Hindawi
Complexity
Volume 2021, Article ID 5554316, 12 pages
https://doi.org/10.1155/2021/5554316

Research Article

A Statistical Image Feature-Based Deep Belief Network for Fire Detection

Dali Sheng,[1] Jinlian Deng,[2] Wei Zhang,[2] Jie Cai,[2] Weisheng Zhao,[2] and Jiawei Xiang[3]

[1]*Department of Police Command and Tactics, Zhejiang Police College, Hangzhou 310053, China*
[2]*Department of Mechanical Engineering, Zhejiang Institute of Mechanical & Electrical Engineering, Hangzhou 310053, China*
[3]*College of Mechanical and Electrical Engineering, Wenzhou University, Wenzhou 325035, China*

Correspondence should be addressed to Jinlian Deng; dengjinlian@zime.edu.cn and Jiawei Xiang; wxw8627@163.com

Received 2 February 2021; Revised 4 July 2021; Accepted 29 July 2021; Published 5 August 2021

Academic Editor: Osnat Mokryn

Detecting fires is of significance to guarantee the security of buildings and forests. However, it is difficult to fast and accurately detect fire stages in complex environment because of the large variations of the fire features of color, texture, and shapes for flame and smoke images. In this paper, a statistic image feature-based deep belief network (DBN) is proposed for fire detections. Firstly, for each individual image, all the statistic image features extracted from a flame and smoke image in time domain, frequency domain, and time-frequency domain are calculated to construct training and testing samples. Then, the constructed samples are fed into DBN to classify the multiple fire stages in complex environment. DBN can automatically learn fault features layer by layer using restricted Boltzmann machine (RBM). Experiments using the benchmark data of three groups of fire and fire-like images are classified by the present method, and the classification results are also compared with those commonly used support vector machine (SVM) and convolutional deep belief networks (CDBNs) to manifest the superiority of the classification accuracy.

1. Introduction

Accurately and timely detection of fires is important to save life, property, and economic losses. Recently, fire detection methods have been developed to monitor forest fires, civil infrastructure, and industrial fires [1–6]. 23,535 fire incidents of buildings in 18 cities around the world in the year 2017 were reported by the International Association of Fire and Rescue Services (CTIF) [7]. Therefore, accurately and timely detection of fires using sensors is of great significance to protect the social security [8].

Fire features, such as heat, gas, flame, and smoke, are the most commonly used in fire detection techniques to monitor fire. Point sensors are the most directly and commonly used fire sensing techniques for engineering applications [9–11]. However, point sensors might lead to high false positives using signal processing techniques to extract features under complicated fire environment [11]. Fire images are the effective and direct representations of fires under complex outside environment, which has attracted numerous researchers to construct modern fire alarm systems. Smokes or flames are the first signal when there is a fire. Therefore, effective smoke or flame detection plays an important role in fire detection.

The color, location, height, and optical density are the features of a smoke for the combination of gases, airborne solid, and liquid particulates [12–15]. However, color-based smoke detection is always reliable for the confusion of gray or black color smoke with nonsmoke pixels of other objects [16]. The fuel and oxidant exothermic reactions create flames with various colors, flickering motions, and dynamic texture feature. The spatial and temporal wavelet analysis combined with Weber contrast analysis and color segmentation was presented to enhance the dynamic texture feature of flames [17]. However, though flames have distinct color features, failed to recognize fire similar color objects in the backgrounds, which lead to a difficult problem on how to choose a suitable color model. In general, once there are fires occurred in buildings and forest, they can be reflected in the measured smokes and flames with certain characteristics.

However, the smoke or flame features are usually too weak to be observed due to the strong background noises and other disturbances.

Traditional machine learning methods are support vector machine (SVM) [18], fuzzy neural networks [19], smoke segmentation-based local binary pattern Silhouettes coefficient variant (LSPSCV) [20], color segmentation-based radial basis function (RBF) nonlinear Gaussian kernel-based binary SVM [21], color segmentation-based fuzzy model [22], and fire frame segmentation-based Markov random field [23]. However, most existing studies on fire detection train models using smoke or flame from videos are often difficult to track more complicated smoke situations. Moreover, traditional machine learning methods only consider fire smoke or flame in an ideal background without too much disturbances. Therefore, these artificial intelligent fire detection models still suffer from obvious shortcomings to learn complicated nonlinear relationships and thus have limited representation capacity.

Deep learning, as a new branch of machine learning, has been shown the excellent ability in learning features from raw images [24]. The most obvious property of deep learning models is the multiple layer structures. With multiple hidden layers stacked hierarchically, the deep learning model can realize very complicated transformation and abstraction of raw images [25–28]. Deep belief network (DBN) is a kind of the generative deep learning model with powerful feature learning ability [29].

Pundir and Raman adopted DBN to recognize fires in an accurate and robust way for varieties of scenarios such as wildfire-smoke video, hill base smoke video, and indoor or outdoor smoke videos [30]. DBN was carefully designed to extract the nonlinear features for a better description of the important trends in a combustion process [31]. Kaabi et al. [32] developed a Gaussian mixture model (GMM) and the corresponding energy attitude of smoke region based on RGB rules to preprocess smoke or flame images for the DBN classification. Wang et al. [33] extended DBNs to classify coal-fired boilers to further predict NOx emission. The main advantage of the intelligent diagnosis solution is that DBN does not rely on manual feature extraction and selection.

In this paper, a new fire detection method based on DBN is proposed to detect fires using smoke and flame images. Firstly, all the statistic image features extracted from the raw images are obtained to characterize the fire status. Secondly, the training and testing samples are constructed by the statistic image features. Finally, DBN is employed to identify the fire status using smoke or flame images in the strong background noises and other disturbances. Comparing with other commonly used methods, the classification accuracy ratio of the proposed method is manifested by using the open-access experimental investigations.

2. A Basic Theory of DBN

DBN is a kind of neural network with multiple hidden layers which allow DBN to learn complex functions which could then complete data transformation and abstraction by successive learning process. The main architecture of DBN includes an ensemble of stacked RBMs. Every two-layer neural network composes an RBM. Here, DBN is composed of three stacked RBMs and an output layer. The learning process of a DBN consists of two stages: one is pretraining every individual RBM layer by layer in an unsupervised manner; the other is to fine tune the whole network using the back propagation algorithm in a supervised manner. More details can be seen in References [34–36].

3. Statistical Image Feature-Based DBN

3.1. Overview. The diagram of the fire detection method is shown in Figure 1. The general procedures are summarized as follows:

Step 1. Predefine smoke or flame patterns.

Step 2. Collect the raw images of the smoke or flame patterns using the image capture system, respectively. Moreover, raw images of unknown status are also collected using the same image capture system.

Step 3. Calculate feature descriptor from the raw images using feature descriptors, i.e., color moments in HSV space, statistical features of gray-scale image, statistical features of gray-level co-occurrence matrix, local binary pattern histogram, and wavelet transform decomposition (WTD)-based statistical features of gray-scale image, which will be discussed in details late.

Step 4. Construct the training samples for the known smoke or flame patterns and the testing samples for unknown smoke or flame patterns using the feature descriptors.

Step 5. Train each individual RBM layer by layer and then fine tune DBN (DBN learning).

Step 6. Obtain the unknown smoke or flame patterns after DBN learning process.

3.2. Color Descriptor

3.2.1. Color Moments in HSV Space. RGB space is the commonly used color representation. However, researchers [37, 38] found that HSV space is more close to human visual perception and is friendly to the base visual feature extraction. The HSV refers to hue (0 to 360°, H), saturation (0 to 1, S), and value (0 to 1, V). In HSV space, colors are represented as combination of the three channels. For any channel of HSV color space, the first-order moment (denoted as M_1) is calculated by

$$M_1(i) = \frac{1}{N}\sum_{j=1}^{N} p_{i,j}, \tag{1}$$

the second-order moment (denoted as M_2) is

$$M_2(i) = \sqrt{\left(\frac{1}{N}\sum_{j=1}^{N}\left(P_{i,j} - \text{mean}(i)\right)^2\right)}, \tag{2}$$

and the third order moment (denoted as M_3) is

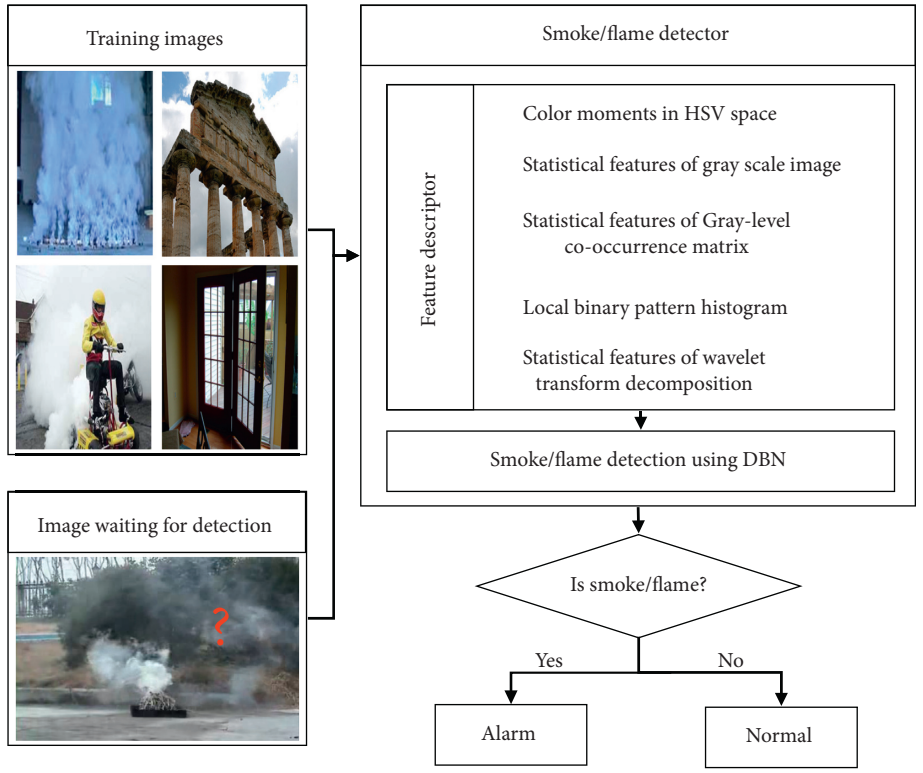


Figure 1: Flowchart of the proposed method.

$$M_3(i) = \sqrt[3]{\left(\frac{1}{N}\sum_{j=1}^{N}\left(P_{i,j} - \text{mean}(i)\right)^3\right)}. \quad (3)$$

Figure 2(a) shows a smoke image in RGB space, and Figure 2(b) shows the same image in HSV space. It does not show much difference in both spaces, except for a certain region of concern. Figure 2(c) shows the detection result of the smoke region in RGB space, and Figure 2(d) shows the same result in HSV space. It can be seen that HSV space shows a better property. Therefore, the first, second, and third moments M_1, M_2, and M_3 in HSV space (H, S, V) will generate 9×1 statistical features vector *F1*.

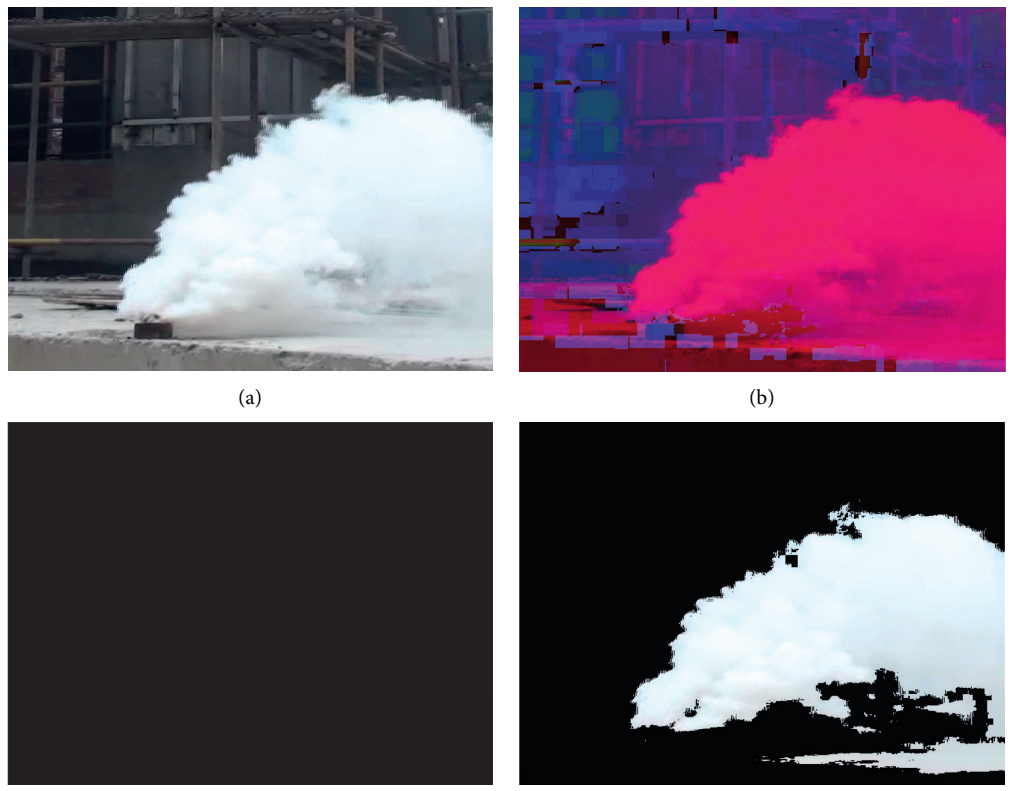
3.2.2. Statistical Features in Gray-Scale Image. Typical statistical features in gray-scale image [39] are considered in this part, as listed in Table 1. Images in gray-level bins of frequency distribution are represented by gray-level histogram, and it counts the pixels which are similar and stores it. The histogram analyses statistics of a single pixel with a certain gray level. It can reflect the changes occurred in the translation, rotation, and angle on individual parts of an image. Entropy is calculated based on the gray-level histogram. Therefore, 7 statistical features of gray-scale images will generate 7×1 statistical features vector *F2*.

3.3. Texture Descriptor

3.3.1. Gray-Level Co-Occurrence Matrix. Texture can be characterized by the direction, adjacent interval, and change range of images. There will be a certain gray-level relationship between two pixels separated by a certain distance in the image space, i.e., the spatial correlation characteristics of the gray level in the image. The gray-level co-occurrence matrix (GLCM) [40] is a tool to describe such relations by measuring the co-occurrence frequencies among pairs of pixels of images in gray level. The GLCM is obtained by the statistics of two pixels with a certain distance in the image. The elements in the diagonals of the GLCM tend to have larger values when the image is composed of pixel blocks with similar gray values, while the elements that deviate from the diagonal will have larger values when the image pixels change locally.

GLCM reflects the comprehensive information about images in gray-level co-occurrence matrix. Through the gray-level co-occurrence matrix, we can analyze the local pattern and arrangement rules of the image. In order to describe the texture, we usually obtain statistics on the basis of the GLCM. Table 2 gives the five typical statistics used in GLCM, which formed 5×1 statistical features vector *F3*.

3.3.2. Local Binary Pattern Histogram. Local binary pattern (LBP) is an operator that is used to describe the local texture features of an image [41]. It has the characteristics of multiresolution, invariant gray scale, and rotation; therefore, it is a good measure for image feature extraction, as shown in Figure 3. LBP is often alongside with histograms to represent images as a feature vectors, i.e., LBPH. As a visual descriptor, LBP can be obtained directly using scikit-image [42] package, which needs at least 2



(a) (b) (c) (d)

Figure 2: (a) A smoke image in RGB. (b) A smoke image HSV. (c) Target region in RGB. (d) Target region in HSV.

Table 1: Statistical features of gray-scale image.

Symbol	Name	Equation
s_1	Arithmetic mean	$x_m = \sum_{i=1}^{N} x_i/N$
s_2	Median	x_{qva}
s_3	Standard deviation	$x_{ava} = \sum_{i=1}^{N} \lvert x(i)\rvert/N$
s_4	Skewness	x_{std}
s_5	Kurtosis	$x_{\text{std}} = \sqrt{\sum_{i=1}^{N} (x(i) - xm)^2/N}$
s_6	Entropy	x_{sra}
s_7	Energy	$x_{\text{sra}} = (\sum_{i=1}^{N} \sqrt{\lvert x(i)\rvert}/N)^2$

Note. P_n is the n-th pixel in the gray-scale image; N is the total pixel number. In the equation of median, P denotes the ordered list of P_n, and x_{var} and $x_{\text{var}} = \sum_{i=1}^{N} (x_i - x_m)^2/N$ denote the floor and ceiling functions. In the equation of entropy, k denotes the number of gray-scale value, and f_k denotes frequency of k-th value.

Table 2: Typical features of GLCM.

Symbol	Name	Equation
g_1	Angular second moment	x_{ske}
g_2	Contrast	$x_{\text{ske}} = \sum_{i=1}^{N} (x(i) - x_m)^3/N$
g_3	Inverse difference moment	x_P
g_4	Entropy	$x_P = \max\lvert x(i)\rvert$
g_5	Correlation	x_{kur}

Note. $x_{\text{kur}} = \sum_{i=1}^{N} (x(i) - x_m)^4/N x_{\text{std}}^4$ denotes the element of GLCM, where $x_{\max} = \max(x(i))$ and $\text{SF} = x_{\text{rms}}/(\sum_{i=1}^{N} \lvert x(i)\rvert/N)$.

parameters, the number of circularly symmetric neighbor set points and the radius of circle. Histograms are then extracting using the generated LBP; we set 18 bins to get the final histogram of the raw image, i.e., a feature vector, denoted as $b1$ to $b18$.

From Figure 3, we can see that different parts of the frequency histogram form a certain texture of the image. Therefore, local binary pattern histogram is chosen here as features for smoke/flame detection. $F4$ is an 18×1 statistical features vector combined by the frequency histogram.

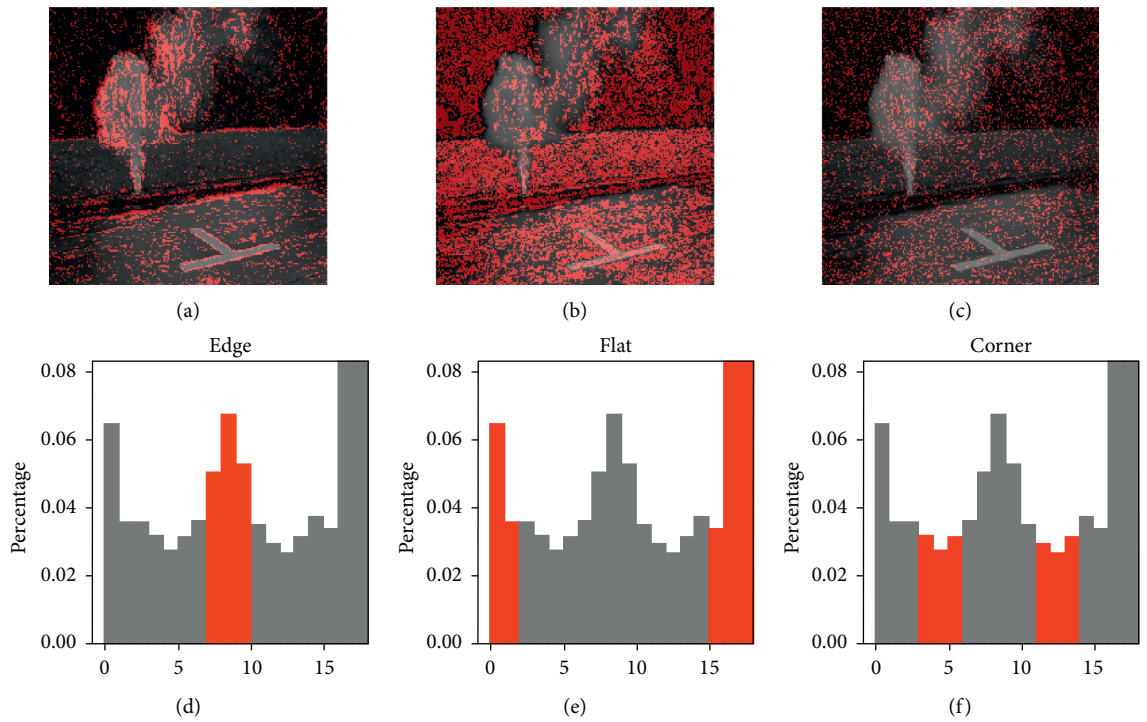


Figure 3: Edge-like, flat, and corner-like regions and their LBP histograms.

3.4. WTD-Based Statistical Features of Gray-Scale Image. WTD can be understood as a successive low-pass and high-pass filters. Therefore, image after WTD is a set of subimages which contains different details of the raw image. Figure 4 shows a 3rd level WTD, L3 is the low resolution approximation of the original image, and H3_h, H3_v, and H3_d; H2_h, H2_v, and H2_d; and H1_h, H21_v, and H1_d are the wavelet subimages of horizontal details, vertical details, and diagonal details at the 3rd level decomposition. Due to its simplicity and efficiency, the Haar wavelet is widely used, especially in the image processing fields such feature detection and image compression. Therefore, the Haar wavelet is chosen in the present method. Figure 4 shows the corresponding subimages after the 3rd level WTD. The final 7 subimages are denoted as L3, H3_h, H3_v, and H3_d; H2_h, H2_v, and H2_d. As can be seen from Figure 5, the raw fire image and its L3, H3_h, H3_v, H3_d, H2_h, H2_v, and H2_d show different details. We calculate the statistical features of the 7 subimages as listed in Table 1. Finally, we get 7 feature vectors for each of the subimage, which are denoted as W1 to W7. In the present investigations, we only focus on lower frequency band, and the subimages of H1_h, H21_v, and H1_d are not considered.

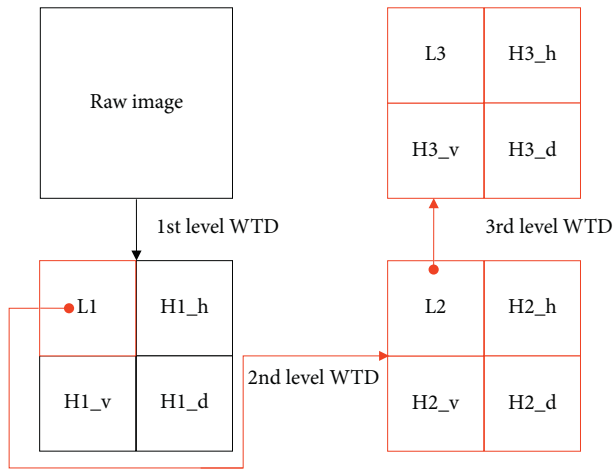


Figure 4: A description of 3-level WTD of an image.

4. Experimental Investigations

4.1. Data and Evaluation Metrics. In our experiment, two case studies on smoke detection and fire detection are carried out. The benchmark data are collected from smoke dataset [43] and fire dataset [44]. In smoke dataset, the first 500 images are characterized by the presence of the smoke and the last 500 images are normal scene images which contain roads, trees, white curtains, lights, jets of water, and so on. In fire dataset, the first 1048 images are characterized by the presence of the fire and the last 1048 images are normal images with natural landscape and some of which contain sunlight that appeared with a large similarity to fires. Table 3 lists the brief information of the two datasets. Figure 6 shows some of the images. We can see that it shows great difference between images, which means it is very challenging to detect smoke from the images.

To evaluate the performance of the present method, we use the same evaluation metrics mentioned in [40], which are detection rate, i.e., the true positive rate (TPR), false alarm rate (i.e., FPR), and average accuracy rate (AAR) given as follows:

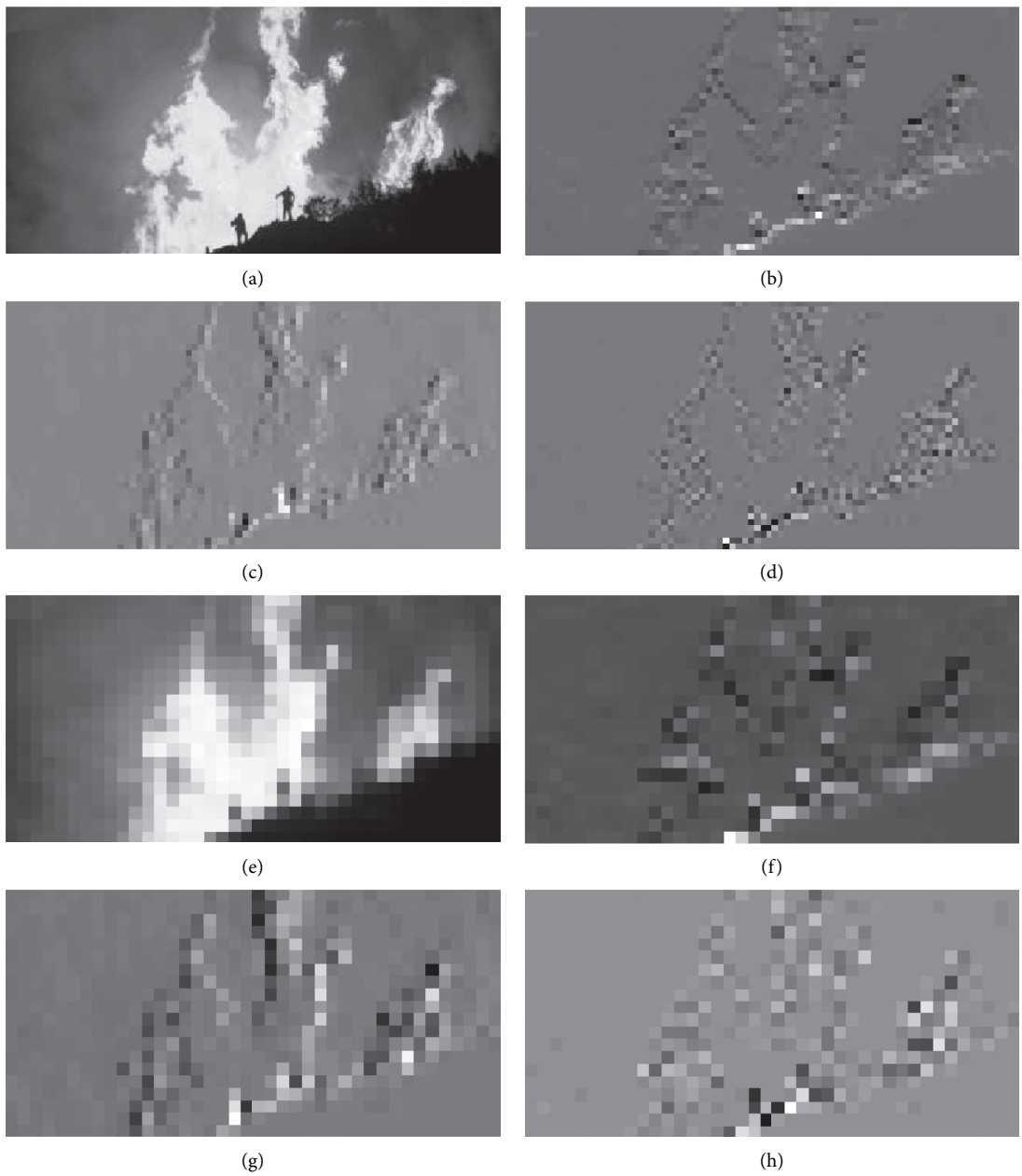


Figure 5: Subimages after WTD: (a) raw; (b) H2_h; (c) H2_v; (d) H2_d; (e) L3; (f) H3_h; (g) H3_v; (h) H3_d.

Table 3: Image datasets for the experiment.

Case	Number of positive image	Number of negative image	Total number
Smoke detection	500	500	1000
Fire detection	1048	1048	2096

$$\mathrm{TPR} = \frac{\mathrm{TP}}{(\mathrm{TP} + \mathrm{FN})}, \tag{4}$$

$$\mathrm{FPR} = \frac{\mathrm{FP}}{(\mathrm{FP} + \mathrm{TN})}, \tag{5}$$

$$\mathrm{AAR} = \frac{\mathrm{TP} + \mathrm{TN}}{(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN})}, \tag{6}$$

in which TP, TN, FP, and FN are the true positive, true negative, false positive, and false negative.

4.2. Result and Discussion. We first apply the model to smoke detection. In the case study, we conduct ablation study w.r.t various feature combinations and comparison study with other methods. The smoke and nonsmoke images are divided into training set and testing set with ratio equals

FIGURE 6: Raw images of the two cases: (a) positive image in smoke detection; (b) negative image in smoke detection; (c) positive image in fire detection; (d) negative image in fire detection.

to 0.6. Features are extracted from the images to feed into the deep belief networks, i.e., F_1, F_2, F_3, F_4, F_5, F_6 (F_1 and F_2), F_7 (F_3 and F_4), F_8 (F_1, F_2, F_3, and F_4), and F_9 (F_1, F_2, F_3, F_4, and F_5). Details are shown in Table 4.

Figure 7 shows the results of the ablation study over three metrics with respect to (w.r.t.) 9 statistical features vectors. Figure 8 is the average accuracy and the standard deviation over 9 statistical features vectors F_1, F_2, F_3, F_4, F_5, F_6, F_7, F_8, and F_9. The combination of color features, textual features, and wavelet features shows the best performance.

Then, we conduct the comparison study with another two machine learning methods. One is the convolutional deep belief networks (CDBNs) [45]. The reason we choose the model for comparison is that it could learn the hierarchical representations from raw images in an unsupervised learning manner without any label information, which is somewhat like the way we extract the statistical features. The other one is the support vector machine (SVM) [46]. Codes are released by the authors. DBN used in the proposed method is simply set to 200-100-50, while the CDBN is a two-layer unsupervised feature learning model followed by a softmax classifier, whose parameters are listed in Table 5. And the kernel type of the SVM is sigmoid, with gamma set to 1 and cost set to 10. The others take the default values. The inputs to the models are all same except for the CDBN.

We take 50 trails to obtain average results of the three evaluation metrics over the three comparison methods, as shown in Figure 9. The metric values of TPR, FPR, and AAR of the proposed method are 97.14%, 5.28%, and 95.88%, respectively. The metric values of TPR, FPR, and AAR of CDBN are 87.61%, 10.64%, and 88.42%, respectively. The metric values of TPR, FPR, and AAR of SVM are 64.39%, 34.63%, and 64.80%, respectively. Table 6 lists the best results among the 50 trails. Besides, the best performance of the average classification accuracy rate over the three methods is 97.68%, 92.72%, and 71.25%, respectively. Compared with the CDBN and SVM models, it can be seen that the present statistic image feature-based DBN obtained the best performance. The comparisons with more advanced models are shown in Table 7. We have to admit that these SOTA models have strong feature learning ability from raw images but may suffer from a large quantity of training parameters and limited training samples, as well as the various backgrounds, diverse angles, and different lights.

Fire detection is conducted in the following part using the proposed method. Same as the previous case in the smoke detection, both ablation study and comparison study are conducted. The fire and nonfire images are also divided into training set and testing set with ratio equals to 0.6. Figure 10 shows the results of the ablation study over three metrics w.r.t. 9 feature combinations. The features are calculated using color descriptor, texture descriptor, and

Table 4: Typical features of GLCM.

Symbol	Name	Element
F_1	Color moments in HSV space	$[M_1, M_2, M_3]$
F_2	Statistical features of gray-scale image	$[s_1, s_2, s_3, s_4, s_5, s_6, s_7]$
F_3	Statistical features of the gray-level co-occurrence matrix	$[g_1, g_2, g_3, g_4, g_5]$
F_4	Local binary pattern histogram	$[b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8, b_9, b_{10}, b_{11}, b_{12}, b_{13}, b_{14}, b_{15}, b_{16}, b_{17}, b_{18}]$
F_5	Statistical features of wavelet transform decomposition	$[W_1, W_2, W_3, W_4, W_5, W_6, W_7]$
F_6	Combinatorial features	$[F_1, F_2]$
F_7	Combinatorial features	$[F_3, F_4]$
F_8	Combinatorial features	$[F_1, F_2, F_3, F_4]$
F_9	Combinatorial features	$[F_1, F_2, F_3, F_4, F_5]$

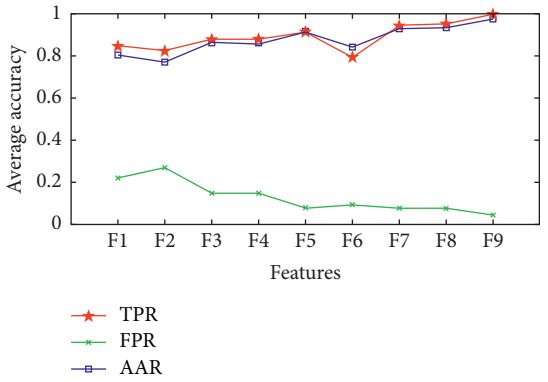


Figure 7: Performance metrics for smoke detection w.r.t combinations of features.

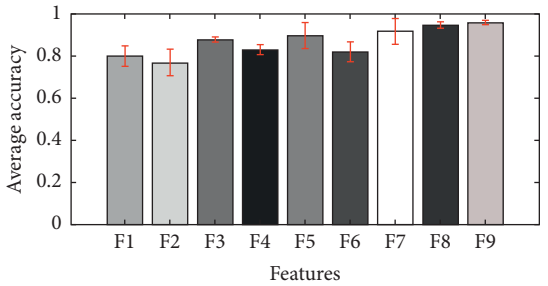


Figure 8: Average accuracy and the std for smoke detection w.r.t combinations of features.

Table 5: Parameters of CDBN.

The 1st layer		The 2nd layer	
Size_of_input	28×28	Size_of_input	10×10
Number_of_filter	9	Number_of_filter	16
Size_of_filter	7	Size_of_filter	5
Size_of_pool	2	Size_of_pool	2
Learning_rate	$3e-4$	Learning_rate	$3e-4$

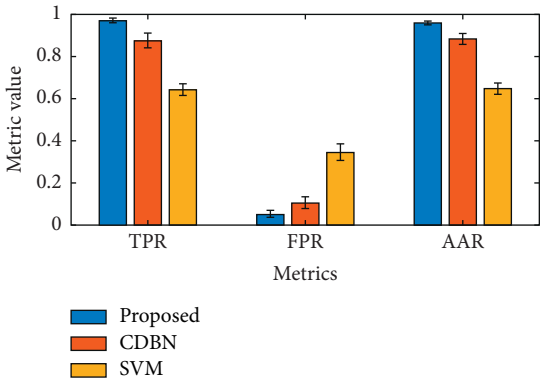


Figure 9: Performance metrics of the comparison study for smoke detection.

Table 6: Precision, recall, and $F1$-score of the comparison study for smoke detection.

Pattern	Proposed (%)			CDBN (%)			SVM (%)		
	Precision	Recall	$F1$-score	Precision	Recall	$F1$-score	Precision	Recall	$F1$-score
Smoke	100	94.87	97.37	89.39	95.68	92.43	71.29	71.64	71.46
Nonsmoke	95.35	100	97.62	96.04	90.23	93.05	71.21	70.85	71.03
Average	97.68	97.44	97.50	92.72	92.96	92.74	71.25	71.25	71.25

Table 7: Precision, recall, and $F1$-score of the comparison study for smoke detection.

Pattern	AlexNet (%)			VGG16 (%)			VGG19 (%)		
	Precision	Recall	$F1$-score	Precision	Recall	$F1$-score	Precision	Recall	$F1$-score
Smoke	93.94	93.00	93.47	96.97	96.00	96.48	98.94	93.00	95.88
Nonsmoke	93.07	94.00	93.53	96.04	97.00	96.52	93.40	99.00	96.12
Average	93.50	93.50	93.50	96.50	96.50	96.50	96.17	96.00	96.00

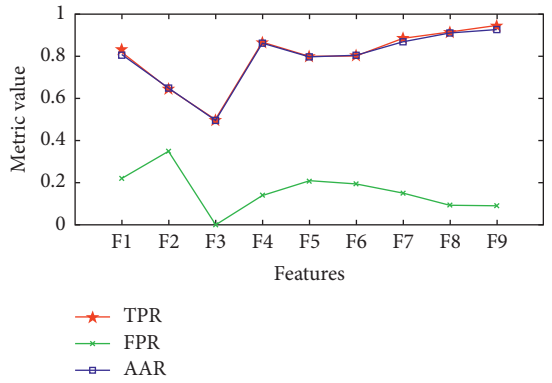


Figure 10: Performance metrics for fire detection w.r.t combinations of features.

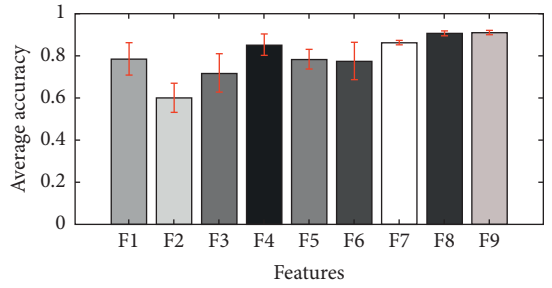


Figure 11: Average accuracy and the std for fire detection w.r.t combinations of features.

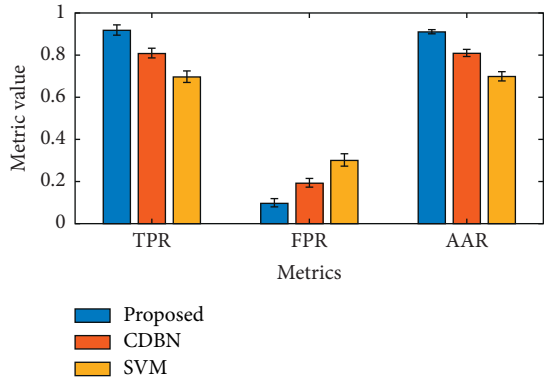


Figure 12: Performance metrics of the comparison study for smoke detection.

Table 8: Precision, recall, and $F1$-score of the comparison study for fire detection.

Pattern	Proposed (%)			CDBN (%)			SVM (%)		
	Precision	Recall	$F1$-score	Precision	Recall	$F1$-score	Precision	Recall	$F1$-score
Fire	94.33	90.76	92.51	84.34	86.27	85.29	75.53	75.35	75.44
Nonfire	90.97	94.47	92.69	84.65	82.54	83.59	74.58	74.76	74.67
Average	92.65	92.62	92.60	84.50	84.41	84.44	75.06	75.06	75.06

Table 9: Precision, recall, and $F1$-score of the comparison study for fire detection.

Pattern	AlexNet (%)			VGG16 (%)			VGG19 (%)		
	Precision	Recall	$F1$-score	Precision	Recall	$F1$-score	Precision	Recall	$F1$-score
Fire	92.06	82.86	87.22	88.94	91.90	90.40	91.35	90.48	90.91
Nonfire	84.42	92.86	88.44	91.63	88.57	90.07	90.57	91.43	91.00
Average	88.24	87.86	87.83	90.28	90.24	90.24	90.96	90.95	90.95

statistical features with DWT. Feature vectors F_1, F_2, F_3, F_4, F_5, F_6, F_7, F_8, and F_9 in the case are the same as those in the previous case.

Figure 11 shows the average accuracy and the standard deviation over 9 feature vectors, and F_9 has the best performance in average accuracy.

In the comparison study of fire detection, we also take 50 trails. Figure 12 shows the average results of the three evaluation metrics. The metric values of TPR, FPR, and AAR of the proposed method are 91.68%, 9.66%, and 90.94%, respectively. The metric values of TPR, FPR, and AAR of CDBN are 80.69%, 19.10%, and 80.79%, respectively. The metric values of TPR, FPR, and AAR of SVM are 69.56%, 29.96%, and 69.76%, respectively. Table 8 lists the best results among the 50 trails in the fire detection. Besides, the best performance of the average classification accuracy rate over the three methods is 92.65%, 84.50%, and 75.06%, respectively. The present statistic image feature-based DBN shows comparable results over CDBN and SVM models. SOTA results are given in Table 9, which shows the similar results as in the smoke detection case.

5. Conclusion

Fires in buildings and forests are difficult to be fast and accurately detected due to the complex environment. A statistic image feature-based DBN is proposed for fire detections to reduce the feature variations influence of color, texture, and shapes for flame and smoke images. Experiments using open-access benchmark data of fire and fire-like images are performed to verify the effectiveness of the present method. The relative optimal combinations of statistic image features are analyzed using the different combinations of statistic image features. The best performance of the average classification accuracy rate using the present method could achieve 97.68 and 92.65%, respectively, for detecting smoke and fire using the benchmark data. Compared with two typical classification models of SVM and CDBN, the classification accuracy rates exceed 4.96%; 26.43% and 8.15%; 17.59%, respectively. The results show that the present method can learn effective flame and smoke features with high accuracy in complicated classifications. Moreover, this method is expected to fast detect fire stages with the help of small samples of statistic image features.

Data Availability

The fire dataset is available in https://mivia.unisa.it/datasets/video-analysis-datasets/fire-detection-dataset/ (accessed on 8 September 2020).

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was funded in part by the Zhejiang Provincial Key Research and Development Project under Grant 2020C03096 and Zhejiang Provincial Science Foundation under Grant GF19F020010.

References

[1] R. Sahal, S. H. Alsamhi, J. G. Breslin, and M. I. Ali, "Industry 4.0 towards forestry 4.0: fire detection use case," *Sensors*, vol. 21, no. 3, p. 694, 2021.

[2] C. Yuan, Y. Zhang, and Z. Liu, "A survey on technologies for automatic forest fire monitoring, detection, and fighting using unmanned aerial vehicles and remote sensing techniques," *Canadian Journal of Forest Research*, vol. 45, no. 7, pp. 783–792, 2015.

[3] A. Gaur, A. Singh, A. Kumar, A. Kumar, and K. Kapoor, "Video flame and smoke based fire detection algorithms: a literature review," *Fire Technology*, vol. 56, no. 5, pp. 1943–1980, 2020.

[4] Y. Luo, L. Zhao, P. Liu, and D. Huang, "Fire smoke detection algorithm based on motion characteristic and convolutional neural networks," *Multimedia Tools and Applications*, vol. 77, no. 12, pp. 15075–15092, 2018.

[5] H. Yin, Y. Wei, H. Liu, S. Liu, C. Liu, and Y. Gao, "Deep convolutional generative adversarial network and convolutional neural network for smoke detection," *Complexity*, vol. 2020, Article ID 6843869, 12 pages, 2020.

[6] R. Wang, Y. Li, H. Sun, and K. Yang, "Multisensor-weighted fusion algorithm based on improved AHP for aircraft fire detection," *Complexity*, vol. 2021, Article ID 8704924, 10 pages, 2021.

[7] N. N. Brushlinsky, M. Ahrens, S. V. Sokolov, and P. Wagner, *Word Fire Statistics Report No 24. Center of Fire Statistics of CTIF*, State Fire Academy of Emercom of Russia, Moscow, Russia, 2020.

[8] A. Gaur, A. Singh, A. Kumar et al., "Fire sensing technologies: a review," *IEEE Sensors Journal*, vol. 19, no. 9, pp. 3191–3202, 2019.

[9] E. G. Bakhoum, "High-sensitivity miniature smoke detector," *IEEE Sensors Journal*, vol. 12, no. 10, pp. 3031–3035, 2012.

[10] J. Gubbi, S. Marusic, and M. Palaniswami, "Smoke detection in video using wavelets and support vector machines," *Fire Safety Journal*, vol. 44, pp. 110–1125, 2009.

[11] X. Liu, S. Cheng, H. Liu, S. Hu, D. Zhang, and H. Ning, "A survey on gas sensing technology," *Sensors*, vol. 12, no. 7, pp. 9635–9665, 2012.

[12] M. J. Hurley, D. T. Gottuk, J. R. Hall et al., *SFPE Handbook of Fire Protection Engineering*, Springer, Berlin, Germany, 2016.

[13] S. Wu, Y. Yang, Z. Fang, and Y. Fang, "Real-time image smoke detection using staircase-searching based dual threshold AdaBoost and dynamic analysis," *IET Image Process*, vol. 9, pp. 849–856, 2015.

[14] J. Rong, D. Zhou, W. Yao, W. Gao, J. Chen, and J. Wang, "Fire flame detection based on GICA and target tracking," *Optics & Laser Technology*, vol. 47, pp. 283–291, 2013.

[15] S. Wang, Y. He, J. J. Zou, D. Zhou, and J. Wang, "Early smoke detection in video using swaying and diffusion feature," *Journal of Intelligent & Fuzzy Systems*, vol. 26, no. 1, pp. 267–275, 2014.

[16] Z. Zhou, Y. Shi, Z. Gao, and S. Li, "Wildfire smoke detection based on local extremal region segmentation and surveillance," *Fire Safety Journal*, vol. 85, pp. 50–58, 2016.

[17] S. Ye, Z. Bai, H. Chen, R. Bohush, and S. Ablameyko, "An effective algorithm to detect both smoke and flame using color and wavelet analysis," *Pattern Recognition and Image Analysis*, vol. 27, no. 1, pp. 131–138, 2017.

[18] D. Xiong and L. Yan, "Early smoke detection of forest fires based on SVM image segmentation," *Journal of Forest Science*, vol. 65, no. 4, pp. 150–159, 2019.

[19] M. Lazarevska, A. T. Gavriloska, M. Laban, M. Knezevic, and M. Cvetkovska, "Determination of fire resistance of eccentrically loaded reinforced concrete columns using fuzzy neural networks," *Complexity*, vol. 2018, Article ID 8204568, , 2018.

[20] Q. Li, H. Liu, J. Zhang, and P. Zeng, "Target segmentation of industrial smoke image based on LBP Silhouettes coefficient variant (LBPSCV) algorithm," *IET Image Processing*, vol. 14, no. 12, pp. 2879–2889, 2020.

[21] M. R. Islam, M. Amiruzzaman, S. Nasim, and J. Shin, "Smoke object segmentation and the dynamic growth feature model for video-based smoke detection systems," *Symmetry*, vol. 12, no. 7, Article ID 1075, 2020.

[22] M. J. Sousa, A. Moutinho, and M. Almeida, "Classification of potential fire outbreaks: a fuzzy modeling approach based on thermal images," *Expert Systems with Applications*, vol. 129, pp. 216–232, 2019.

[23] M. Ajith and M. Martinez-Ramon, "Unsupervised segmentation of fire and smoke from infra-red videos," *IEEE Access*, vol. 7, pp. 182381–182394, 2019.

[24] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[25] J. Toledo-Castro, P. Caballero-Gil, N. Rodriguez-Perez, I. Santos-Gonzalez, C. Hernandez-Goya, and R. Aguasca-Colomo, "Forest fire prevention, detection, and fighting based on fuzzy logic and wireless sensor networks," *Complexity*, vol. 2018, Article ID 1639715, , 2018.

[26] S. Saponara, A. Elhanashi, and A. Gagliardi, "Real-time video fire/smoke detection based on CNN in antifire surveillance systems," *Journal of Real-Time Image Processing*, vol. 18, no. 3, pp. 889–900, 2020.

[27] X. Wu, Y. Park, A. Li, X. Huang, F. Xiao, and A. Usmani, "Smart detection of fire source in tunnel based on the numerical database and artificial intelligence," *Fire Technology*, vol. 57, no. 2, pp. 657–682, 2020.

[28] J. J. Wang, J. Yu, and Z. He, "DECA: a novel multi-scale efficient channel attention module for object detection in real-life fire images," *Applied Intelligence.*, 2021.

[29] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[30] A. S. Pundir and B. Raman, "Deep belief network for smoke detection," *Fire Technology*, vol. 53, no. 6, pp. 1943–1960, 2017.

[31] Y. Liu, Y. Fan, and J. Chen, "Flame images for oxygen content prediction of combustion systems using DBN," *Energy & Fuels*, vol. 31, no. 8, pp. 8776–8783, 2017.

[32] R. Kaabi, M. Bouchouicha, A. Mouelhi, M. Sayadi, and E. Moreau, "An efficient smoke detection algorithm based on deep belief network classifier using energy and intensity features," *Electronics*, vol. 9, no. 9, p. 1390, 2020.

[33] F. Wang, S. Ma, H. Wang, Y. Li, and J. Zhang, "Prediction of NOx emission for coal-fired boilers based on deep belief network," *Control Engineering Practice*, vol. 80, pp. 26–35, 2018.

[34] G. E. Hinton, "A practical guide to training restricted Boltzmann machines," *Lecture Notes in Computer Science*, vol. 7700, pp. 599–619, 2012.

[35] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[36] S. Wang, J. Xiang, Y. Zhong, and H. Tang, "A data indicator-based deep belief networks to detect multiple faults in axial piston pumps," *Mechanical Systems and Signal Processing*, vol. 112, pp. 154–170, 2018.

[37] A. M. Treisman and G. Gelade, "A feature-integration theory of attention," *Cognitive Psychology*, vol. 12, no. 1, pp. 97–136, 1980.

[38] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397–1409, 2013.

[39] W.-S. Tang, S.-H. Jiang, and S.-L. Wang, "Gray scale potential: a new feature for sparse image," *Neurocomputing*, vol. 116, pp. 112–121, 2013.

[40] X. Zhang, J. Cui, W. Wang, and C. Lin, "A study for texture feature extraction of high-resolution satellite images based on a direction measure and gray level co-occurrence matrix fusion algorithm," *Sensors*, vol. 17, no. 7, p. 1474, 2017.

[41] S. Yasmin, R. K. Pathan, M. Biswas, M. U. Khandaker, and M. R. I. Faruque, "Development of a robust multi-scale featured local binary pattern for improved facial expression recognition," *Sensors*, vol. 20, no. 18, p. 5391, 2020.

[42] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias et al., "The scikit-image contributors scikit-image: image processing in Python," *PeerJ*, vol. 2, Article ID e453, 2014.

[43] Z. Yin, B. Wan, F. Yuan, X. Xia, and J. Shi, "A deep normalization and convolutional neural network for image smoke detection," *IEEE Access*, vol. 5, pp. 18429–18438, 2017.

[44] Fire Dataset Available online: https://github.com/LeadingIndiaAI/Forest-Fire-Detection-through-UAV%20-imagery-using-CNNs (accessed on 8 September 2020).

[45] H. Lee, R. Grosse, and R. Ranganath, "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations," in *Proceedings of the 26th annual*

international conference on machine learning, pp. 609–616, Montreal Quebec Canada, June 2009.

[46] C. C. Chang and C. J. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, p. 27, 2011.