

Article

Early Fire Detection on Video Using LBP and Spread Ascending of Smoke

Jesus Olivares-Mercado ¹, Karina Toscano-Medina ¹, Gabriel Sánchez-Perez ¹, Aldo Hernandez-Suarez ¹, Hector Perez-Meana ¹ and Ana Lucila Sandoval Orozco ² and Luis Javier García Villalba ^{2,*}

¹ Instituto Politecnico Nacional, ESIME Culhuacan, Mexico City 04440, Mexico; jolivares@ipn.mx (J.O.-M.); ltoscano@ipn.mx (K.T.-M.); gasanchezp@ipn.mx (G.S.-P.); ahernandezs1325@alumno.ipn.mx (A.H.-S.); hmperezm@ipn.mx (H.P.-M.)

² Group of Analysis, Security and Systems (GASS), Department of Software Engineering and Artificial Intelligence (DISIA), Faculty of Computer Science and Engineering, Office 431, Universidad Complutense de Madrid (UCM), Calle Profesor José García Santesmases, 9, Ciudad Universitaria, 28040 Madrid, Spain; asandoval@fdi.ucm.es

* Correspondence: javiergv@fdi.ucm.es

Received: 19 April 2019; Accepted: 29 May 2019; Published: 13 June 2019



Abstract: This paper proposes a methodology for early fire detection based on visual smoke characteristics such as movement, color, gray tones and dynamic texture, i.e., diverse but representative and discriminant characteristics, as well as its ascending expansion, which is sequentially processed to find the candidate smoke regions. Thus, once a region with movement is detected, the pixels inside it that are smoke color are estimated to obtain a more detailed description of the smoke candidate region. Next, to increase the system efficiency and reduce false alarms, each region is characterized using the local binary pattern, which analyzes its texture and classifies it by means of a multi-layer perceptron. Finally, the ascending expansion of the candidate region is analyzed and those smoke regions that maintain or increase their ascending growth over a time span are considered as a smoke regions, and an alarm is triggered. Evaluations were performed using two different classifiers, namely multi-Layer perceptron and the support vector machine, with a standard database smoke video. Evaluation results show that the proposed system provides fire detection accuracy of between 97.85% and 99.83%.

Keywords: smoke detection; Multi-Layer Perceptron; Artificial Neural Network; Local Binary Pattern; Support Vector Machines

1. Introduction

Earlier, faster and accurate fire detection is very important to safeguard human lives, wildlife, forests, etc., because it allows a timely response by emergency services. Smoke, which is the first symptom in flameless combustion in most fires, has various important characteristics that can be used for early fire detection: low temperature, movement, gray tones, dynamic texture, and ascending expansion. Thus, intelligent and automated video surveillance systems, which have been a topic of active research and development during the last years, exploit those qualities for early fire detection. These systems have important advantages over traditional systems, such as quick detection, not requiring smoke contact, can operate in open spaces or adverse air currents, remote sensing capabilities, and take advantage of the existing infrastructure in many buildings and public areas. These systems are based on visual fire detection techniques that use standard video surveillance

cameras, together with sophisticated computational analysis algorithms that segment the images to describe and recognize the visual fire characteristics to determine their presence.

Numerous computer vision algorithms for fire detection have been proposed, which can be classified into two broad categories: flame-based and smoke-based video detection algorithms. In both cases, the primarily extracted features are the movement and color together with a mixture of additional parameters such as texture, expansion, silhouette, blinking, direction, transparency, changes in geometry, and self-similarity, among others. In most fires, the smoke appears before the flames, because the expanding movement increases the probability of falling into the visual field of the camera. Thus, this work focuses on the detection of smoke to activate the fire alarm.

Several algorithms have been reported in the literature regarding early fire detection. T. H. Cheng et al. [1] proposed an early fire warning, based on video processing using two decision rules: a static rule-based decision and a dynamic decision rule. Here, the static decision rule follows the chromaticity of gray-colored smoke, which states that, in the Red–Green–Blue (RGB) color space, the three components must have very similar values, while the dynamic decision rules depend on the attributes of smoke spread, where silhouette disorder is analyzed against one circumference. It can be considered as a disadvantage since there are several objects whose relationship with the circumference may be similar to smoke. Yuan [2] proposed a model of cumulative movement based on the integral image by estimating the orientation based of fast-moving smoke, however, this estimation is not accurate enough because it is based on the sum of the blocks. To solve this limitation, an accumulation of orientation in time is performed, which offsets the results of the inaccuracy of the orientation, considering also the movement, color and direction. Then, to determine the movement, the differentiation of successive images is used. In the color components, it applies the rules proposed in [1] and adds an extra rule that expands to bluish shades detection. A drawback of this scheme is that it considers only the analysis of the upward movement of the smoke, which is not always the case. Töreyn et al. [3,4] proposed a fire detection scheme using the characteristics of semi-transparency and the changing shape of smoke over time. It is based on the fact that the edges of an object covered by smoke lose their sharpness, resulting in a decrease of high frequencies. The authors analyzed the images with wavelet transform and monitored the slow and gradual reduction of wavelet components using hidden Markov models with a temporary wavelet analysis. This can be considered as a flaw because the processing time is considerable, which complicates its implementation in real time [3,4]. Yu et al. [5] analyzed only the texture and proposed a combined use of wavelets and Gray Level Co-occurrence Matrix (GLCM) with four parameters. Specifically, they considered entropy, contrast, the second angular momentum, difference inverse of the moment and the correlation among them in four different directions together with an Artificial Neural Network (ANN) classifier. A weakness of this scheme is that it only examines the texture, which may produce false positives and negatives due to the complexity of the description of the dynamics of the smoke texture. Maruta et al. [6] used GLCM with fourteen statistical measures describing smoke texture. To determine the movement with a difference of 1-s delay between images, they described the texture region with movement and used support vector machines to establish whether the texture corresponds to smoke. If the texture corresponds to smoke, the resulting regions are accumulated and, if a given threshold is exceeded, an alarm is activated. Millan et al. [7] detected the presence of smoke using video sequences captured by IP cameras, in which some important characteristics of smoke, such as the properties of color, movement and growth, are used. The Discrete Cosine Transform (DCT) is used to increase the detection accuracy without requiring the inverse DCT operation. This system provides detection results with a false negative error rate of 4% and a false positive rate of 2%. Aponte et al. [8] introduced a hardware solution using different sensors such as CO, smoke and temperature sensors. Xuehui et al. [9] used a camera and different characteristics such as texture, wavelet, and color edge orientation and proposed an AdaBoost-like classifier.

The aim of the algorithms mentioned above is to avoid confusing the properties of smoke with other objects in the scene that contain some similar features. However, different situations complicate

smoke detection causing false alarms, which still must be resolved. To solve some of these problems, this paper proposes a smoke detection scheme using some particular properties and characteristics of smoke. The proposed system employs a mix of smoke features such as motion, color and texture, which are processed in cascade, to detect the presence of smoke in the scene. For motion detection, the proposed algorithm combines the accumulation of movement by differentiating successive frames subtracting the adaptable background. The color detection is determined by using two rules that define shades of gray smoke. The texture characterization is performed by Local Binary Pattern (LBP) and Local Binary Pattern Variance (LBPV) [10] using a multiscale analysis. Next, the regions that meet these three characteristics are considered as candidate regions with smoke, which are ultimately tested to determine if the expansion through the temporal blocks overlap. The system was evaluated using standard videos, activating the alarm in an average of 6 s in videos containing smoke. The proposed system provides a detection rate of about 98%.

The rest of the paper is organized as follows. Section 2 describes the overall scheme consisting of motion detection, color detection, texture classification analysis, growth analysis, direction analysis and alarm activation. In Section 3, the results are presented and discussed. Finally, the paper concludes in Section 4.

2. Proposed System

The block diagram of the proposed system is shown in Figure 1. Here, the system firstly captures the video signal, which is segmented into frames. The received frames are then inserted into the motion detection step, which combines the movement accumulation by estimating the difference between successive images together with the adaptive subtraction background. Then, the same frame is inserted into the color detection stage, which is determined using two rules defining the gray tones of the smoke. Then, it is inserted into the texture detection stage, to characterize the texture using the LBP and LBPV based on multiscale analysis. Finally, the regions that meet these three characteristics are considered as candidate regions, in which the presence of smoke might be found. These are finally analyzed to determine whether the expansion is upward, by temporarily overlapping blocks. The next sections describe in depth each block of the proposed scheme.

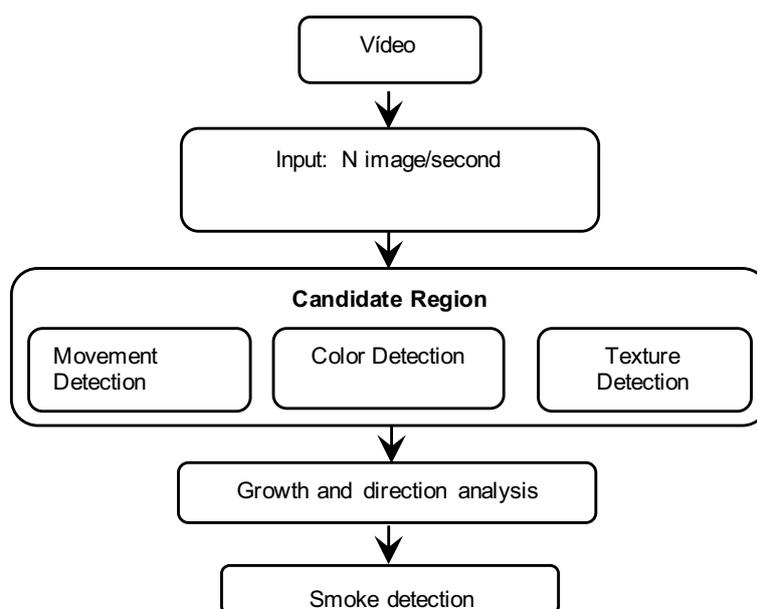


Figure 1. Proposed system.

2.1. Motion Detection

The smoke, although sometimes not visible, is always moving because it is warmer than the surrounding air and presents an ascending and expansive movement as well as agglomerations. This movement is very special and can be subjected to extensive analysis to evaluate the influence of several factors that affect the precise detection of the movement: partial occlusions, repetitive motion and low contrast. For instance, at the limits of the smoke cloud, it tends to be diluted with an increased transparency, camouflaging the background, while, within the cloud, under some lighting conditions and in certain regions, the smoke has a uniform intensity.

A hybrid algorithm for smoke motion detection is proposed, which combines a buildup of the absolute differences between consecutive frames with an adaptive background subtraction. In this way, it is possible to compensate the limitations of each technique due to the fact that the background subtraction is very sensitive to dynamic changes of lighting conditions. Besides that, it is also very sensitive to subtle movements and effective for detecting smoke in the presence of light, where it is almost transparent. Because the mean accumulation of differences between frames ($D(x, y)$) and background extraction ($D_F(x, y)$) are robust to changes in lighting and highly adaptable to dynamic changes, the movement detection is done as follows:

$$D(x, y) = \frac{\sum_{t=1}^{\tau} |I_t(x, y) - I_{t-1}(x, y)|}{\tau - 1} \quad (1)$$

$$D_F(x, y) = |B_{\tau}(x, y) - I_t(x, y)| \quad (2)$$

$$B_{\tau}(x, y) = \begin{cases} \alpha B_{\tau}(x, y) + (1 - \alpha) I_{\tau}(x, y) & \text{stationary} \\ B_{\tau}(x, y) & \text{movement} \end{cases} \quad (3)$$

where $I_t(x, y)$ is the intensity value of the pixel at position (x, y) of the t th frame of the image sequence ($1 < t \leq \tau$), τ is the number of frames to be considered in the sequence, B_{τ} is the image adaptable background, and α is a time constant that determines how fast the new information replaces the previous observations.

$$U = \begin{cases} \mu(D) + \sigma(D) & \text{if } U^{\min} \leq \mu(D) + \sigma(D) \leq U^{\max} \\ U^{\min} & \text{if } \mu(D) + \sigma(D) < U^{\min} \\ U^{\max} & \text{if } \mu(D) + \sigma(D) > U^{\max} \end{cases} \quad (4)$$

$$U_F = \begin{cases} \mu(D_F) + \sigma(D_F) & \text{if } U_F^{\min} \leq \mu(D_F) + \sigma(D_F) \leq U_F^{\max} \\ U_F^{\min} & \text{if } \mu(D_F) + \sigma(D_F) < U_F^{\min} \\ U_F^{\max} & \text{if } \mu(D_F) + \sigma(D_F) > U_F^{\max} \end{cases} \quad (5)$$

$$D'(x, y) = \begin{cases} 1 & \text{if } D(x, y) > U \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

$$D'_F(x, y) = \begin{cases} 1 & \text{if } D_F(x, y) > U_F \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

$$M(x, y) = \begin{cases} 1 & \text{if } D'(x, y) \vee D'_F(x, y) \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

where μ is the mean and σ is the standard deviation of D and D_F . U and U_F are the threshold value for D and D_F , respectively, which are within the interval defined by $[U^{\min}, U^{\max}]$, $[U_F^{\min}, U_F^{\max}]$. The thresholds U^{\min} , U^{\max} , U_F^{\min} and U_F^{\max} were determined experimentally using cross-validation. Operator \vee corresponds to the logical operator "OR". The result is an $M(x, y)$ binary image where the pixels with movement correspond to logical 1. Several sequences of images, taken in different scenarios, were analyzed, since lighting factors showed marked variations in each scenario and

complex situations. These factors are also variations with common noises present in motion detection, in particular reflections and shadows. The general thresholds obtained experimentally for different lighting conditions are shown in Table 1.

Table 1. Thresholds obtained for movement detection.

Movement Threshold	Minimum	Maximum
U	3.77	10.11
U_F	7.03	21.18

This method recognizes general movement patterns, even in low illumination and low contrast, as shown in Figure 2, namely, semitransparent smoke regions, as well as detects isolated pixels that are considered as noise since they are produced by the shadows, video compression, or camera quantification. It also has a low computational cost. However, several factors do not allow a complete removal of residual noise without using a post-processing stage. To this end, we use morphological and median filtering as the post-processing method. In this research, the morphological erosion operation was used. The mathematical formulation applied to the $M(x, y)$ image for erosion is given as:

$$M \ominus E = \{(x', y') | (x' + i, y' + j) \in P_M(i, j), \forall (i, j) \in P_E\} \quad (9)$$

where moving image $M(x, y)$ is to be filtered; E is the reference structure or structural element, where the structural element of the diamond was used for this research; \ominus is the erosion symbol; and P_M and P_E are the pixels sets with value 1 of $M(x, y)$ and E , respectively.

The median filter, with a window of 7×7 , was used to remove the remaining isolated pixels that are considered as noise. In addition, the criterion for selecting candidate regions of movement for further analysis is that they must contain at least the same number of pixels as the block used in the training of classifiers, i.e., 22×22 pixels. These dimensions were chosen because they have a good statistical consistency and are less prone to noise from other textures captured on a larger block. The approved regions are obtained using a bounding box algorithm, which contains a binary region in the smallest possible rectangle avoiding the noise related to the shape of smoke, such as holes and omitted parts. Figure 2 shows the results of the detected motion regions, marked with white boxes in some of the videos used, which have different variations in lighting and contrasts, translucent smoke and smoke camouflaged by the background.

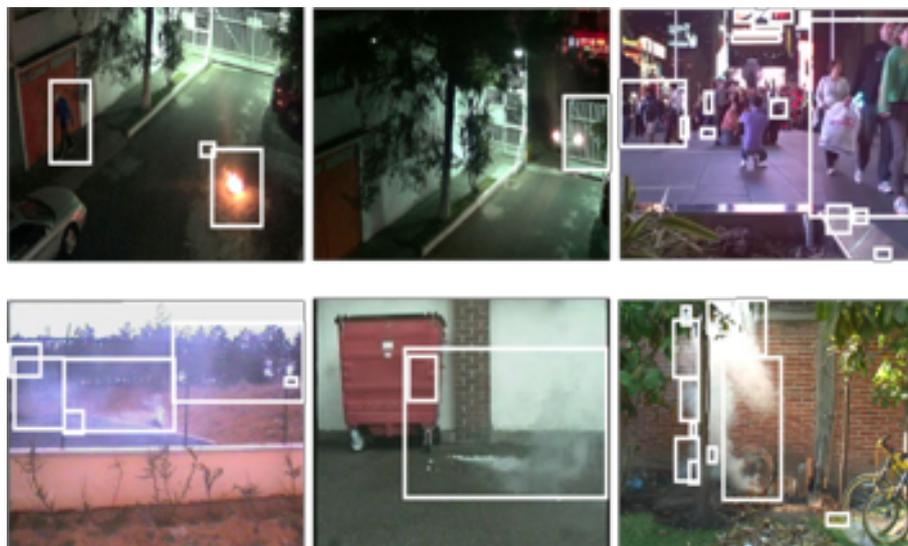


Figure 2. Result of motion detection, where white boxes indicate regions with motion.

2.2. Color Detection

Color is an important feature of most objects in the real world, thus provides important additional information about an object under analysis. The color of smoke depends on several factors: the temperature of combustion, the combustible material, and the lighting, among others. However, in uncontrolled combustion, smoke usually presents colors ranging from light colors due to micro-droplets produced during combustion without flame and pyrolysis of fuel, to black due to the carbonaceous solid particles or soot produced during combustion with flames [11]. In [1], it is established that the smoke varies from white to bluish white smoke when the temperature is low and black-gray to black when combustion presents flames. This observation is the guideline for determining a suitable threshold that restricts the very dark tones, because the main objective of this research is to detect fires in the early stages of combustion. The characteristic color of smoke together with the movement serves to segment more specific regions. In the RGB color space, gray colors are located on the diagonal between black and white, thus the values of the three RGB components should be similar to each other. To detect the characteristic color of smoke, the following rules are defined:

1. The absolute difference of the maximum and minimum values among the three components (RGB) must be less than a predetermined threshold U_1^c .
2. The average of the three components must be greater than the threshold U_2^c .

The first rule allows us to detect gray colors defined in Equation (10). The second rule defines the range of colors, especially very dark colors, defined by Equation (11).

$$C^R(x, y) = \begin{cases} 1 & \text{if } |\max(R(x, y), G(x, y), B(x, y)) - \min(R(x, y), G(x, y), B(x, y))| < U_1^c \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

$$C^P(x, y) = \begin{cases} 1 & \text{if } \frac{R(x, y) + G(x, y) + B(x, y)}{3} > U_2^c \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

where C^R and C^P are the conditions that must be satisfied for gray and dark colors, respectively; U_1^c and U_2^c are thresholds used for determining gray and dark colors, respectively. and $R(x, y)$, $G(x, y)$ and $B(x, y)$ are the values of the color components at the previously segmented position (x, y) of the region with movement.

The optimal thresholds used for color detection were found to be $U_1^c = 30.70$ and $U_2^c = 89.04$, respectively, where the first one was obtained by the ratio between the mean of intensity and the first standard deviation, while the second one was calculated subtracting the first standard deviation from the mean intensity. Thus, a pixel is considered to have a smoke-like color if Equations (10) and (11) are satisfied simultaneously:

$$C^H(x, y) = \begin{cases} 1 & \text{if } C^R(x, y) \wedge C^P(x, y) \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

Color detection is performed only in the regions in which some movement has been detected in the previous stage. Figure 3 shows the evaluation results in which the proposed system is required to detect the color of smoke. Evaluation results show that, in most cases, an accurate detection is achieved, although, in some instances, when the smoke regions are highly transparent, such that the background color is more noticeable than the color of smoke and therefore the region under analysis does not meet the established rules, the system fails to accurately detect the color of smoke.



Figure 3. Smoke detection results based on the characteristic color of smoke.

2.3. Texture Analysis and Classification

Among the most important attributes of an object is its texture, which is related to the visual properties of the observed object. These properties are derived from the surface roughness, the structure or composition of the material from which it is made. The visual texture is defined as the characteristic variation in the intensity that allows describing and recognizing a particular region of a given image. That is, texture does not exist in an area in which the image has a uniform intensity, while, if the intensity change does not have enough regularity, the texture cannot be characterized properly, thus an accurate recognition and segmentation is not possible. These regularity and randomness degrees should be measured and compared when a texture is characterized.

The visibility of smoke texture mainly depends on the concentration and distribution of the smoke cloud, which in turn depend on many factors such as the chemical composition of the fuel and combustion rates, the magnitude and direction of the airstream, room temperature, atmospheric pressure and if it is a stratified or uniform environment. Sometimes the texture is more visible in the boundary layer of the cloud of smoke than inside it because of its density, thus it is not possible to capture the texture of the brightly lit and transparent smoke region in the image under analysis. The texture is itself a complicated to measure characteristic. In addition, textures in the real world are affected by variations in lighting, perspective, viewing distance, additive noise, resolution, shadows, possible occlusions, and mutability in appearance over time, among others. Moreover, a robust operator invariant to lighting, rotation, scaling and observation points, among others, is required. The Local Binary Pattern (LBP) is an operator that combines features of statistical and structural texture analysis methods, and can conveniently characterize the regularity and randomness of the pixels in an image. It is independent of linear changes in the average intensity, robust to rotation and supports multi-scale analysis. In addition, it is complemented with contrast information, which enhances the description. These qualities make it suitable for describing smoke texture.

2.3.1. Local Binary Pattern

LBP was originally proposed for texture classification by T. Ojala et al. [12]. This operator compares a central pixel with its neighbors and, if the intensity of the pixel under analysis is smaller than the central pixel of the window, it is assigned a zero, otherwise the central pixel is equal to one, as shown in Equations (13) and (14).

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (13)$$

$$s(g_p - g_c) = \begin{cases} 1 & \text{if } g_p - g_c \geq 0 \\ 0 & \text{if } g_p - g_c < 0 \end{cases} \quad (14)$$

where g_c is the value of the central pixel and $g_p (p = 0, \dots, P - 1)$ is the value of neighboring pixels equally spaced in a circle of radius $R (R > 0)$. Thus, if we have a neighborhood with eight neighbors, we have 256 different patterns. Typically, the histogram of LBP values, given by Equation (15), is used to obtain the feature vectors used to describe the image under analysis:

$$H_{P,R}(k) = \sum_{i=1}^N \sum_{j=1}^M w(LBP_{P,R}^{type}(i, j), k) \quad (15)$$

where

$$w(x, y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

$k \in [0, K]$, K is the number of possible patterns according to the type of LBP used, while N and M are the height and width of the image under analysis, respectively.

Several variations of LBP have been proposed such as uniform LBP, an LBP invariant to rotation, an LBP uniform and invariant to rotation and an LBP that is not redundant [13]. The uniform rotation invariant patterns describe basic structures such as flat areas, points, edges, curves, corners and end of lines, regardless of their orientation. Smoke texture changes constantly without following a definite pattern, therefore does not have a defined orientation and contrast, thus they can be discarded. This is the purpose of using non-redundant LBP patterns that consider identical spatial structures that are complementary in their contrast.

2.3.2. Texture Description

The texture description is performed with a multi-scale analysis using feature vectors derived from the histograms of the distribution of LBP and the Variance of the Local Binary Patterns (LBPV) [14], in different samples and radius, as shown in the following equations:

$$H_{P,R}^{riu2}(k) = \sum_{x=\lceil R \rceil}^{N-\lceil R \rceil} \sum_{y=\lceil R \rceil}^{M-\lceil R \rceil} w(LBP_{P,R}^{riu2}(x, y), k), \quad k \in [0, K] \quad (17)$$

$$HV_{P,R}^{riu2}(k) = \sum_{x=-\lceil R \rceil}^{N-\lceil R \rceil} \sum_{y=-\lceil R \rceil}^{M-\lceil R \rceil} W(LBP_{P,R}^{riu2}(x, y), k), \quad k \in [0, K] \quad (18)$$

where $w(x, y)$ is given by Equation (16) and

$$W(LBP_{P,R}(i, j), k) = \begin{cases} VAR_{P,R}(i, j) & LBP_{P,R}(i, j) = k \\ 0 & \text{otherwise} \end{cases} \quad (19)$$

$$VAR_{P,R} = \frac{1}{P} \sum_{p=0}^{P-1} (g_p - u)^2 \quad (20)$$

with $P = 8$ samples and $R = 1$ pixel distance in the first scale. In addition, an analysis is performed on the second scale with the texture image previously filtered with a Gaussian low-pass filter. For the filter design, Equations (21)–(24) are used, where $w_2 = 3$ is obtained by Equation (23), $\sigma_2 = 0.30$ is obtained by (24) and $R = 2.5$ is obtained by (22). N and M are the width and height of the image block

texture, respectively, and K is the number of possible patterns, nine patterns in this case and additional patterns to the box labeled “miscellaneous”. The radius is given by:

$$r_e = r_{e-1} \left(\frac{2}{1 - \sin(\pi/P_e)} - 1 \right), \quad e \in \{2, \dots, E_s\} \quad (21)$$

where E_s is the total number of scales and P_e is the number of elements at a particular scale. The filter is useful only when the radius is greater than one, thus in the calculations the first radius considered is $r_1 = 1.5$ for $P = 8$. The radio operator on the scale LBP e when $e \geq 2$ is between r_e and r_{e-1} :

$$R_e = \frac{r_e + r_{e-1}}{2} \quad (22)$$

The width and height of each spatial Gaussian filter in the scale e are given by:

$$w_e = 2 \left\lceil \frac{r_e - r_{e-1}}{2} \right\rceil + 1 \quad (23)$$

The standard deviation of the Gaussian filter, considering a standard normal distribution in the scale e , is estimated by:

$$\sigma = \frac{r_e - r_{e-1}}{2 \times \sqrt{-2 \ln(1 - p)}}, \quad p \in [0, 1] \quad (24)$$

where p is the probability that the filter mass falls into the circle area of radius r , which is given by $p = 0.95$.

For the texture description, the histograms of LBP^{riu2} and $LBPV^{riu2}$ are concatenated together with the intensity average μ^I and its standard deviation σ^I resulting in the feature vector with 14 elements. It should also be noted that the fewer elements there are in a feature vector, the more efficient is the separation of different classes.

$$V = \{H_{8,1}^{riu2}, \log(HV_{8,1}^{riu2}), H_{8,2.5}^{riu2}, \log(HV_{8,2.5}^{riu2}), \mu^I, \sigma^I\} \quad (25)$$

2.3.3. Texture Classification

To determine which classification method provides better performance, several classification methods were evaluated, including Multilayer Perceptron (MLP) and Support Vector Machine (SVM). For texture classification, the regions that meet the movement and color requirements are divided into small blocks to analyze their texture using LBP. A block is a candidate for texture analysis if it has at least 50% of the specific color according to Equation (13). Figure 4 shows the results of texture detection, where the red blocks indicate regions that meet the motion and color characteristics and MLP classifies them as smoke texture. The third and fourth ones in Figure 4 shows some noisy blocks, i.e., false positives. However, the size of these regions decrease or remain constant such that they disappear and reappear throughout the video. Smoke zones that are not classified as smoke texture are also observed; this is because the regions within the cloud are smooth and lack texture.

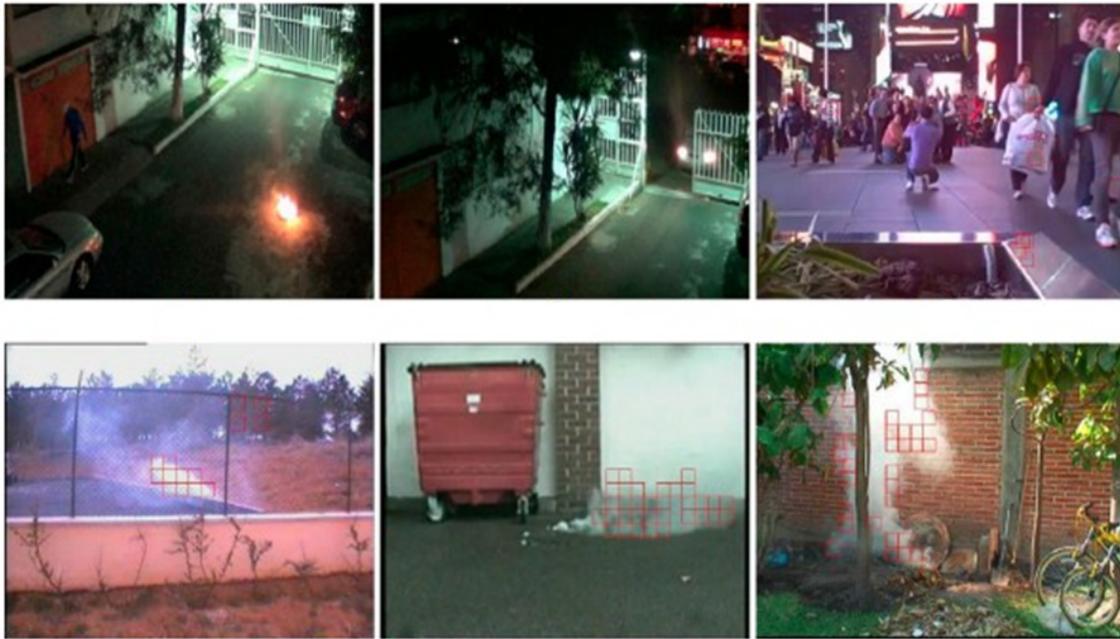


Figure 4. Smoke blocks detected that meet the characteristics of movement and color.

2.3.4. Growth and Direction Analysis

In several cases, the texture of a smoke region's border is poorly defined. Then, it is possible to find some regions with similar texture to smoke or at least equivalently described by LBP. This fact inevitably produces incorrect classification. To improve the system performance and reduce the false alarms, an analysis of the ascendant and expanding moving agglomerations over time is proposed. To this end, a series of images is analyzed considering the difference between frames, searching for regions with smoke texture that increase their size in an upward direction. If such region maintains its upward growth over time, it can be considered as smoke region and then an alarm is triggered. This is accomplished by overlapping and connecting blocks with smoke texture and eliminating noisy blocks, where the noisy blocks are those classified as smoke, but scattered and isolated and disappearing and reappearing during the video sequence. This can be accomplished using the scheme proposed by Millan et al. [7], which is given as follows:

$$\hat{A}_t^k = \begin{cases} \hat{A}_{t-1}^k \cup \hat{A}_t^k & \text{if } \hat{A}_{t-1}^k \cap \hat{A}_t^k \neq \emptyset \\ \hat{A}_t^k & \text{otherwise} \end{cases}, \quad k = 1, 2, \dots, K \quad (26)$$

$$\hat{H}_t = \bigcup_{k=1}^K \hat{A}_t^k \quad (27)$$

where \hat{A}_t^k is an area or region classified as smoke texture in the frame t , \hat{A}_{t-1}^k is a region marked as smoke texture at frame $t - 1$ and k is the label of each region. \hat{H}_t is the resulting image with new overlapping or connected regions. H_t is the image with smoke texture regions, which is analyzed to determine the regions that overlap with the same region of the previous picture at time $t - 1$. The connected regions at time t acquire the same label as the previous region.

2.3.5. Movement Direction

The motion vector is estimated as the difference of the k th centroid estimated at time t and $t - \tau$, that is \hat{A}_t^k and $\hat{A}_{t-\tau}^k$, respectively. If the angle θ_t of motion vector is between 0° and 180° , then it is considered the region moves upward, that is

$$E_t^k(\hat{A}_t^k) = \begin{cases} 1 & \text{if } 0 \leq \theta_t \leq 180 \\ 0 & \text{otherwise} \end{cases} \quad (28)$$

where E_t^k indicates that region \hat{A}_t^k is moving in the accepted angles, as shown in Figure 5.

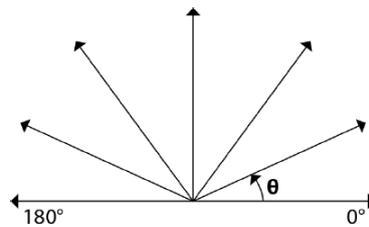


Figure 5. Typical directions for the spread of smoke.

Additionally, the direction of displacement of centroids is accumulated in a histogram $h^k(E^k, \varepsilon)$ during a period $\varepsilon = r * \tau$, whose dominant direction considers the final moving direction of the region.

$$E'^k = \arg \max \{h^k(E^k, \varepsilon)\} \quad (29)$$

where $\arg \max(h)$ is the most frequently depicted direction in the histogram, where 1 is used for upward directions and 0 for the remaining ones.

2.3.6. Growth of Candidate Region

To determine if a region is a smoke candidate region, the number of elements belonging to each region \hat{A}_t^k is estimated as follows:

$$R_t^k = \text{number}((x, y) | (x, y) \in \hat{A}_t^k(x, y)), \quad k = 1, 2, \dots, K \quad (30)$$

where (k) denotes the number of elements in y . The number of elements belonging to the region under analysis must be greater or equal to the number of elements of the same region at time $t - \tau$. That is, the size of the region under analysis must increase or at least maintain the same size such that R_t^k becomes:

$$R_t^k = \begin{cases} 1 & \text{if } R_t^k \geq R_{t-\tau}^k \\ 0 & \text{otherwise} \end{cases} \quad (31)$$

Figure 6 shows a sequence of growing regions, where the upward expansion is analyzed using the temporal overlap of the blocks with smoke texture, where the centroid is marked with a red asterisk.

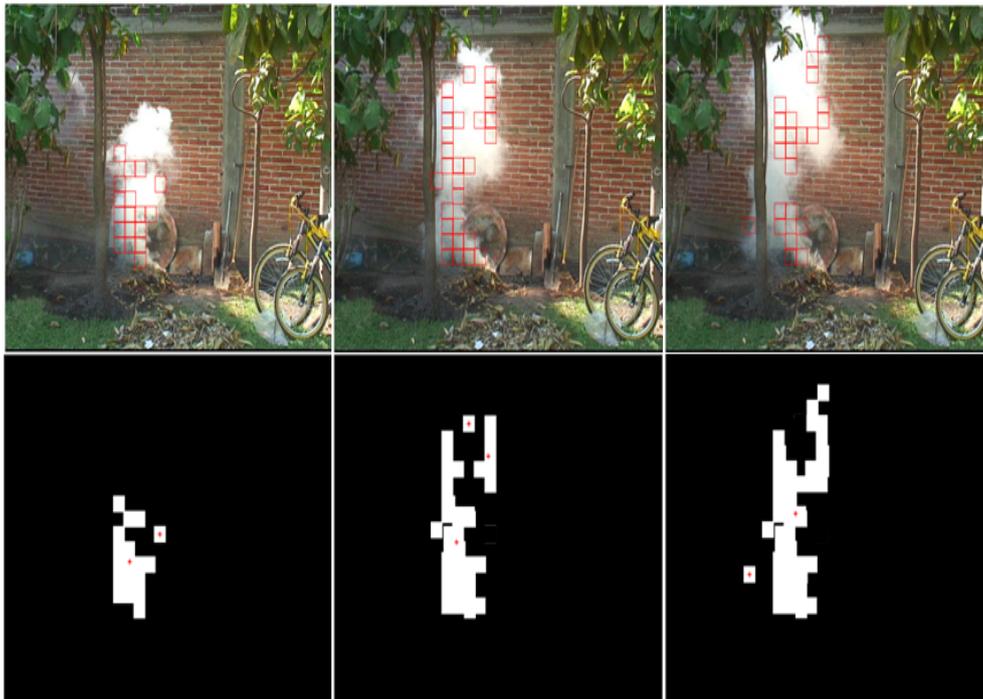


Figure 6. Upward expansion of smoke.

2.3.7. Alarm Activation

To determine whether the fire alarm must be activated, the upward motion and the growth of a specific region (F_t^k) must satisfy:

$$F_t^k = \begin{cases} 1 & \text{if } R_t^k \wedge E_t^k \\ 0 & \text{otherwise} \end{cases} \quad (32)$$

where \wedge denotes the logical operator AND, during a time interval given in advance. The alarm is activated if Equation (33) is satisfied during monitoring time ε , given by $\varepsilon = r \times \tau$

$$FSA = \bigcup_{t=1}^{\varepsilon} F_t^k \quad (33)$$

where FSA means fire smoke alarm, τ is the previous frame number, and r is a factor that determines the monitoring time. If r is small, the alarm will be more sensitive prone to false alarms, while, if it is very large, it will take longer to be activated.

3. Results

To evaluate the effectiveness, robustness and efficiency of the proposed system, several smoke detections were performed using standard databases [15,16]. To this end, two supervised classifiers, namely *MLP_BP* and SVM, were used, because, when the input and output are known, the neural network and SVM are very efficient [17]. *MLP_BP* was used with different inputs and outputs depending on the experiment, using a *tansig* (hyperbolic tangent sigmoid transfer function), 30 hidden layer neurons, 0.05 minimum error and 1000 iterations in the case of non-convergence. The SVM was configured with a Gaussian radial basis and default values in the library *libsvm* [18]. In the first test only, the texture classification capability of Yuan's method [14] and the proposed scheme, given by Equation (34) with mean value μ^I and standard deviation σ^I of the pixel's intensity, were compared. Two multilayer perceptron back propagation neural networks (*MLP_BP*) were trained with 1000 images that contain smoke and 1000 that contain different textures, segmented into

square blocks of 48 pixels, while 840 images not used during training, 300 containing smoke and 540 not containing smoke, were used for testing. The evaluation results obtained are shown in Table 2.

$$V = \{H_{8,1}^{riu2}, \log(HV_{8,1}^{riu2}), H_{8,2.5}^{riu2}, \log(HV_{8,2.5}^{riu2}), \mu^I, \sigma^I\} \quad (34)$$

Table 2. Comparison of the proposed method and Yuan’s method using only the texture feature.

Method	Classification	False Positives	False Negatives	Activation Time
Proposed	94.77%	4.40%	0.83%	7.76 s
Yuan	92.63%	2.26%	5.11%	19.26 s

Table 2 shows that the proposed system presented a higher classification rate and less time activation than the scheme proposed by Yuan [2], which uses a three-level pyramid decomposition reducing the image size at each level resulting in a characteristic vector of 210 elements. On the other hand, in the proposed system, only two levels are used without image sub-sampling and a characteristic vector of 42 items.

In the second test, 2800 images of 22×22 pixels divided into two groups were used, 1400 images containing smoke and 1400 not containing smoke. These dimensions were chosen because this image size has a good statistical consistency and is less prone to noise from other textures captured in a larger block. In total, 700 images with smoke and 700 smokeless ones were used for training and the remaining 1400 were used for testing. *MLP_BP* and SVM were used, together with the kernel of Gaussian radial basis function, using only the texture. Evaluation results provided in Table 3 show that a similar performance was obtained using both classifiers. For the image description, Equation (35) was used, reducing the size of the feature vector because only 12 elements were required.

$$V = \{H_{8,1}^{riu} + \log(HV_{8,1}^{riu2}) + H_{8,2.5}^{riu2} + \log(HV_{8,2.5}^{riu2}), \mu^I, \sigma^I\} \quad (35)$$

Table 3. Comparison of a database with two classifiers using only the texture feature.

Method	Average Classification	False Positives	False Negatives
MLP_BP	92.80%	2.30%	4.90%
SVM	92.14%	5.00%	2.85%

The third test was a full evaluation of the proposed algorithm, i.e., images that meet the three characteristics of smoke described above. Movement, color and texture were considered to look for candidate regions to smoke. These regions were then monitored for a certain time looking for upward expansions. If any of these regions maintained an upward growth during a certain time, it was considered as a smoke region and then the alarm was activated. To obtain the evaluation results shown in Table 3, the *MLP_BP* was trained using 250 images with smoke and 250 smokeless images of size 22×22 pixels whose features vector were obtained using Equation (35). These 500 images were obtained from the 12 videos shown in Figure 7. The performance of the proposed algorithm was evaluated using six videos with smoke and six smokeless videos, which are available at [12,13]. Figure 7 shows some of the videos used in the evaluation. The first line shows the videos with smoke and the second line the videos without smoke. Here, the smokeless scenarios had properties and characteristics similar to those with smoke, such as color, texture or upward movement.



Figure 7. Videos used in the testing of smoke. The first line shows videos with smoke and the second line smokeless videos.

Tables 4 and 5 show the smoke texture evaluation results obtained using the videos with smoke and those without smoke, respectively. In both situations, 700 images with smoke and 700 without smoke were used, each with 22×22 pixels. Evaluation results show that the proposed system provided false positive and false negative rates lower than 1% and a detection rate higher than 98%. In addition, the proposed system activated the alarm in about 6 s when the video contained smoke, while it was not activated in the case of smokeless videos. Here, the false positives were caused by image blocks with smoke textures that were not connected to regions that grow or that appeared and disappeared over the time, thus did not activate the alarm. On the other hand, the false negatives were due to blocks with smoke, the color of which was almost transparent and the characterized texture corresponded mainly to backward objects. Table 6 shows some metrics of the algorithm performance.

Table 4. Percentages texture classification with smoke and smoke detection alarm activation.

Smoke Videos	Frame Rate (FPS)	Smoke Classification	False Positives	False Negatives	Alarm Activation	Activation Time
Bard	30	98.75%	0%	1.25%	YES	4 s
Dump	10	98.58%	1.42%	0%	YES	6.5 s
House	30	98.45%	0.83%	0.71%	YES	4 s
Trailer	12	99.68%	0.19%	0.23%	YES	4 s
Window	15	98.64%	0.62%	0.74%	YES	13.5 s
Grill	25	98.81%	0.94%	0.25%	YES	4 s
Average	–	98.81%	0.66%	0.53%	–	6 s

Table 5. Percentage of smokeless texture classification.

Videos without Smoke	Frame Rate (FPS)	Classification without Smoke	False Positives	False Negatives	Alarm Activation
Gray Car	30	97.85%	2.15%	0%	NO
Day Fount	30	99.43%	0.57%	0%	NO
Night Fount	30	99.83%	0.16%	0%	NO
Pigeons	30	99.02%	0.98%	0%	NO
Tunnel Lopez	25	99.73%	0.27%	0%	NO
Night Car	10	97.76%	2.23%	0%	NO
Average	–	98.94%	1.06%	0%	–

Table 6. General metrics of the algorithm performance.

Accuracy	Precision	Recall	F1 Score
98.79%	98.15%	99.42%	98.78%

The alarm activation time depends on the threshold value ϵ , which is selected to determine if there is smoke in the scene. The value of ϵ depends on the desired number of frames required to activate the

fire alarm. If we consider a threshold that is too small, the system may produce a large number of false alarms. On the other hand, if the threshold value is too large, the system response may not be able to respond as quickly as required in early fire detection applications. In our experiments, a threshold of 4 s was selected such that the time period for the movement analysis was 0.5 s, i.e., $\tau = fps/2$, where fps is the number of frames per second. In the videos without smoke, the alarm was not activated because false positives that may occur were not connected in clusters. Note that these detection rates heavily depended on the training and the parameters used in the neural network that must be considered: the lighting conditions, the contrast in the scene, occlusions in the foreground and various factors that affect smoke detection. The boundary regions containing the smoke column were properly detected by the proposed algorithm because of their characteristics, i.e., expansion, dilution, translucency and texture.

4. Conclusions

This study proposes an early smoke detection method using image from video sequences, which is based on detecting the motion, color and texture properties, estimated in cascade to delimit the candidate smoke region, where the texture is analyzed using the LBP and LBPV operators. Finally, if the candidate regions maintain an upward expansion, the presence of smoke is determined and an alarm is activated. The evaluation results show that the proposed system provides a good detection rate of smoke, with enough robustness against several scenarios. Evaluation results show that the proposed system provides fairly good detection results: when the input videos contained smoke, the alarm was turned on in all cases, while, in the videos without smoke, the alarm was not turned on. Some factors that affected detection were low quality videos and the distance to the target, i.e., the viewing distance to the place in which smoke was produced. On the other hand, the main goal is its application in places such as offices, houses, and streets as well as in surveillance; however, its main drawback is the distance between the camera and smoke, because, in the case that the fire event occurred far from the camera, the methodology proposed is not accurate enough to detect it, thus this method cannot be applied in a forest fire. Even though the algorithm provides good results in the environments described above, the performance of the proposed algorithm must be analyzed in the cases of dust or aerosols as well as when there are changes in lighting, especially at night. These issues are under study and will be reported as part of future work.

Author Contributions: All authors contributed equally to this work.

Funding: This research received no external funding.

Acknowledgments: This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 700326 (website: <http://ramses2020.eu>). In addition, the authors thank the National Science and Technology Council of Mexico (CONACyT), and the Instituto Politecnico Nacional for the financial support for this research.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chen, T.; Yin, Y.; Huang, S.; Ye, Y. The smoke detection for early fire-alarming system base on video processing. In Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia, Pasadena, CA, USA, 18–20 December 2006; pp. 427–430.
2. Yuan, F. A fast accumulative motion orientation model based on integral image for video smoke detection. *Pattern Recognit. Lett.* **2008**, *29*, 925–932. [[CrossRef](#)]
3. Töreyn, B.U.; Dedeoğlu, Y.; Cetin, A.E. Wavelet based real-time smoke detection in video. In Proceedings of the 2005 13th European Signal Processing Conference, Antalya, Turkey, 4–8 September 2005; pp. 1–4.
4. Toreyn, B.U.; Dedeoglu, Y.; Cetin, A.E. Contour based smoke detection in video using wavelets. In Proceedings of the 2006 14th European Signal Processing Conference, Florence, Italy, 4–8 September 2006; pp. 1–5.

5. Yu, C.; Zhang, Y.; Fang, J.; Wang, J. Texture analysis of smoke for real-time fire detection. In Proceedings of the 2009 Second International Workshop on Computer Science and Engineering, Qingdao, China, 28–30 October 2009; Volume 2, pp. 511–515.
6. Maruta, H.; Nakamura, A.; Kurokawa, F. A new approach for smoke detection with texture analysis and support vector machine. In Proceedings of the 2010 IEEE International Symposium on Industrial Electronics, Bari, Italy, 4–7 July 2010; pp. 1550–1555.
7. Millan-Garcia, L.; Sanchez-Perez, G.; Nakano, M.; Toscano-Medina, K.; Perez-Meana, H.; Rojas-Cardenas, L. An early fire detection algorithm using IP cameras. *Sensors* **2012**, *12*, 5670–5686. [[CrossRef](#)] [[PubMed](#)]
8. Luis, J.A.; Galán, J.A.G.; Espigado, J.A. Low Power Wireless Smoke Alarm System in Home Fires. *Sensors* **2015**, *15*, 20717–20729, doi:10.3390/s150820717. [[CrossRef](#)] [[PubMed](#)]
9. Wu, X.; Lu, X.; Leung, H. A video based fire smoke detection using robust AdaBoost. *Sensors* **2018**, *18*, 3780. [[CrossRef](#)] [[PubMed](#)]
10. Mäenpää, T.; Ojala, T.; Pietikäinen, M.; Soriano, M. Robust texture classification by subsets of local binary patterns. In Proceedings of the 15th International Conference on Pattern Recognition, Barcelona, Spain, 3–7 September 2000.
11. Mulholland, G.W. Smoke production and properties. In *SFPE Handbook of Fire Protection Engineering*; Springer: New York, NY, USA, 2002; Volume 3.
12. Ojala, T.; Pietikäinen, M.; Mäenpää, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *7*, 971–987. [[CrossRef](#)]
13. Song, K.-C.; Yan, Y.-H.; Chen, W.-H.; Xu, Z. Research and perspective on local binary pattern. *Acta Autom. Sin.* **2013**, *39*, 730–744. [[CrossRef](#)]
14. Yuan, F. Video-based smoke detection with histogram sequence of LBP and LBPV pyramids. *Fire Saf. J.* **2011**, *46*, 132–139. [[CrossRef](#)]
15. Bilkent EE Signal Processing group. Sample Fire and Smoke Video Clips. Available online: <http://signal.ee.bilkent.edu.tr/VisiFire/Demo/SampleClips.html> (accessed on 1 September 2018).
16. SEPI ESIME Culhuacan. Articles Videos Smoke@ONLINE. 2016. Available online: http://www.posgrados.esimecu.ipn.mx/index.php?option=com_content&view=article&id=222&itemid=0 (accessed on 1 September 2018).
17. Olivares-Mercado, J.; Aguilar-Torres, G.; Toscano-Medina, K.; Nakano-Miyatake, M.; Perez-Meana, H. GMM vs SVM for Face Recognition and Face verification. In *Reviews, Refinements and New Ideas in Face Recognition*; IntechOpen: Rijeka, Croatia, 2011; pp. 1–338.
18. Chang, C.C.; Lin, C.J. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 27. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).