict the existence of smoke, for application in video smoke detection, with the system achieving excellent performance with the best accuracy rate of 98.12% on a dataset created from images.

The introduction of a combination of object level, pixel level saliency maps through the deep saliency network enables highly accurate detection of smoke regions in image, while remaining invariant to major factors. However, these saliency maps, produced through corresponding saliency CNNs and usage of a fully convolutional network (FCN) makes the architecture complex. As a result, the model size is increased and with the increase in layers, it is computationally expensive to process.

6. Challenges and Potential Suggestions

Table 1 summarizes all the works and compares each algorithm used in smoke and fire flame detection. Figure 6 briefly mentions all the algorithms and models discussed in this survey paper. The inference that can be arrived from this comprehensive analysis of the techniques used for fire and smoke detection is that the recent technologies that are used, notably the usage of convolutional neural networks have led to great progress and has increased the early detection rate by significant amounts. There is, however, a scope for improvement due to certain existing challenges which are briefed below.

6.1. Existing Challenges

6.1.1. Limited to no Existing Fire and Smoke Datasets The first and foremost important challenge that greatly hinders the development of deep learning methods, convolutional neural networks in the field of fire flame and smoke detection is the limited and or lack of curated datasets that can be used as a benchmark and for training and testing of the neural network. So far, most of the research work that has been done is on datasets that were created with images and videos that were captured by the author(s) and or taken from publicly available datasets and other sources. The annotation of fire, flame and smoke in the videos and images has been a hugely resource consuming task. Due to this and various other reasons, datasets are posing as a huge obstacle in the efficient and fast development of specialized deep learning techniques. Incidentally, we can turn these challenges into opportunities for future research. Specifically, we can leverage effective data augmentation techniques for ample data generation and develop automated annotation methods. Devise automated methods for quantification of smoke and fire flame detection and research design/development of CNN based solutions for simultaneous multi-spot detection. In the coming section, we provide potential future research directions for handling these challenges.

6.1.2. Data Augmentation For automatic data generation, an important direction of future research will be to devise unique data augmentation methods for data generation. Also, it will be vital to develop automated annotation methods for facilitating weakly supervised learning as manual annotation would not be feasible on such a big scale.

6.1.3. Quantifying Smoke and Fire Flame Intensity Provided ample grading data from domain experts, there is a need to design methods that can interpolate on that data and learn to quantify the intensity of smoke and fire flame. This problem can be solved with effective data augmentation techniques, and especially generative adversarial networks may prove to be quite useful for this problem, due to their recent success in similar settings. According to this proposal, the generator network generates a rating for the smoke and fire flame intensity, where the discriminator network attempts to discriminate it in terms of whether it is at which level—low, medium and heavy.

6.1.4. Simultaneous Multi-Spot Detection To counter the challenge of simultaneous multi-spots of smoke and fire flame detection, it is vital to resort to CNNs, owing to their inherent capabilities and success in recent literature. CNN's are an inherent choice for simultaneous detection of multi-class categories, which is considered a fact after their triumph on large scale data having thousands of classes. Hence, the design and development of CNN based machine vision systems for robust visual inspection of leather and hides is an important future research direction.

For challenging cases, where an image scene contains several spots of smoke and fire flame having a high degree of variability, the traditional CNN based methods may not obtain the best results. This can be achieved through an ensemble of CNNs.

6.1.5. Performance Issues As a resulting factor of the issue of non-availability of standard datasets, one method cannot be judged as better than other and

Table 1
Recent Algorithms Proposed for Flame or Fire and Smoke Detection

Years	Authors	Basic methodology	Models	Performance	Key methodology points
2020	Zhang et al.	Deep Learning	Dual-Channel Covolution Neural Network (DC-CNN)	99.33% detection rate	Latest Dual Channel CNN approach with one channel using transfer learning and AlexNet, used for extracting generalized features and the other channel used to detect specific features
2018	Hu et al.	Deep Learning	Spatio-Temporal based Convolutional Neural Network	97% detection rate	For video smoke detection, used a multitask learning strategy by estimating optical flow and capturing inter-frame motion features and intra-frame appearance features
2018	Yin et al.	Deep Learning	Recurrent Neural Networks (RNNs)	95% true positive rate	A deep convolutional motion space network for capturing space and motion context information. RNNs and temporal pooling model used for training
2019	Xu et al.	Deep Learning	Deep Saliency Network	98.12% detection rate	Smoke saliency map extracted using a combination of pixel and object level salient CNNs
2017	Luo et al.	Deep Learning	Motion Characteristics and CNNs	99.8% detection rate	Moving object detection algorithm deployed to identify suspected regions. CNN feature extraction and identification performed
2018	Zhong et al.	Deep Learning	Deep CNNs	97.64% accuracy rate	Candidate target extraction algorithm used for suspected flame area recognition and feature map extraction. Deep CNN model used for classification of these feature maps
2017	Filonenko et al.	Deep Learning	CNNs with Recurrent Neural Networks (RNNs)	91.41% accuracy rate	CNN is used for building low level features, RNN finds the relation between features in different frames of the same event
2018	Yuan et al.	Deep Learning	Deep Multi-Scale CNN (DMCNN)	97.28% detection rate	A DMCNN is built with several parallel convolutional layers with same number of filters but different kernel sizes. Global averaging pooling is used for extracting scale invariant features from the cascaded basic blocks
2018	Russo et al.	Image Processing	Local Binary Pattern (LBP)	93.27% detection rate	Approximate Median Filtering applied to subtract background from input frame. Shape based filtering method applied to obtain region of interest. LBP values and histograms are built from region of interest to construct feature vectors. Bhattacharyya coefficient applied for verification
2019	Gong et al.	Image Processing	Smoke and Fire Features, Multifeature fusion of flame	95.29% accuracy rate	Motion detection and colour detection done at preprocessing. Centroid of flame region calculated based on an developed stabilization algorithm. Area, shape and spatial variability extracted and SVM used for prediction

employed in the detection process. Also, due to the usage of customized datasets, one method might be better in detecting fire alone, but not smoke and vice-versa. The reduction in false positive cases and improvement in accuracy is still awaited.

6.1.6. Interoperability Issues The fact that the system cannot be trusted to give the fire suppression systems total control is also a main problem. This system is completely autonomous, and works on camera inputs and takes decision on the algorithm. If suppose, the system considers a white object or heavy fog as smoke or an orange object as fire flame, the system would falsely trigger the suppression systems causing unexpected damage. If coupled with the existing sensor based systems directly, it would either render both of the systems inefficient and unreliable.

6.2. Potential Suggestions

To improve the early detection of fires, which will most certainly reduce the risk it poses for human lives and monetary things, researchers can develop new methodologies for fire and smoke detection centered on CNNs. Some of the potential suggestions for making the methodologies better are:

1. Designing CNNs with completely new architectures with smoke and fire detection focus.
2. Developing CNNs in similar areas.
3. Advanced image processing techniques added into CNNs.
4. Building hybrid CNNs with each CNN specialising on a specific area of fire and smoke detection.
5. Using transfer learning.

To engage the fire suppression systems and alert fire tenders, it is highly required that these systems be integrated with the existing smoke and fire sensors for improving the detection rate and provide a non-failure system in any situation. As far as the detection of fire goes, it is necessary that the CNNs detect both fire and smoke and prevent fog or any other smoke like particles such as clouds and fire like orange lights, get falsely detected as smoke or fire. More focus should be engaged in the reduction of these false detections for increasing the robustness of and trust on these vision-based fire and smoke detection systems.

7. Conclusion

The early precise detection of smoke or fire is the best approach to identify and tackle a fire that already has started , before it gets out of hand. To this specific purpose, many CNNs and image processing based methods and systems have been proposed for detecting smoke and or fire in images and videos. Convolutional neural networks, which are a very significant and important deep learning frameworks have been analysed in brief, with their application and advantages to this specific detection task discussed.

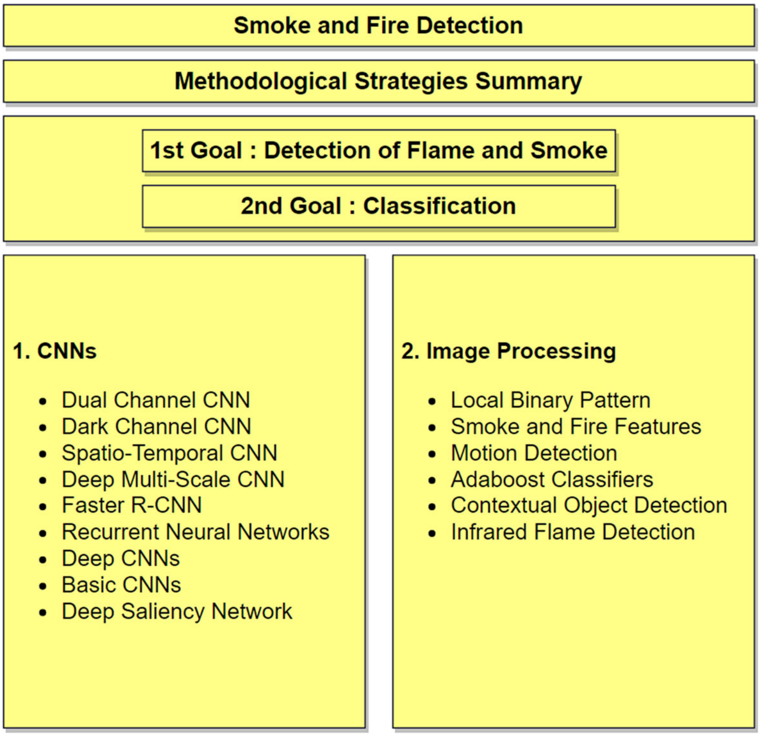

Figure 6. Summary of the different categories of smoke and flame detection methodologies.

A comprehensive analysis of these CNNs and image processing methods and systems has been presented in this paper detailing their methodologies, datasets and their performances. The methods and systems that have been discussed are notable works published from 1996 onwards.

The basic inference that can be made is that the application of CNNs in detecting smoke or fire in images and videos has yielded notably significant high performances. These existing methods and systems have drawbacks and challenges with them. Thus, the improvisation of algorithms in CNNs promises a hope for improving the early detection and controlling of fires, and related detection based areas of research such as object detection and identification as well. In this work, we highlighted the challenges that exist in the design and development of CNN based solutions for smoke and fire flame detection, where ample training data, quantification of smoke and fire flame intensity and high variability of multi-spots are some of the greatest challenges. We also presented research directions for fellow researchers that should be investigated in the future to overcome these challenges and enable advancements in this very important area of research.

References

1. Brownlee J (2020) How do convolutional layers work in deep learning neural networks? Machine Learning Mastery. https://machinelearningmastery.com/convolutional-layers-for-deep-learning-neural-networks/, Accessed 12 Feb 2020
2. Calderara S, Piccinini P, Cucchiara R (2010) Vision based smoke detection system using image energy and color information. Mach Vision Appl 22(4):705–719. https://doi.org/10.1007/s00138-010-0272-1
3. Cetin AE (2008) Visifire dataset. Bilkent EE Signal Processing Group. http://signal.ee.bilkent.edu.tr/VisiFire/, Accessed 23 Feb 2020
4. Chan WS, Burge JW (1996) Imaging flame detection system. US Patent 5937077A, 25 Apr 1996
5. Chen T, Yin Y, Huang S, Ye Y (2006) The smoke detection for early fire-alarming system base on video processing. In: 2006 international conference on intelligent information hiding and multimedia, pp 427–430. https://doi.org/10.1109/IIH-MSP.2006.265033
6. Chunyu Y, Jun F, Jinjun W, Yongming Z (2009) Video fire smoke detection using motion and color features. Fire Technol 46(3):651–663. https://doi.org/10.1007/s10694-009-0110-z
7. DeepQuestAI (2019) Fireflame dataset. GitHub. https://github.com/DeepQuestAI/Fire-Smoke-Dataset. Accessed 22 Feb 2020
8. Di Lascio R, Greco A, Saggese A, Vento M (2014) Improving fire detection reliability by a combination of videoanalytics. In: Image analysis and recognition, pp 477–484. https://doi.org/10.1007/978-3-319-11758-4_52
9. Evarts B (2020) Fire loss in the united states during 2018. NFPA. https://www.nfpa.org/-/media/Files/News-and-Research/Fire-statistics-and-reports/US-Fire-Problem/osFireLoss.pdf. Accessed 10 Feb 2020
10. Enis AC, Kosmas D, Benedict G, Nikos G, Osman G, Habiboglu YH, Toreyin BU, Steven V (2013) Video fire detection—review. Digital Signal Process 23(6):1827–1843. https://doi.org/10.1016/j.dsp.2013.07.003
11. Filonenko A, Hernáindez DC, Jo K (2015) Smoke detection for static cameras. In: 2015 21st Korea–Japan joint workshop on frontiers of computer vision (FCV), pp 1–4. https://doi.org/10.1109/FCV.2015.7103719
12. Filonenko A, Kurnianggoro L, Jo KH (2017) Smoke detection on video sequences using convolutional and recurrent neural networks. In: Computational collective intelligence. Springer, Berlin, pp 558–566. https://doi.org/10.1007/978-3-319-67077-5_54
13. Foggia P, Saggese A, Vento M (2015) Real-time fire detection for video-surveillance applications using a combination of experts based on color, shape, and motion. IEEE Trans Circuits Syst Video Technol 25:1545–1556. https://doi.org/10.1109/TCSVT.2015.2392531
14. Gao Y, Cheng P (2019) Forest fire smoke detection based on visual smoke root and diffusion model. Fire Technol 55(5):1801–1826. https://doi.org/10.1007/s10694-019-00831-x
15. Gaur A, Singh A, Kumar A, Kumar A, Kapoor K (2020) Video flame and smoke based fire detection algorithms: a literature review. Fire Technol 56(5):1943–1980. https://doi.org/10.1007/s10694-020-00986-y
16. Gong F, Li C, Gong W, Li X, Yuan X, Ma Y, Song T (2019) A real-time fire detection method from video with multifeature fusion. Comput Intell Neurosci . https://doi.org/10.1155/2019/1939171

17. Gu K, Xia Z, Qiao J, Lin W (2020) Deep dual-channel neural network for image-based smoke detection. IEEE Trans Multimed 22(2):311–323. https://doi.org/10.1109/tmm.2019.2929009
18. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR), pp 770–778. https://doi.org/10.1109/CVPR.2016.90
19. He K, Zhang X, Ren S, Sun J (2016) Identity mappings in deep residual networks. In: Computer vision—ECCV 2016. Springer International Publishing, Cham, pp 630–645. https://doi.org/10.1007/978-3-319-46493-0_38
20. Horng WB, Peng JW (2008) A fast image-based fire flame detection method using color analysis. Tamkang J Sci Eng 11:273–285
21. Hu Y, Lu X (2018) Real-time video fire smoke detection by utilizing spatial-temporal convnet features. Multimed Tools Appl 77(22):29283–29301. https://doi.org/10.1007/s11042-018-5978-5
22. Jadon A (2019) Firenet dataset. GitHub. https://github.com/arpit-jadon/FireNet-LightWeight-Network-for-Fire-Detection. Accessed 22 Feb 2020
23. 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 preprint arXiv:1905.11922
24. Jia Y, Yuan J, Wang J, Fang J, Zhang Q, Zhang Y (2015) A saliency-based method for early smoke detection in video sequences. Fire Technol 52(5):1271–1292. https://doi.org/10.1007/s10694-014-0453-y
25. Jian W, Wu K, Yu Z, Chen L (2018) Smoke regions extraction based on two steps segmentation and motion detection in early fire. In: MIPPR 2017: pattern recognition and computer vision, international society for optics and photonics, SPIE, vol 10609, pp 281–288. https://doi.org/10.1117/12.2285697
26. Jiang X, Hu C, Fan Z, Zhang P (2015) Research on flame detection method by fusion feature and sparse representation classification. Int J Comput Commun Eng 5(4):238–245. https://doi.org/10.17706/ijcce.2016.5.4.238-245
27. Jinlan L, Lin W, Ruliang Z, Chengquan H, Yan R (2016) A method of fire and smoke detection based on surendra background and gray bitmap plane algorithm. In: 2016 8th international conference on information technology in medicine and education (ITME), pp 370–374. https://doi.org/10.1109/ITME.2016.0089
28. Kim B, Lee J (2019) A video-based fire detection using deep learning models. Appl Sci 9:2862. https://doi.org/10.3390/app9142862
29. KMU-CVPR (2012) Kmu fire & smoke database. KMU CVPR Lab. https://cvpr.kmu.ac.kr/Dataset/Dataset.htm. Accessed 22 Feb 2020
30. Ko B, Kwak JY, Nam JY (2012) Wildfire smoke detection using temporospatial features and random forest classifiers. Opt Eng 51(1):017208-1–017208-10. https://doi.org/10.1117/1.oe.51.1.017208
31. Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. Commun ACM 60(6):84–90. https://doi.org/10.1145/3065386
32. Lee Y, Kim T, Shim J (2017) Smoke detection system research using fully connected method based on adaboost. J Multimed Inf Syst 4(2):79–82. https://doi.org/10.9717/JMIS.2017.4.2.79
33. Liping Z, Hongqi L, Fenghui W, Jie L, Ali S, Hong Z (2018) A flame detection method based on novel gradient features. J Intell Syst 29(1):773–786. https://doi.org/10.1515/jisys-2017-0562

34. Liu Z, Yang X, Liu Y, Qian Z (2019) Smoke-detection framework for high-definition video using fused spatial- and frequency-domain features. IEEE Access 7:89687–89701. https://doi.org/10.1109/access.2019.2926571
35. Liu ZG, Yang Y, Ji XH (2015) Flame detection algorithm based on a saliency detection technique and the uniform local binary pattern in the ycbcr color space. Signal Image Video Process 10(2):277–284. https://doi.org/10.1007/s11760-014-0738-0
36. Luo S, Yan C, Wu K, Zheng J (2015) Smoke detection based on condensed image. Fire Safety J 75:23–35. https://doi.org/10.1016/j.firesaf.2015.04.002
37. Luo Y, Zhao L, Liu P, Huang D (2018) Fire smoke detection algorithm based on motion characteristic and convolutional neural networks. Multimed Tools Appl 77:15075–15092. https://doi.org/10.1007/s11042-017-5090-2
38. Marbach G, Loepfe M, Brupbacher T (2006) An image processing technique for fire detection in video images. Fire Saf J 41(4):285–289. https://doi.org/10.1016/j.firesaf.2006.02.001
39. Matlani P, Shrivastava M (2017) A survey on video smoke detection. Inf Commun Technol Sustain Dev . https://doi.org/10.1007/978-981-10-3932-4_22
40. Pundir AS, Raman B (2019) Dual deep learning model for image based smoke detection. Fire Technol 55(6):2419–2442. https://doi.org/10.1007/s10694-019-00872-2
41. Qixing Z (2018) Sklfs dataset. Fire Detection Research Group. http://smoke.ustc.edu.cn/datasets.htm. Accessed 22 Feb 2020
42. Razmi SM, Saad N, Asirvadam VS (2010) Vision-based flame analysis using motion and edge detection. In: 2010 international conference on intelligent and advanced systems, pp 1–4. https://doi.org/10.1109/ICIAS.2010.5716222
43. Rouse M (2015) Matlab definition. WhatIs.com. https://whatis.techtarget.com/definition/MATLAB. Accessed 24 Feb 2020
44. Russo AU, Deb K, Tista SC, Islam A (2018) Smoke detection method based on LBP and SVM from surveillance camera. In: 2018 international conference on computer, communication, chemical, material and electronic engineering (IC4ME2), pp 1–4. https://doi.org/10.1109/IC4ME2.2018.8465661
45. Sharma J, Granmo OC, Goodwin M, Fidje JT (2017) Deep convolutional neural networks for fire detection in images. In: Engineering applications of neural networks. Springer, Cham, pp 183–193. https://doi.org/10.1007/978-3-319-65172-9_16
46. Shi X, Lu N, Cui Z (2019) Smoke detection based on dark channel and convolutional neural networks. In: 2019 5th international conference on big data and information analytics (BigDIA), pp 23–28. https://doi.org/10.1109/BigDIA.2019.8802668
47. Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556
48. Szegedy C, Wei Liu, Yangqing Jia, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: 2015 IEEE conference on computer vision and pattern recognition (CVPR), pp 1–9. https://doi.org/10.1109/CVPR.2015.7298594
49. Tang T, Dai L, Yin Z (2017/09) Smoke image recognition based on local binary pattern. In: Proceedings of the 2017 5th international conference on mechatronics, materials, chemistry and computer engineering (ICMMCCE 2017). Atlantis Press, pp 1118–1123. https://doi.org/10.2991/icmmcce-17.2017.199
50. Tao C, Zhang J, Wang P (2016) Smoke detection based on deep convolutional neural networks. In: 2016 international conference on industrial informatics—computing technology, intelligent technology, industrial information integration (ICIICII), pp 150–153. https://doi.org/10.1109/ICIICII.2016.0045

51. Toulouse T, Rossi L, Campana A, Celik T, Akhloufi MA (2017) Computer vision for wildfire research: an evolving image dataset for processing and analysis. Fire Saf J 92:188–194. https://doi.org/10.1016/j.firesaf.2017.06.012
52. Vijayalakshmi SR, Muruganand S (2018) Fire recognition based on sensor node and feature of video smoke. In: 2018 international conference on advanced computation and telecommunication (ICACAT), pp 1–7. https://doi.org/10.1109/ICACAT.2018.8933629
53. Wang L, Li A (2017) Early fire recognition based on multi-feature fusion of video smoke. In: 2017 36th Chinese control conference (CCC), pp 5318–5323. https://doi.org/10.23919/ChiCC.2017.8028197
54. Wang S, He Y, Yang H, Wang K, Wang J (2017) Video smoke detection using shape, color and dynamic features. J Intell Fuzzy Syst 33(1):305–313. https://doi.org/10.3233/jifs-161605
55. Wang Y, Wu A, Zhang J, Zhao M, Li W, Dong N (2016) Fire smoke detection based on texture features and optical flow vector of contour. In: 2016 12th world congress on intelligent control and automation (WCICA), pp 2879–2883. https://doi.org/10.1109/WCICA.2016.7578611
56. Wikipedia (2020a) Convolutional neural network. Wikipedia. https://en.wikipedia.org/wiki/Convolutional_neural_network. Accessed 10 Feb 2020
57. Wikipedia (2020b) Cuda. Wikipedia. https://en.wikipedia.org/wiki/CUDA. Accessed 24 Feb 2020
58. Woodford C (2020) How smoke detectors work. ExplainThatStuff. https://www.explainthatstuff.com/smokedetector.html. Accessed 10 Feb 2020
59. Wu X, Lu X, Leung H (2018) A video based fire smoke detection using robust adaboost. Sensors 18(11):3780. https://doi.org/10.3390/s18113780
60. Xu G, Zhang Q, Liu D, Lin G, Wang J, Zhang Y (2019) Adversarial adaptation from synthesis to reality in fast detector for smoke detection. IEEE Access 7:29471–29483. https://doi.org/10.1109/ACCESS.2019.2902606
61. Xu G, Zhang Y, Zhang Q, Lin G, Wang Z, Jia Y, Wang J (2019) Video smoke detection based on deep saliency network. Fire Saf J 105:277–285. https://doi.org/10.1016/j.firesaf.2019.03.004
62. Xu Z, Wanguo W, Xinrui L, Bin L, Yuan T (2019b) Flame and smoke detection in substation based on wavelet analysis and convolution neural network. In: ICIAI 2019: proceedings of the 2019 3rd international conference on innovation in artificial intelligence, pp 248–252. https://doi.org/10.1145/3319921.3319962
63. Xuan Truong T, Kim JM (2012) Fire flame detection in video sequences using multistage pattern recognition techniques. Eng Appl Artif Intell 25(7):1365–1372. https://doi.org/10.1016/j.engappai.2012.05.007
64. Yadav G, Gupta V, Gaur V, Bhattacharya M (2012) Optimized flame detection using image processing based techniques. Indian J Comput Sci Eng 3:202–211
65. Ye W, Zhao J, Wang S, Wang Y, Zhang D, Yuan Z (2015) Dynamic texture based smoke detection using surfacelet transform and hmt model. Fire Safety J 73:91–101. https://doi.org/10.1016/j.firesaf.2015.03.001
66. Yin M, Lang C, Li Z, Feng S, Wang T (2018) Recurrent convolutional network for video-based smoke detection. Multimed Tools Appl 78(1):237–256. https://doi.org/10.1007/s11042-017-5561-5
67. Yin Z, Wan B, Yuan F, Xia X, Shi J (2017) A deep normalization and convolutional neural network for image smoke detection. IEEE Access 5:18429–18438. https://doi.org/10.1109/ACCESS.2017.2747399

68. Yu C, Mei Z, Zhang X (2013) A real-time video fire flame and smoke detection algorithm. Procedia Eng 62:891–898. https://doi.org/10.1016/j.proeng.2013.08.140
69. Yuan F (2011) Video-based smoke detection with histogram sequence of lbp and lbpv pyramids. Fire Safety J 46(3):132–139. https://doi.org/10.1016/j.firesaf.2011.01.001
70. Yuan F (2020) Video smoke detection dataset. State Key Lab of Fire Science, University of Science and Technology of China. http://staff.ustc.edu.cn/~yfn/vsd.html
71. Yuan F, Zhang L, Wan B, Xia X, Shi J (2018) Convolutional neural networks based on multi-scale additive merging layers for visual smoke recognition. Mach Vision Appl 30(2):345–358. https://doi.org/10.1007/s00138-018-0990-3
72. Zeng J, Lin Z, Qi C, Zhao X, Wang F (2018) An improved object detection method based on deep convolution neural network for smoke detection. In: 2018 international conference on machine learning and cybernetics (ICMLC). vol 1, pp 184–189. https://doi.org/10.1109/ICMLC.2018.8527037
73. Zhaa X, Ji H, Zhang D, Bao H (2018) Fire smoke detection based on contextual object detection. In: 2018 IEEE 3rd international conference on image, vision and computing (ICIVC), pp 473–476. https://doi.org/10.1109/ICIVC.2018.8492823
74. Zhang F, Qin W, Liu Y, Xiao Z, Liu J, Wang Q, Liu K (2020) A dual-channel convolution neural network for image smoke detection. Multimed Tools Appl . https://doi.org/10.1007/s11042-019-08551-8
75. Zhong Z, Wang M, Shi Y, Gao W (2018) A convolutional neural network-based flame detection method in video sequence. Signal Image Video Process 12(8):1619–1627. https://doi.org/10.1007/s11760-018-1319-4
76. Zhou Z, Shi Y, Gao Z, Li S (2016) Wildfire smoke detection based on local extremal region segmentation and surveillance. Fire Saf J 85:50–58. https://doi.org/10.1016/j.firesaf.2016.08.004