

Combining multi-channel color space with local binary co-occurrence feature descriptors for accurate smoke detection from surveillance videos

Author

Alamgir, Nyma, Kien, Nguyen, Chandran, Vinod, Boles, Wageeh

Published

2018

Journal Title

Fire Safety Journal

Version

Accepted Manuscript (AM)

DOI

[10.1016/j.firesaf.2018.09.003](https://doi.org/10.1016/j.firesaf.2018.09.003)

Rights statement

© 2018 Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, providing that the work is properly cited.

Downloaded from

<http://hdl.handle.net/10072/414571>

Griffith Research Online

<https://research-repository.griffith.edu.au>

Combining Multi-Channel Color Space with Local Binary Co-occurrence Feature Descriptors for Accurate Smoke Detection from Surveillance Videos

Nyma Alamgir*, Kien Nguyen, Vinod Chandran, and Wageeh Boles

School of Electrical Engineering and Computer Science, Queensland University of Technology, 2 George St, Brisbane 4000.

Abstract

Fire is one of the most devastating hazards that can cause serious damage to human life, health and property. As smoke is often an initial sign of fire, smoke detection using surveillance cameras is key to providing early alarm in open space environments. In this paper, we propose a new feature extraction method that combines local binary patterns with co-occurrence of texture features in RGB color space to characterize the diverse manifestations of smoke. The proposed RGB color based Local Binary Co-occurrence Patterns (RGB_LBCoP) extracts smoke features from candidate smoke regions which are extracted by Fuzzy C-Means (FCM) algorithm. Subsequently, Support Vector Machine (SVM) is used for training and classification based on these features. The major benefit of the proposed feature descriptor is the ability to incorporate local and global texture properties of smoke along with color information. This property enables the detection of smoke in complex environments and provides insensitivity to illumination changes. For validation, performance of the proposed method is compared with other LBP variants and Grey-level co-occurrence matrix (GLCM). Experimental analysis on publicly available smoke video datasets demonstrates that the proposed algorithm outperforms the other methods by achieving an average True Positive Rate (TPR) of 92.02%.

Keywords: *Smoke Detection, RGB Color Space, Local Binary Co-occurrence Patterns, Local and Global Features, Surveillance Videos*

1. Introduction

Traditional smoke detection sensors which are cheap and simple to use, detect the presence of smoke, heat or radiation by sampling smoke particles, atmospheric temperature and relative humidity [1]. However, these sensors suffer from the propagation delay of the smoke particles as they trigger when sufficient amounts of smoke particles reach near the sensors. Hence, smoke sensors demand conditions like proximity to fires as well as good operating conditions. Further, they are not suitable for application in open space environments. Furthermore, these types of point sensors cannot provide the actual position and size of the fires. In contrast, video cameras can be used to overcome the aforementioned limitations by monitoring the volumes of fire with information about their size and growth rate. Hence, video surveillance cameras, which are widely used in security applications, can be used in fire monitoring systems for open spaces. For open spaces, initially smoke might appear in the surveillance cameras as fire is obstructed by the foliage in forests and infrastructures in built-up areas. Therefore, detecting smoke can give an early warning for fire hazards.

Various approaches have been proposed in recent years with the aim of efficient detection of smoke in video sequences so that it can be applicable for real time applications. One of the main challenges of video based smoke detection is to extract smoke features as it presents diversely with quite chaotic variations in color, shade, motion and density. For this reason, feature extraction has been widely investigated in the literature. Cetin et al [2] summarized the existing smoke and fire detection techniques based on video surveillance cameras and computer vision methods. The authors provided a comprehensive overview about color, motion, flicker, dynamic texture and spatio-temporal descriptor features which are currently available in the state-of-the-art works. Töreyn et al.[3] used temporal and spatial wavelet analysis to detect semi-transparent smoke for a static camera. Gubbi et al. [4] characterized smoke by detailed wavelet features which were extracted from the three levels of wavelet transformation. Then these wavelet features were used by SVM to detect smoke. A model based on illumination invariant color descriptors to detect smoke, was also proposed in [5]. Chunyu et

al. [6] combined color based decision rule and motion features to detect smoke from videos. Qureshi et al. [7] presented QuickBlaze, a vision sensor based smoke and flame detection technique that can be applicable for open or closed indoor and outdoor environments. Rule based thresholding was applied to calculate features like turbulence, growth and flow rate for detecting smoke and fire. Jia et al. [8] also used color and motion features for saliency map calculation in order to detect smoke. Lee et al. [9] and Tao et al. [10] used convolutional neural networks for image based smoke detection. A deep domain based method was proposed by Xu et al. which extracted powerful learned features to detect smoke from videos comprised of synthetic and real images [11]. However, convolutional neural networks require a large amount of data which makes it computationally expensive in both space and time.

Limited application of static texture features for video-based smoke and fire detection has been reported so far in the literature. Prema et al. [12] utilized static texture features along with dynamic features for video-based fire detection. Static features are extracted by a hybrid method combining LBP, GLCM and DWT. Additionally, both static and dynamic texture features were also combined for video-based smoke detection in [13] which suffered from a substantial false alarm rate. It should be noted here that, both [12] and [13] have utilized dynamic features along with the static features to achieve good detection accuracy. This shows the limitation of single use of static features in video-based smoke and fire detection.

Dynamic textures, defined by textures with motion, play an important role in video-based smoke detection as smoke is a moving object. Grey Level Co-Occurrence Matrices (GLCM) were used in [14, 15] to extract texture features for smoke detection. Prema et al. [14] evaluated the performance of smoke detection using surveillance CCTV cameras by incorporating color, spatio-temporal and GLCM texture features. Nguyen et al. [15] also used GLCM features for optical smoke detection. Local Binary Patterns (LBP) and their variants are one of the powerful texture descriptors, which are widely used in the field of image and video processing, as it is insensitive to image rotation and illumination variations and is cheap to compute. Yuan et al. [16] computed the histograms of LBP and local binary patterns based on variance (LBPV) pyramids and then neural network classifier was used to differentiate between smoke and non-smoke regions. A variant of LBP, named Local Binary Motion Patterns (LBMP), was proposed by Zhao et al. [17] to define the dynamic texture features of smoke which is then classified by the Adaboost algorithm. However, the common limitations of these approaches include: i) they were applied on grey scale images while discarding the discriminative color features. ii) They are susceptible to high false alarm rates, because contrast is highly affected by gray scale variations and iii) correlation between LBP features were not considered.

To improve the discriminative power of the LBP texture features, the correlation between two features were considered which has been successfully used in other applications like image classification [18], face and texture recognition [19], smoke image detection [1] and medical image retrieval [20]. Qi et al. [18] proposed a novel Pairwise Rotation Invariant Co-occurrence Local Binary Pattern (PRICoLBP) technique that considered spatial co-occurrence information between two LBP features and then applied it to six different but related applications—texture, material, flower, leaf, food, and scene classification. Nosaka et al. [19] computed the co-occurrence among multiple LBPs that provides more details of the image and successfully applied it to gray scale face and texture recognition. Local Ternary Co-occurrence Patterns (LTCoP) was proposed for medical image retrieval [20] which includes the co-occurrence of similar ternary edges based on the gray values of the center pixel and its neighboring pixels. Yuan et al. [1] proposed an extension of Local Ternary Patterns (LTP) by encoding high order directional derivatives at each pixel of grayscale image. This computationally complex algorithm was applied for image based smoke detection and image classification. However, LTP is sensitive to noise and can be easily affected by intensity changes [21].

Motivated by the higher accuracy for the combination of the co-occurrence with LBPs, this work proposes a new feature extraction technique by computing the co-occurrence of LBP for RGB color space. When spatial correlation between two LBP features is considered, it provides more information for complex structures in an image [18]. In addition, the spatial co-occurrence of two LBPs can provide both local and global information of the image regions that helps to detect subtle textural changes of the smoke regions. Furthermore, when color information is added with the co-occurrence of LBP features, it provides details of spatial information for a complex image texture.

The main contributions of this paper can be summarized as: 1) We have introduced a new feature extraction technique, Local Binary Co-occurrence Patterns for RGB color plane (RGB_LBCoP) which is successfully applied for video-based smoke detection. 2) We consider $LBP_{P,R}^{riu2}$ as it is an excellent measure of image texture that helps to detect complex structures. We calculate the $LBP_{P,R}^{riu2}$ for each of the three channels of RGB color space. In this way, discriminative color features are incorporated with texture features. 3) We measure the co-occurrence of the adjacent LBPs for each channel and compute six co-occurrence features. Then we concatenate the six features of each channel to get the feature vectors of 18 descriptive features along with color. 4) We evaluate the performance of the proposed RGB_LBCoP technique in comparison with other feature extraction techniques like GLCM, LBP, CoLBP, RGB_LBP. In addition, the performance of the proposed technique was compared with other state-of-the-art smoke detection algorithms and VisiFire software.

The rest of the paper is organized as follows. Section 2 contains the background information related to the proposed smoke detection approach. Section 3 briefly describes about the proposed video based smoke detection algorithm. Section 4 gives experimental result analysis with various comparisons. The conclusions of this study are given in Section 5.

2. Background Information

This section provides a brief overview of the algorithms utilized in our proposed smoke detection method.

2.1. Fuzzy c-means algorithm

The Fuzzy c-means (FCM) algorithm is an unsupervised iterative clustering method which produces optimal clustering. It was developed by Bezdek et al. and has been successfully used in pattern recognition and image processing applications [22].

Let an unlabeled data set $X = (x_1, x_2, x_3 \dots x_n)$ represent the intensity of the image pixels where n is the number of pixels. The FCM algorithm tries to cluster the data set X into c number of clusters. The standard objective function is defined by:

$$J_m = \sum_{i=1}^c \sum_{(k=1)}^n \mu_{ik}^m d^2(x_k, v_i) \quad (1)$$

where $d^2(x_k, v_i)$ is the Euclidian distance between the data point x_k and the centroid value v_i of the i th cluster, μ_{ik} is the degree of membership for data point x_k in the i -th cluster. Here parameter m is called fuzzy factor that controls the fuzziness of the resulting partition along with the constraint $m \geq 1$. The optimal number of clusters is produced by local minimization of the objective function, J_m . Objective function is minimized by repeatedly adjusting the values of μ_{ik} and v_i based on the following equations:

$$\mu_{ik} = \left[\sum_{j=1}^c \left(\frac{d^2(x_k, v_i)}{d^2(x_k, v_j)} \right)^{\frac{1}{m-1}} \right]^{-1} \quad (2)$$

$$v_i = \frac{\sum_{k=1}^n \mu_{ik}^m x_k}{\sum_{k=1}^n \mu_{ik}^m} \quad (3)$$

The clustering iterations of FCM algorithm are terminated when the termination condition $\max_{1 \leq i \leq c} \{ \| (v_i^t - v_i^{t-1}) \| \} < \epsilon$ is satisfied. Here v_i^{t-1} is the centre of the previous iteration and ϵ is the termination threshold value. Finally all pixels are distributed into clusters according to the value of degree of membership, μ_{ik} .

2.2. Local Binary Patterns(LBP) operators

An LBP texture analysis operator, proposed by Ojala et al. [23], is a grey-scale invariant texture descriptor. LBP code is computed by comparing the value of center pixel with the value of the neighboring pixels. The LBP operator generates code for each pixel of an image by thresholding its P neighbors in a circle of radius R as per the following equation:

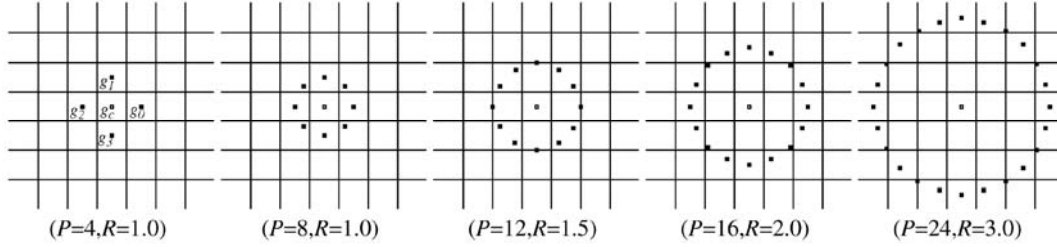


Figure 1: Symmetric neighbors for different values of P and R [23]

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (4)$$

where g_c denotes the gray scale value of the center pixel, g_p is the gray scale value of its neighboring pixel, and s is a function defined by $s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$. The individual LBP code is capable of describing the texture information around the center pixel. Figure 1 shows different combinations of the circularly symmetric pixels on the circumference of a circle with radius, R.

2.3. Support Vector Machines (SVM)

Support Vector Machines (SVM) are based on a supervised learning algorithm proposed by Vapnik [24]. It analyzes data and recognize patterns which help SVM for being popular in various fields of classification and pattern recognition [25]. SVM classifies the data into two classes by finding the best hyperplane that separates the data into two classes. The optimal hyperplane maximizes the margin of separation between itself and the support vectors (minimally separated data points from the two classes).

3. The Proposed Smoke Detection Algorithm

The proposed video based smoke detection algorithm consists of four major stages: 1) detection of moving regions 2) selection of candidate smoke regions 3) feature extraction by proposed color and texture based method, and 4) detection of smoke by Support Vector Machine- (SVM-) based classification. A flowchart of the proposed video based smoke detection algorithm is shown in Figure 2

3.1. Detection of Moving Regions

In surveillance videos, smoke appears as a moving object of constantly changing shape depending on the wind and burning materials. Therefore, the first step of the proposed smoke detection algorithm aims at detecting the moving object within the video sequences. For detecting moving regions, approximate median subtraction method is employed due to its simplicity and higher accuracy [26]. Approximate median method detects the moving objects by subtracting the background. The extracted moving regions usually have some small blobs which are removed using median filtering. This process enhances the quality of the detected moving regions. However, the extracted moving regions include smoke but also other moving objects that can be refined through further processing.

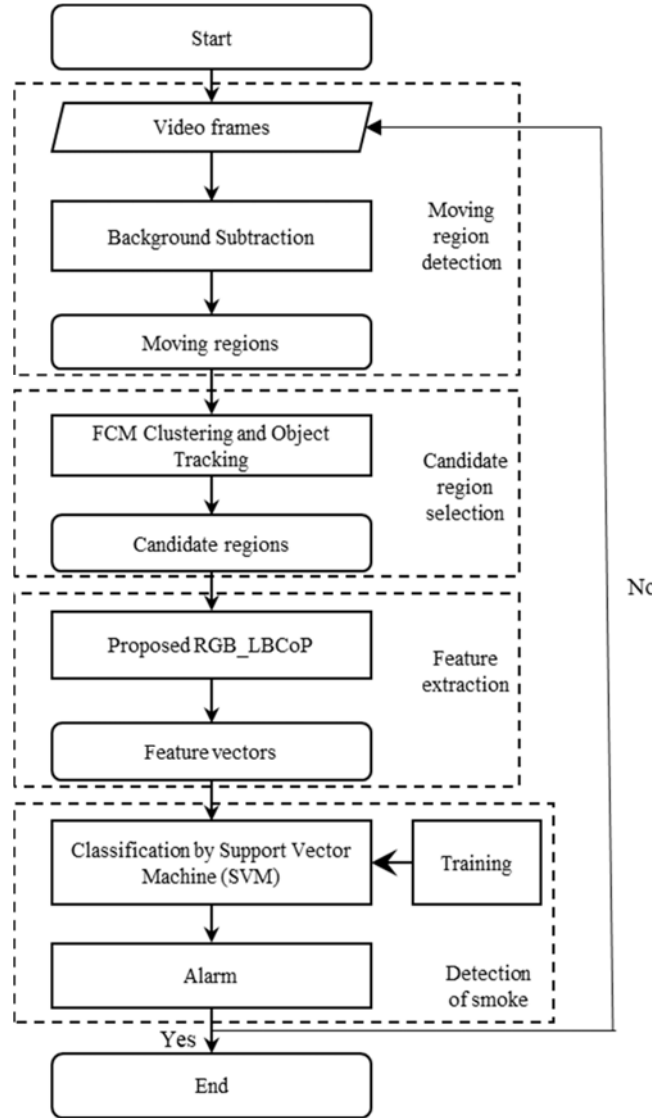


Figure 2: Flowchart of proposed smoke detection

3.2. Candidate Region Detection

After identifying the moving regions in a frame, non-smoke moving objects need to be filtered out in order to decrease the false alarm rate. In this regard, this work utilizes Fuzzy c-means (FCM) algorithm to identify smoke candidate regions from the moving regions of the video. FCM separates the smoke regions from the non-smoke moving regions based on color and intensity. All the pixels of the moving regions are partitioned into some clusters according to the membership value μ_{ik} . The overall computational steps are given below:

- Step 1: Compute the number of clusters c and initialize the centroid value v_i
- Step 2: Calculate the membership value, μ_{ik} , using equation (2)
- Step 3: Update the centroid value using equation (3)
- Step 4: If the termination condition is satisfied stop the iteration otherwise repeat from step 2
- Step 5: Assign all the pixels to the clusters based on the maximum membership value.

It is essential to select the optimal number of clusters and the initial values of the cluster centroid to achieve good clustering accuracy. A cluster is detected as a smoke candidate region when the centroid values are close to the color values of smoke [25]. Otherwise, the cluster is considered as a non-smoke moving region. After detecting the cluster with smoke candidate regions, we have applied the object tracking module on this cluster to estimate the location of the objects in the candidate regions [15, 27].

This object tracking module provides small surrounding rectangles which define the area and position of the candidate objects. To avoid selecting an insignificantly small candidate objects, we have defined a minimum number of pixels for the regions surrounded by green rectangular box. As there is no consensus on this issue in the literature, we have selected 150 pixels (10x15 pixels) as the optimal number based on exhaustive experimentation on our video dataset. Then these rectangular objects are used as the candidate smoke regions for the feature extraction step. Figure 3 illustrates the results of smoke candidate region detection process. The screenshot of the frame with moving region detection is shown in Figure 3 (a). The filtered moving regions are shown in Figure 3 (b). Figure 3 (c) and 3(d) represent the FCM clusters with non-smoke moving regions whereas Figure 3(e) shows the FCM cluster with smoke candidate regions. Figure 3 (f) shows green boxes around the estimated smoke candidate objects.

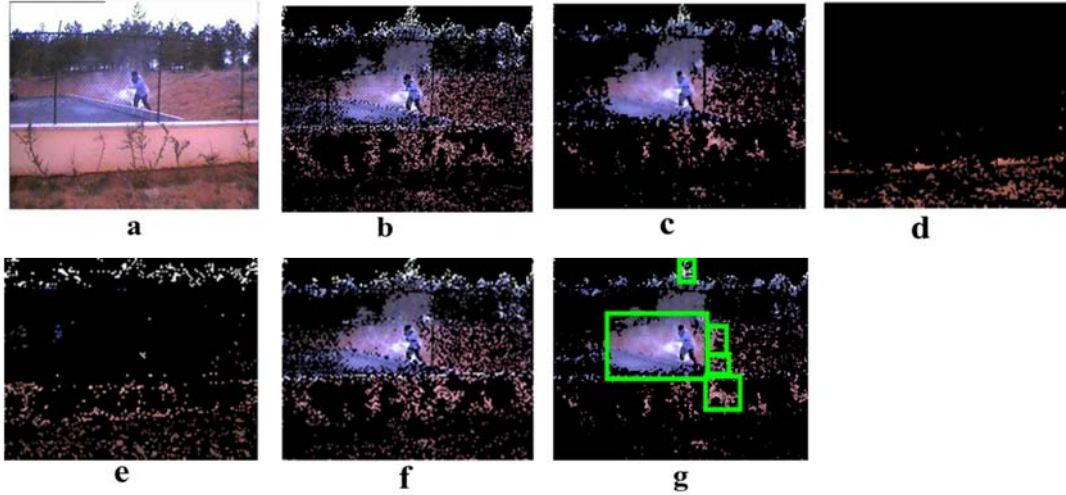


Figure 3: Candidate region selection: (a) Original frame, (b) Moving region detection; (c) Moving regions after median filtering, (d) FCM cluster with non-smoke regions (e) FCM cluster with non-smoke regions, (f) FCM cluster with smoke regions, and (g) Estimated smoke candidate regions using object tracking (area inside the green boxes)

3.3. Feature Extraction

After detecting the smoke candidate regions using FCM and object tracking module, some smoke-like moving objects can be present due to the similarity of their natures with smokes. These smoke-like moving objects can be moving clouds, smoke-like color vehicles, a person wearing grayish black or white clothes, and moving tree leaves. On the other hand, smoke varies rapidly in color, texture and shape and it also blurs the objects which makes the image features more complex [1]. Therefore, it is a challenging task to detect smoke from other smoke-like objects. To handle this challenge, we propose a new feature extraction technique which combines color information along with local and global texture features.

The existing literature confirms the important role of color information in detecting smoke. Furthermore, combination of different texture features (local and global) is expected to bring together complementary information to increase the discriminability leading to higher accuracies in complex scenarios. Motivated by the high accuracy of texture based feature combination, this work proposes a Local Binary Co-occurrence Patterns based on the RGB color space (RGB_LBCoP) by fusing $LBP_{p,R}^{riu2}$ and co-occurrence of the LBP features. The overall procedure of the proposed feature extraction method is presented in Figure 4.

RGB color information along with other motion features were utilized for vision based smoke detection [6, 28]. In contrast, Nguyen et al. analyzed the competency of HSI color space for video based optical smoke detection [15]. Based on our comparative study among the different color spaces, provided in Table 3 (of Section 4), RGB color space exhibits higher accuracy. Because, RGB color space has the ability to handle a wide range of colors and also different combinations of red, green and blue colors help to perceive different colors. Moreover, previous research suggests that, due to the

greyish color of a smoke region, magnitudes of its three components (R, G and B) are almost equal [29, 30]. Therefore, a non-smoke region is likely to be detected by dissimilar magnitude of its three components (R, G and B).

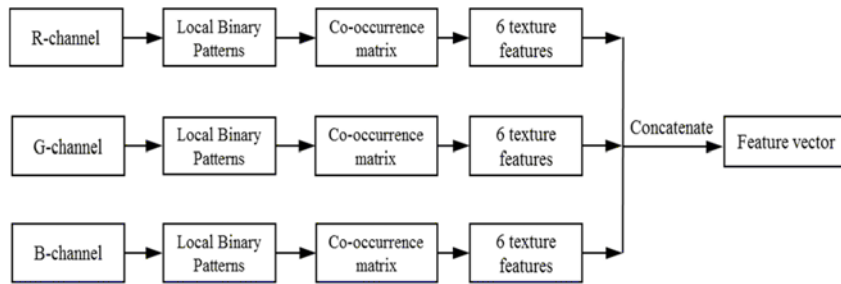


Figure 4: Proposed Local Binary Co-occurrence Patterns in the RGB color space (RGB_LBCoP)

The LBP operator has a number of mapping methods such as uniform, rotation invariant, and rotation invariant uniform. The uniformity is measured by the number of spatial transitions (bitwise 0/1 changes) within the circular neighborhoods [23]. To achieve rotation invariance, the LBP code is circularly rotated into its minimum value [23]. The rotation invariant uniform pattern, $LBP_{P,R}^{riu2}$ ('u2' corresponds to two 0/1 transitions in the pattern), is considered as an excellent image texture descriptor as it has the ability to detect complex microstructures within an image while showing insensitivity to illumination changes. As a result, these properties of the LBP operator can be useful to identify a smoke region, within an image, irrespective of its contrast.

To utilize both the texture and color characteristics of smoke, this study applies LBP operator in all the three channels of the RGB color space. In this respect, the proposed feature extraction method first separates the R, G and B channels and then applies $LBP_{P,R}^{riu2}$ operator in each of the three channels.

The $LBP_{P,R}^{riu2}$ operator can provide detailed information about the local image texture while it is not suitable for measuring the global characteristics of the image. One of the well-known global feature descriptors is Gray Level Co-occurrence Matrix (GLCM), proposed by Haralick et al. [31]. The GLCM computes how often a pair of pixels with the spatial relationship occurs within an image. However, any kind of illumination variation brings great changes in the co-occurrence matrix [18]. To achieve illumination invariance while obtaining global features, spatial co-occurrence of the $LBP_{P,R}^{riu2}$ codes is calculated which measures how often a combination of LBP codes occur in an image. It provides higher order statistical local and global information [32]. This information aids in detecting smoke even in complex environments.

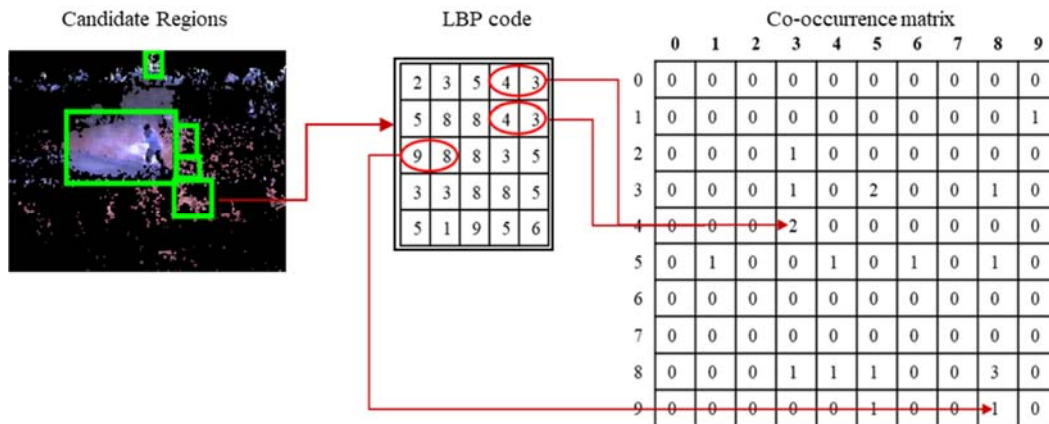


Figure 5: Process to calculate the co-occurrence of the LBP codes considering two neighboring pixels in the horizontal direction

In this work, co-occurrence is measured from the spatial relationship of two horizontally adjacent LBP codes. The process of calculating co-occurrence value is illustrated in Fig.5. From the co-occurrence matrix, nearly up to the 20 second order statistical texture features can be extracted [33]. According to previous research, six features are found to be most effective, among these 20 features, in case of smoke detection [34-36]. Therefore, these six second order texture features, which include

contrast, dissimilarity, homogeneity, difference variance, inverse difference normalized, and inverse difference moment normalized, are chosen to differentiate smoke from other non-smoke regions. The computational steps for feature extraction are given below.

- 1) Extract the R, G, and B channels from the $N \times M$ candidate region
- 2) For each channel, calculate $LBP_{P,R}^{riu^2}$ for each pixel considering $P = 8$ and $R = 1$ which results in $(N - 2) \times (M - 2)$ LBP coded matrix. This (P, R) combination ensures maximum localization within the neighborhood [23]
- 3) Calculate the co-occurrence between two adjacent LBP codes in the horizontal direction within the $(N - 2) \times (M - 2)$ matrix which results in a 10×10 co-occurrence matrix
- 4) Extract six second order texture features (as mentioned above) from the co-occurrence matrix
- 5) Concatenate the features obtained for all the three (R, G and B) channels to create the final feature vector (consists of 18 features)

This feature vector is utilized by the SVM for classification as described in the next section.

3.4. Classification by Support Vector Machine

Support Vector Machine (SVM) is a supervised machine learning algorithm which is extensively used in video-based fire and smoke detection due to its robust performance for two classes (smoke or non-smoke). For non-linear data, a kernel function is used to map the initial data into a high dimensional feature space. Several kernel functions like sigmoid, polynomial, and radial basis function (RBF) were used in different applications. This work utilized RBF kernel as it performs better for video based smoke detection [25]. The RBF kernel function is defined as:

$$k(x, y) = \exp\left(\frac{-\|x - y\|^2}{2\sigma^2}\right) \text{ for } \sigma > 0 \quad (5)$$

Here two input vectors represented by x, y are the input patterns, and σ is set by the user that determines the width of the kernel function. It is important to carefully select the sigma value. With large values of sigma, the function draws an oval around the points without defining any pattern. In contrast, small sigma values can be the reason of overtraining in which the basis function is wrapped tightly around the data points [25]. In this work, the sigma parameter is tuned using pattern search based optimization tool [37]. In the optimization process, an objective function is formulated to account for classification error of 10-fold cross validation. Then the pattern search-based technique is used to minimize the objective function by utilizing the initial value for kernel parameters (randomly chosen from a normal distribution). This optimization process is run for 20 times and, finally, the sigma value is selected which corresponds to the minimum objective function among the 20 trials.

4. Experimental Analysis

The proposed algorithm was implemented in the MATLAB environment with Intel® Core™ i7 processor and an image size of 320×240 . Eight positive smoke video streams and three negative videos were used from the publicly available dataset [2, 38-40]. The video streams have been chosen carefully to include not only smoke but also contain scenes like bright sun, smoke colored background, mountain with fog, clouds etc. Furthermore, to check the validity of the proposed approach, we have selected videos with different colors of smoke including white, ash, whitish ash, blue, black and deep gray. The descriptions of the videos along with smoke color are reported in Table 1 while the screenshots of the frames are shown in Figure 6. The training dataset was obtained by randomly selecting the frames from six videos (Video 1 to Video 6). The training datasets contain 2060 samples in which 1030 are smoke samples and the remaining are non-smoke ones.

To train the SVM, samples from Video 1 to 6 are utilized. Testing is performed using the remaining videos. However, while testing a video from the training group (Video 1 to 6), its samples are excluded from the training database to ensure fair analysis. We experimentally adjusted some threshold values and selected some parameter settings. The fuzzy factor, m and the termination threshold, ϵ were set to 2 and 0.0001 for Fuzzy c-mean based candidate region selection [34]. According to the optimization result, the sigma value is set as 3.75. The SVM is then trained using feature vectors obtained from the training dataset.



Figure 6: Screenshots of the video streams used in this work. Positive videos: Movie 1-13 and Negative Videos: Movie 14-19

Table 1: Descriptions of the videos used for the experiments

Name	Number of Frames	Description
Video 1	630	Smoke on a red ground including other moving objects
Video 2	900	A bluish color smoke generated in a red ground
Video 3	900	Ash coloured smoke near a red dustbin while background exhibits similar color of smoke
Video 4	1725	A thin and blurred whitish ash smoke in the car park. Background also has almost similar color features
Video 5	240	Ash color smoke is visible from a window. The video includes bright reflection of the sun and also other moving objects
Video 6	6025	Small size and very slow moving white smoke on the mountains including moving clouds and fog
Video 7	3168	Slow moving white smoke on the mountains including moving clouds, leaves of the tress
Video 8	7608	White smoke on the mountains including moving clouds, leaves of the tress
Video 9	7500	White smoke on the mountains including moving clouds, leaves of the tress
Video 10	210	Thick black, white and gray coloured smoke in the forest
Video 11	1375	Thick white coloured smoke in the rural-urban interface area
Video 12	750	Thick dark gray and black smoke in the rural-urban interface area
Video 13	150	Thick dark gray and black smoke in the structure
Video 14	840	Smoke coloured moving clouds and reflection of light
Video 15	2784	Moving clouds, fog and tress
Video 16	528	Bright sparkling lights
Video 17	150	Walking person and reflection of light
Video 18	888	Black clouds, fog and mountains
Video 19	7505	Red reflection and cloud in the mountains

4.1. Performance of the Proposed Feature Extraction Method

We have evaluated the performance of the proposed feature extraction method and have compared it with other texture feature descriptors such as Grey Level Co-occurrence Matrix (GLCM), Rotation Invariant Local Binary Patterns (LBP_riu2), Co-occurrence of Local Binary Patterns (CoLBP), Local Binary Patterns in the RGB color space (RGB_LBP). For extracting the local binary pattern-based feature, the $LBP_{P,R}^{riu2}$ operator is utilized. In case of CoLBP, we first extracted the LBP features from the grey scale image and then measured the co-occurrence of the LBP codes.

For RGB_LBP, we have computed the LBP features in the three (R, G, and B) channels and then concatenated the features obtained from the three channels to form a single feature vector. While performing smoke detection utilizing the aforementioned feature extraction methods, we have maintained the same processing stages for moving regions detection, candidate regions selection and classification. This way, we were able to compare performance of the different feature extraction methods in smoke detection. From each positive video, only that segment is chosen for experimental analysis which records from smoke initiation until the occurrence of fire. For negative video, the whole segment is used for calculating false alarm rate.

Table 2: True Positive Rate (TPR) of the Proposed Method and Compared with the other Feature Extraction Methods

	GLCM	LBP_riu2	CoLBP	RGB_LBP	Proposed
Video 1	95	68.31	67	96.87	95.78
Video 2	99.6	100	96.55	100	100
Video 3	65.72	45.46	100	75.5	90
Video 4	60.23	71.69	49.05	87.5	94.44
Video 5	87.54	45.29	80	68.29	96.29
Video 6	93.75	63.64	27.27	55.12	45
Video 7	59.41	45.01	65	89.38	95.71
Video 8	78.82	64.93	89.74	78.7	88.89
Video 9	98.34	96.79	63.16	96.28	97.30
Video 10	80	66.67	58.33	79.89	95.46
Video 11	92.18	90.18	60	99.12	100
Video 12	100	95.43	78.95	100	100
Video 13	85.79	87.4	72.13	91.25	97.45
Average	84.33	72.37	69.78	85.99	92.02

The accuracy is measured by computing the true positive rate (TPR) of the methods [41].

$$TPR = \frac{\text{Number of TP}}{\text{Number of TP} + \text{Number of FN}} \quad (6)$$

Table 2 shows the performance of the proposed feature extraction method along with the other four techniques. The proposed feature descriptor (RGB_LBCoP) – higher accuracy for ten videos out of thirteen videos. This is mainly due to the incorporation of color texture features along with spatial co-occurrence. RGB_LBCoP considers both local and global features of the image along with discriminative color features. In case of Video 2, all techniques achieved higher TPR due to clearly visible smoke which diffuses across the entire frame. In Video 4, smoke is very thin and the color is almost the same as that of background. This complex scenario of smoke was effectively detected by the proposed approach achieving TPR of 94.44% while others detected less than 88%. Hence, sunlight and the dynamic background make Video 5 a complex structure. In spite of this, smoke was successfully detected by our proposed RGB_LBCoP which is depicted in Figure 7. Figure 7(a) represents the original frame of Video 5 whereas 7(b) refers to the candidate regions (area in the green boxes). Figure 7(c) – 7(g) shows the detected smoke regions (area in the red boxes) using GLCM, LBP_riu2, CoLBP, RGB_LBP, and the Proposed one. In Video 6, the slow-moving smoke is of small size and blurred by fog which caused the low TPR of our proposed method in comparison to GLCM. In case of thick black, dark gray and white colored smoke of Video 10 to Video 13, our proposed

method demonstrates almost perfect detection performance. This is due to the fact that the proposed feature descriptor (RGB_LBCoP) combines both color and texture information of all three channels which are of equal nature in case of these smoke colors.

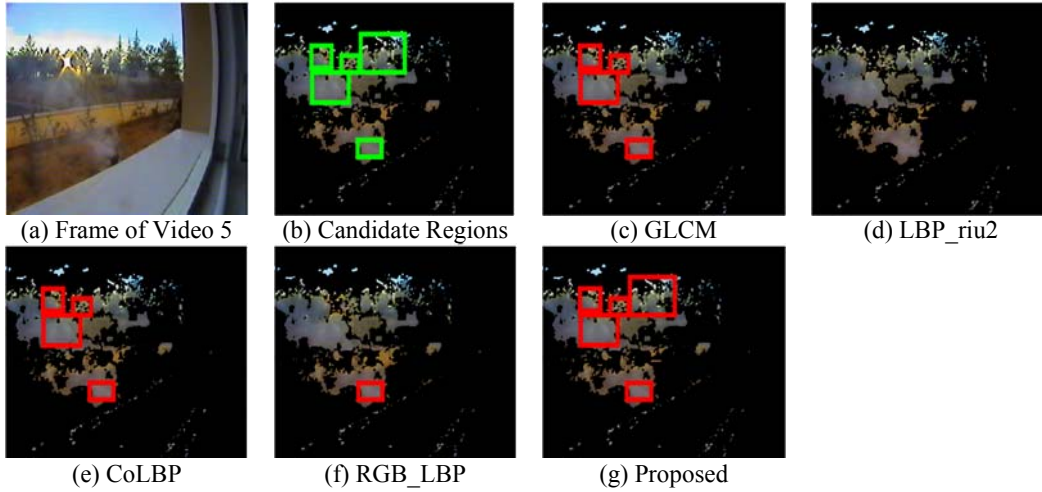


Figure 7: (a) Frame of Video 5, (b) Candidate smoke regions (area in the green boxes), (c) Smoke regions detected by GLCM (area in the red boxes), (d) Smoke regions detected by LBP_riu2 (area in the red boxes), (e) Smoke regions detected by CoLBP (area in the red boxes), (f) Smoke regions detected by RGB_LBP (area in the red boxes) and (g) Smoke regions detected by the proposed RGB_LBCoP (area in the red boxes)

As mentioned earlier, six second order texture features, which include contrast, dissimilarity, homogeneity, difference variance, inverse difference normalized, and inverse difference moment normalized, are extracted from the co-occurrence matrix. Figure 8 (a) presents the distribution of two best features (Feature 7 and 16) out of 18 features for testing the discrimination capacity of the proposed feature descriptors. Feature 7 and 16 represent two second order texture features contrast and difference variance obtained by applying proposed LBCoP on the G and B channel, respectively. For comparison, similar distribution of two best features (Feature 5 and 6) for GLCM is also provided in Figure 8 (b). Here, Feature 5 is the difference variance and Feature 6 is the inverse difference moment normalized. A clearer separation between smoke and non-smoke case is observed for our proposed feature descriptor compared to GLCM.

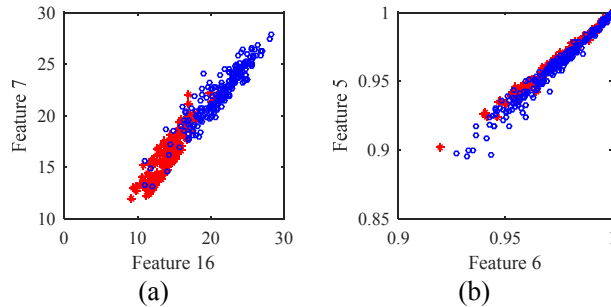


Figure 8: Distribution of two best features (a) Proposed RGB_LBCoP, (b) GLCM. (red star- smoke; blue circles- non-smoke)

Receiver Operating Characteristics (ROC) curves are plotted to demonstrate the performance of the proposed RGB_LBCoP in comparison to the other feature extraction methods. ROC curve represents graphical summaries used to check the quality of the features for an SVM classifier. Figure 9 illustrates the ROC curves of the proposed method which is obtained by plotting TPR against the False Positive Rate (FPR). The FPR can be defined as [41]:

$$FPR = \frac{\text{Number of FP}}{\text{Number of FP} + \text{Number of TN}} \quad (7)$$

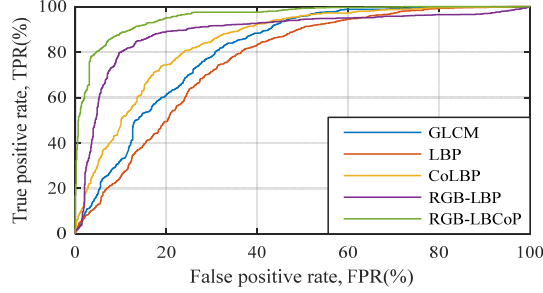


Figure 9: ROC Curves for different features

It is found that the proposed approach generates a higher Area Under the Curve (AUC) of 0.8102 in comparison to the nearest one of RGB_LBP with an AUC of 0.7652. Hence, it is obvious that the proposed RGB_LBCoP performs better than the other feature extraction techniques. In practice, the spatial co-occurrence of $LBP_{P,R}^{riu2}$ provides higher order statistical information that helps to illustrate the subtle difference in a complex image. Thus, the combination of color, local texture and co-occurrence of the proposed approach helps to achieve a higher TPR than the other texture based feature descriptors.

4.2. Comparison of TPR for Different Color Spaces

The purpose of this comparison is to find out the effective color space for the proposed Local Binary Co-occurrence Pattern fused with color information. According to the results in Table 3, the RGB color space achieved better TPR in comparison to the other color spaces. This is because, the RGB visualizes the different color tones of the smoke produced by different combustion materials. In an effort to solidify this claim, we have included different colored smoke in our experimental datasets.

Table 3: Performance Comparison of True Positive Rate (TPR) for Different Color Spaces

	HSI	HSV	YUV	RGB
Video 1	71.43	85.71	91.34	95.78
Video 2	100	100	100	100
Video 3	100	20	28.57	90
Video 4	98	70.59	97.88	94.44
Video 5	95.24	90.38	89.71	96.29
Video 6	41.67	40	42.10	45
Video 7	83.82	92.45	83.47	95.71
Video 8	87.78	66.67	64.78	88.89
Video 9	94.7368	89	93.92	97.30
Video 10	66.6667	50	98.19	95.46
Video 11	90	86.6667	100	100
Video 12	100	96	100	100
Video 13	95.59	92.69	93.84	97.45
Average	86.53	75.39	83.37	92.02

4.3. Comparison with VisiFire Software

This section provides performance comparison between the proposed method and the popular VisiFire software [42]. VisiFire is a windows-based software available for limited academic use. It was developed combining different academic research contributions [3, 39, 43-45]. It has various analyzers for fire and smoke detection such as smoke detector, forest smoke detector, fire and smoke detector, fire detector, and an infrared sensor-based fire detector. We have selected the smoke detector with the default software settings for performance analysis. Using this software, it is easy to detect the first frame in which smoke appears, as the frame number is displayed on the screen. However, any provision for calculating TPR is not available. Due to this reason, we determine the first detection-frame for both VisiFire and the proposed algorithm for the experimental dataset and provide the results in Table 4. Except for Video 6, the proposed method detects smoke earlier than the VisiFire software. It demonstrates the superior detection performance of the proposed method in comparison to VisiFire.

Table 4: Experimental Results for Comparing Proposed Method with VisiFire Software

	First Smoke Frame	VisiFire	Proposed Method
Video 1	36	227	110
Video 2	1	70	3
Video 3	11	110	19
Video 4	280	540	491
Video 5	34	Not detected	61
Video 6	914	1014	1019
Video 7	1	143	71
Video 8	1	91	48
Video 9	1	35	17
Video 10	1	116	38
Video 11	1	35	5
Video 12	1	30	7
Video 13	1	42	25

4.4. Evaluation of the Smoke Detection Algorithm

Performance of the proposed video based smoke detection algorithm is compared with two other state-of-the-art algorithms. One of them detects smoke based on motion features and SVM (Motion_SVM in short) [25] while the other is based on GLCM along with FCM and neural network (GLCM_FCM in short) [15]. In order to evaluate the performance, TPR for the positive videos and false positive rate FPR for negative videos are used.

According to the results listed in Table 5, our proposed RGB_LBCoP achieves a higher average TPR than the other algorithms. The low TPR of Motion_SVM algorithm is due to its assumption that smoke diffuses only upwards. However, in reality, smoke also spreads in other directions. The GLCM_FCM performs better compared to Motion_SVM because it considers texture features along with color based candidate region selection. However, the GLCM_FCM achieves a lower TPR than the proposed algorithm as it extracts features from grey scale image based on GLCM which is highly affected by illumination variations. It should be mentioned here that, the accuracies reported in the original work of Motion_SVM and GLCM_FCM are different from those obtained in our implementation. This is likely due to two reasons: a) We have implemented the algorithms and tested for our experimental videos, b) the calculation method of accuracy used in those works was different from equation (6) and c) the video dataset of those works was limited (less than 6 outdoor videos) and slightly different from the present work.

To further evaluate the effectiveness of the proposed method, six non-smoke videos that include different conditions were also tested. The False Alarm Rate of the negative videos are reported in Table 6. For Video 14, all three algorithms were suffering from false alarms because the video contains smoke colored moving clouds while in between the clouds reflections of sunlight looked similar to smoke. In Video 16, bright colored live billboard was selected as a smoke by Motion_SVM and GLCM_FCM. However, the combination of local and global features by the proposed approach help to detect the complex scenes which leads to achieve zero FPR for Video 16.

Table 5: Performance Comparison of the Proposed Method and other methods using True Positive Rate (TPR)

	GLCM_FCM[15]	Motion_SVM[25]	Proposed
Video 1	94.88	88.6	95.78
Video 2	98.18	91.4	100
Video 3	65.72	74.47	90
Video 4	58.33	81.31	94.44
Video 5	89.94	88.32	96.29
Video 6	93.75	50.12	45
Video 7	59.41	65.74	95.71
Video 8	78.82	71.65	88.89
Video 9	96.36	88.91	97.30
Video 10	93.55	91.38	95.46
Video 11	98.79	93.66	100
Video 12	99	97	100
Video 13	95.38	94.84	97.45
Average	86.32	82.88	92.02

Table 6: Performance Comparison of the Proposed Method and other methods using False Alarm Rate

	GLCM_FCM [15]	Motion_SVM [25]	Proposed
Video 14	25.85	18.9	17.73
Video 15	0	0	3.5
Video 16	3.74	4.8	0
Video 17	0	0	0
Video 18	0.5	2.6	0
Video 19	0	0	0
Average	5.02	4.38	3.53

4.5. Computational Complexity

The proposed video-based smoke detection consists of multiple stages: moving region detection, candidate region selection, feature extraction and classification. The computational complexity is calculated as an average time required to process single frame which is listed in Figure 10. It was calculated using MATLAB (R2015a) environment with Intel® Core™ i7 processor. GLCM_FCM is computationally expensive as it uses neural networks for detecting smoke. The computational complexity of the proposed algorithm can be further improved by using Digital Signal Processors (DSP), and Graphical Processing Units (GPU).

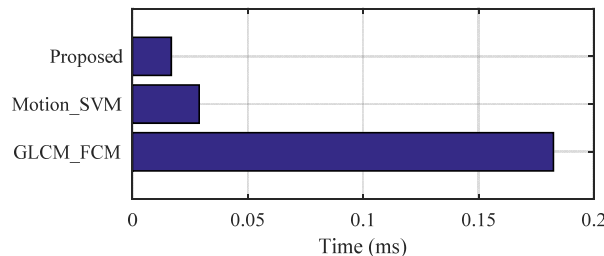


Figure 10: Computational Complexity of the proposed RGB_LBCoP in comparison to GLCM_FCM [15] and Motion_SVM [25]

Both qualitative and quantitative experimental analysis depicts that the proposed technique is a promising approach that can even be further improved to make it suitable for real time application.

5. Conclusion

Local and global texture properties, obtained by the proposed Local Binary Co-occurrence Patterns for the RGB color space (RGB_LBCoP), are exploited efficiently in this work for video based smoke detection. To explore the texture features of smoke, we utilized an efficient co-occurrence encoding scheme along with rotation invariant uniform LBP and extended it to integrate RGB color information. As a result, the proposed approach can not only capture the local and global information, but also measure higher order statistical texture information that helps to describe the complex structure of smoke in video streams. We have investigated the capabilities of our proposed feature extraction method, RGB_LBCoP, and have compared its performance with other conventional texture feature extraction methods such as GLCM and LBP derivatives. In addition, the proposed smoke detection method has been tested on various videos with different environmental conditions and its performance compared with another state-of-the-art methods. According to the comparison results, the proposed method is found to be effective in accurately detecting smoke while demonstrating well-balanced trade-off between the accuracy, false alarm and computational complexity. In future, the proposed method will be extended to incorporate fire detection capability.

6. References

- [1] Yuan F, Shi J, Xia X, Fang Y, Fang Z, Mei T, High-order local ternary patterns with locality preserving projection for smoke detection and image classification, *Information Sciences*, 2016;372: 225-40.
- [2] A. Enis Çetina, KosmasDimitropoulosb, Benedict Gouverneurc, Nikos Grammalidisb, Osman Günaya, Y. Hakan Habibo^glua, B. U^gur Töreayind, Verstockte S, Video Fire Detection - Review, *Digital Signal Processing*, 2013;23: 1827-43.
- [3] Töreayin BU, Dedeoğlu Y, Çetin AE. Wavelet based real-time smoke detection in video. In 2005 13th European Signal Processing Conference, 2005, pp. 1-4.
- [4] Gubbi J, Marusic S, Palaniswami M, Smoke detection in video using wavelets and support vector machines, *Fire Safety Journal*, 2009;44: 1110-15.
- [5] Besbes O, Benazza-Benyahia A, A novel video-based smoke detection method based on color invariants, 2016: 1911-15.
- [6] Chunyu Y, Jun F, Jinjun W, Yongming Z, Video Fire Smoke Detection Using Motion and Color Features, *Fire Technology*, 2010;46: 651-63.
- [7] Qureshi WS, Ekpanyapong M, Dailey MN, Rinsurongkawong S, Malenichev A, O K, QuickBlaze: early fire detection using a combined video processing approach, *Fire Technology*, 2016;52: 1293-317.
- [8] Yang Jia, Jie Yuan, Jinjun Wang, Jun Fang, Zhang Y, Zhang Q, A Saliency-Based Method for Early Smoke Detection in Video Sequences, *Fire Technology*, 2016;52: 1271-92.
- [9] Wonjae L, Seonghyun K, Yong-Tae L, Hyun-Woo L, Min C. Deep neural networks for wild fire detection with unmanned aerial vehicle. In 2017 IEEE International Conference on Consumer Electronics (ICCE), 2017, pp. 252-53.
- [10] Tao C, Zhang J, Wang P. Smoke Detection Based on Deep Convolutional Neural Networks. In 2016 International Conference on Industrial Informatics - Computing Technology, Intelligent Technology, Industrial Information Integration (ICIIII), 2016, pp. 150-53.
- [11] Xu G, Zhang Y, Zhang Q, Lin G, Wang J, Deep Domain Adaptation Based Video Smoke Detection using Synthetic Smoke Images, arXiv preprint arXiv:1703.10729, 2017.
- [12] Emmy Prema C, Vinsley SS, Suresh S, Efficient Flame Detection Based on Static and Dynamic Texture Analysis in Forest Fire Detection, *Fire Technology*, 2018;54: 255-88.
- [13] Wang Y, Wu A, Zhang J, Zhao M, Li W, Dong N. Fire smoke detection based on texture features and optical flow vector of contour. In 2016 12th World Congress on Intelligent Control and Automation (WCICA), 2016, pp. 2879-83.
- [14] Prema CE, Vinsley S, Suresh S, Multi feature analysis of smoke in YUV color space for early forest fire detection, *Fire Technology*, 2016;52: 1319-42.
- [15] Nguyen TKT, Kim J-M, Multistage optical smoke detection approach for smoke alarm systems, *Optical Engineering*, 2013;52(5).
- [16] Yuan F, Video-based smoke detection with histogram sequence of LBP and LBPV pyramids, *Fire Safety Journal*, 2011;46: 132-39.
- [17] Zhao Y, Zhou Z, Xu M, Forest fire smoke video detection using spatiotemporal and dynamic texture features, *Journal of Electrical and Computer Engineering*, 2015;2015: 40.
- [18] Qi X, Xiao R, Li C-G, Qiao Y, Guo J, Tang X, Pairwise rotation invariant co-occurrence local binary pattern, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014;36: 2199-213.
- [19] Nosaka R, Ohkawa Y, Fukui K. Feature extraction based on co-occurrence of adjacent local binary patterns. In Pacific-Rim Symposium on Image and Video Technology, 2011, pp. 82-91.
- [20] Murala S, Wu QJ, Local ternary co-occurrence patterns: a new feature descriptor for MRI and CT image retrieval, *Neurocomputing*, 2013;119: 399-412.
- [21] Rassem TH, Mohammed MF, Khoo BE, Makbol NM. Performance evaluation of Completed Local Ternary Patterns (CLTP) for medical, scene and event image categorisation. In 2015 4th International Conference on Software Engineering and Computer Systems (ICSECS), 2015, pp. 33-38.
- [22] Bezdek JC, Keller J, Krisnapuram R, Pal N, Fuzzy models and algorithms for pattern recognition and image processing, Springer Science & Business Media, 2006.

- [23] Ojala T, Pietikainen M, Maenpaa T, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002;24: 971-87.
- [24] Vapnik VN, An overview of statistical learning theory, *IEEE transactions on neural networks*, 1999;10: 988-99.
- [25] Truong Xuan Tung, Kim J-M, An effective four-stage smoke-detection algorithm using video images for early fire-alarm systems, *Fire Safety Journal*, 2011;46: 276-82.
- [26] Shahbaz A, Hariyono J, Jo KH. Evaluation of background subtraction algorithms for video surveillance. In 2015 21st Korea-Japan Joint workshop on Frontiers of Computer Vision (FCV), 2015, pp. 1-4.
- [27] Maruta H, Nakamura A, Kurokawa F. A new approach for smoke detection with texture analysis and support vector machine. In 2010 IEEE International Symposium on Industrial Electronics, 2010, pp. 1550-55.
- [28] Calderara S, Piccinini P, Cucchiara R, Vision based smoke detection system using image energy and color information, *Machine Vision and Applications*, 2011;22: 705-19.
- [29] Thou-Ho C, Ping-Hsueh W, Yung-Chuen C. An early fire-detection method based on image processing. In 2004, International Conference on Image Processing (ICIP), 2004, pp. 1707-10 Vol. 3.
- [30] Wang Y, Chua TW, Chang R, Pham NT. Real-time smoke detection using texture and color features. In Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012), 2012, pp. 1727-30.
- [31] Haralick RM, Shanmugam K, Textural features for image classification, *IEEE Transactions on systems, man, and cybernetics*, 1973;3: 610-21.
- [32] Islam IU. Feature Fusion for Pattern Recognition. Politecnico di Torino, 2015.
- [33] Albrechtsen F, Statistical texture measures computed from gray level cooccurrence matrices, Image processing laboratory, department of informatics, university of oslo, 2008;5.
- [34] Alamgir N, Boles W, Chandran V. A Model Integrating Fire Prediction and Detection for Rural-Urban Interface. In 2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA) Adelaide, Australia, 2015, pp. 1-8.
- [35] Srisuwan T, Ruchanurucks M. Smoke detection using GLCM, wavelet, and motion. In Fifth International Conference on Graphic and Image Processing, 2014, pp. 90691H-91H-5.
- [36] John Adedapo Ojo, Oladosu JA, Video-based Smoke Detection Algorithms: A Chronological Survey, *Computer Engineering and Intelligent Systems*, 2014;5: 38-50.
- [37] MathWorks. Support Vector Machines (SVM). 2015. Available: <https://au.mathworks.com/help/releases/R2015a/stats/support-vector-machines-svm.html?searchHighlight=Tuning%20an%20SVM%20Classifier>
- [38] Byoung Chul Ko, Kwang-Ho Cheong, Nam J-Y, Fire detection based on vision sensor and support vector machines, *Fire Safety Journal*, 2009;44: 322-29.
- [39] Töreyn BU, Dedeoğlu Y, Güdükbay U, Çetin AE, Computer vision based method for real-time fire and flame detection, *Pattern Recognition Letters*, 2006;27: 49-58.
- [40] Foggia P, Saggese A, Vento M, Real-Time Fire Detection for Video-Surveillance Applications Using a Combination of Experts Based on Color, Shape, and Motion, *Circuits and Systems for Video Technology*, *IEEE Transactions on*, 2015;25: 1545-56.
- [41] Zhao Y, Tang G, Xu M, Hierarchical detection of wildfire flame video from pixel level to semantic level, *Expert Systems with Applications*, 2015;42: 4097-104.
- [42] Cetin AE. Computer Vision based fire detection software-sample video clips. Available: <http://signal.ee.bilkent.edu.tr/VisiFire/>
- [43] Toreyin BU, Dedeoglu Y, Cetin AE. Contour based smoke detection in video using wavelets. In 2006 14th European Signal Processing Conference, 2006, pp. 1-5.
- [44] Toreyin BU, Dedeoglu Y, Cetin AE. Flame detection in video using hidden Markov models. In 2005 IEEE International Conference on Image Processing, 2005, pp. II-1230-3.
- [45] Töreyn BU, Cinbis RG, Dedeoglu Y, Cetin AE. Fire detection in infrared video using wavelet analysis. *Optical Engineering*, vol. 46, p. 9, 2007.