

Technical Note

Early Fire Detection Based on Aerial 360-Degree Sensors, Deep Convolution Neural Networks and Exploitation of Fire Dynamic Textures

Panagiotis Barmpoutis ^{1,*}, Tania Stathaki ², Kosmas Dimitropoulos ¹  and Nikos Grammalidis ¹ 

¹ Information Technologies Institute, Centre for Research and Technology Hellas, 60361 Thessaloniki, Greece; dimitrop@iti.gr (K.D.); ngramm@iti.gr (N.G.)

² Department of Electrical and Electronic Engineering, Faculty of Engineering, Imperial College London, London SW7 2AZ, UK; t.stathaki@imperial.ac.uk

* Correspondence: panbar@iti.gr

Received: 15 August 2020; Accepted: 23 September 2020; Published: 28 September 2020



Abstract: The environmental challenges the world faces have never been greater or more complex. Global areas that are covered by forests and urban woodlands are threatened by large-scale forest fires that have increased dramatically during the last decades in Europe and worldwide, in terms of both frequency and magnitude. To this end, rapid advances in remote sensing systems including ground-based, unmanned aerial vehicle-based and satellite-based systems have been adopted for effective forest fire surveillance. In this paper, the recently introduced 360-degree sensor cameras are proposed for early fire detection, making it possible to obtain unlimited field of view captures which reduce the number of required sensors and the computational cost and make the systems more efficient. More specifically, once optical 360-degree raw data are obtained using an RGB 360-degree camera mounted on an unmanned aerial vehicle, we convert the equirectangular projection format images to stereographic images. Then, two DeepLab V3+ networks are applied to perform flame and smoke segmentation, respectively. Subsequently, a novel post-validation adaptive method is proposed exploiting the environmental appearance of each test image and reducing the false-positive rates. For evaluating the performance of the proposed system, a dataset, namely the “Fire detection 360-degree dataset”, consisting of 150 unlimited field of view images that contain both synthetic and real fire, was created. Experimental results demonstrate the great potential of the proposed system, which has achieved an F-score fire detection rate equal to 94.6%, hence reducing the number of required sensors. This indicates that the proposed method could significantly contribute to early fire detection.

Keywords: remote sensing; fire detection; flame detection; smoke detection; natural disasters; 360-degree sensors; deep convolutional neural networks

1. Introduction

The increasing occurrence of large-scale forest fires in modern society significantly impacts society and communities in terms of remarkable losses in human lives, infrastructures and properties [1]. Depending on burn severity, wildfires also impact environment and climate change, increasing the released quantity levels of CO₂, soot and aerosols and damaging the forests that would remove CO₂ from the air. This results in extremely dry conditions, increasing the risk of wildfires. Furthermore, forest fires lead to runoff generation and to major changes to the soil infiltration [2]. To this end, computer-based early fire warning systems that incorporate remote sensing technologies have attracted

particular attention in the last decade. These early detection systems use individual or networks of ground sensors, unmanned aerial vehicles (UAVs) and satellite-based systems consisting of active or passive sensors. Ground sensors are inflexible, and they need to be carefully placed in order to ensure adequate visibility. Thus, they are usually allocated in watchtowers for monitoring high-risk situations. On the other hand, UAVs cover wider areas and are flexible, allowing the change of monitoring area, but they are affected by weather conditions, while in many cases, their flight time is limited [3]. Satellite surveillance is capable of the detection of both small and large fires. However, in many cases, satellite visit periods are infrequent and satellite data are not immediately available to the early fire detection systems. For example, the two Moderate Resolution Imaging Spectroradiometer (MODIS) satellites cover the Earth four times per day, and Landsat-8 has a 16 day revisit time [4]. Furthermore, sometimes, cloud coverage can reduce the availability of usable satellite data [3]. Regarding the type of sensors that systems are equipped with, active sensor systems provide their own source of energy for the fire detection, and passive sensors detect natural energy that is emitted or reflected by the scene being observed [5]. Most of the active sensors' systems operate in the microwave portion of the electromagnetic spectrum, and passive systems operate in the visible, infrared and microwave portions of the electromagnetic spectrum [5–9].

The main fire detection challenges of the aforementioned systems lie in the modelling of the chaotic and complex nature of the fire phenomenon, in the separation of the fire-emitted radiance from the reflected background radiance and in the occurrence of large variations of either flame or smoke appearance [3,6]. Thus, to address this problem, many fire algorithms combining features that are related with the physical properties of flame, such as color [10,11], spectral [12], spatial [13] and texture characteristics [14], have been developed. More specifically, in Reference [10], a method for fire detection using ordinary motion, color clues and flame flicker in the wavelet domain was proposed. Dimitropoulos et al. [11] extracted more spatio-temporal features such as color probability, contour irregularity, spatial energy, flickering and spatio-temporal energy. A multi-sensor surveillance system, combining optical and infrared cameras and creating wireless sensor networks, was proposed by Grammalidis et al. [12]. Similar to this, a wireless multi-sensor network was developed by Lloret et al. [15] aiming towards the early detection of forest fires. Extending their previous work [13], Dimitropoulos et al. [14] used ground optical cameras and exploited the information that is provided by fire dynamics. Specifically, they adapted spatio-temporal modelling and they performed a dynamic texture analysis based on linear dynamical systems (LDSs), demonstrating that this fusion leads to higher detection rates and less false alarms. Static and dynamic texture analysis of flame for forest fire detection are combined in Reference [16]. Dynamic texture features were derived using two-dimensional (2D) wavelet decomposition in temporal domain and three-dimensional (3D) volumetric wavelet decomposition. Compared to optical or infrared sensors, measuring instruments working at the microwave range are more efficient for fire detection in conditions of insufficient visibility [17]. Varotsos et al. [18] deployed a platform IL-18 equipped with passive microwave radiometers of 1.43, 13.3 and 37.5 GHz frequencies and proposed a decision-making system in order to achieve early detection of forest fires. In contrast to the ground sensor warning systems, many researchers have focused on the several operational satellites aiming for early fire detection on a global scale [19,20]. Jang et al. [21] used the Himawari-8 geostationary satellite data that collects data in 16 bands from visible to infrared wavelengths and a three-step forest fire detection algorithm in order to achieve forest fire detection. Polivka et al. [22] proposed a multi-step method for the detection of fire pixels by employing the Visible Infrared Imaging Radiometer Suite (VIIRS) day-night band (DNB) in conjunction with the mid-wave infrared (MWIR) band.

More recently, the development of deep learning algorithms based on Convolution Neural Networks (CNNs) for fire detection and classification of fire and smoke-colored objects has made contributions to the significant increment of detection accuracy. Compared to the previously discussed image and video processing methods, fire detection using the same remote sensors and methodology and based on deep learning is often more reliable and accurate. More specifically, Sharma et al. [23]

used RGB images and re-tuned two pre-trained CNNs, based on VGG16 and ResNet50 architectures in order to develop a fire detection warning system. It is worth mentioning that for the training, they created an unbalanced dataset including more non-fire images. The main advantage of CNNs is the fact that they automatically extract and learn features [24,25]. Extending deep learning approaches, Barmpoutis et al. [26] combined the power of modern deep learning networks with multidimensional texture analysis based on higher-order LDSs (h-LDSs). The use of UAVs in remote sensing platforms for environmental monitoring has become increasingly popular. Chen et al. [27] used optical and infrared sensors and data in order to train a CNN for fire detection, while Zhao et al. [28] used a UAV equipped with global positioning systems (GPS) and deployed a 15-layered self-learning deep convolutional neural network architecture for fire feature extraction and classification. In Reference [20], satellite optical in combination with neural network architecture were used for forest fire detection.

Unlike the flame, optical cameras with good visibility can easily capture smoke in the first few minutes of a fire from a long distance (more than 50 km away) due to the fact that it moves upwards [4,29]. However, the reliable identification of smoke is also still an open research issue since there are many objects in nature that have similar characteristics with smoke. In addition, the large variations of smoke's appearance as well as environmental changes, including clouds and shadows, make the task of smoke detection even more difficult. More specifically, regarding the fire detection based on the identification of smoke occurrence, Filonenko et al. [30] presented a smoke detection method relying on shape features and color information of smoke regions for surveillance cameras. Their work improved the processing speed of both low-resolution and high-definition videos with the utilization of both CPU and General-Purpose Graphics Processing Unit (GPGPU). In another approach, based on contour analysis and edge regions with decreasing sub-band wavelet energies, Toreyin et al. [31] developed an algorithm for indoor or outdoor smoke detection. Furthermore, an algorithm which extracts optical flows and motion features for the discrimination between smoke regions and similar moving objects was proposed by Chunyu et al. [32]. Based on the note that smoke is a non-rigid dynamical object, Barmpoutis et al. [33] and Dimitropoulos et al. [29] efficiently modeled smoke regions using dynamic texture analysis. Sudhakar et al. [34] proposed a method for forest fire detection through UAVs using color identification and smoke motion analysis. Additionally, Yuan et al. [35] combined two color spaces and used an extended Kalman filter in order to perform smoke detection. More recently, Dai et al. [36] noticed that the smoke screen area changes rapidly in infrared smoke interference image sequences and they used infrared sensors and a super-pixel segmentation method in order to detect smoke. Subsequently, Wang et al. [37], using satellite data, resampled from Suomi National Polar-orbiting Partnership (Suomi NPP) the VIIRS day–night band pixel radiances to the footprint of M-band pixels in order to link visible energy fraction to the emission factors and to estimate the fire burning phase.

Additionally, many deep learning algorithms have been developed for smoke detection, providing a way of avoiding damages caused by fire. More specifically, Yin et al. [38] proposed a deep normalization and a CNN with 14 layers, while Tao et al. [39] adapted the architecture of AlexNet. Luo et al. [40] proposed a smoke detection algorithm combining motion characteristics of smoke and a CNN. As the big challenge of deep learning algorithms is the lack of a significant amount of labelled data, Xu et al. [41] implemented an early smoke detection method using synthesized smoke images for the training of the region-based Faster R-CNN, which provides satisfying results. Similar work can also be seen in Reference [42], where a method of synthesizing smoke images based on the expansion of the domain-invariant smoke feature was introduced.

Although several technologies based on different sensors have been proposed for fire surveillance covering different needs, the developed systems and detection algorithms are far from optimal in respect of the complexity of systems that use a collection of sensors, the higher computational time that is required and the accuracy rates. Moreover, ground- or UAV-based systems have limited field of view or they use specialized aerial hardware with complex standard protocols for data collection,

and satellite data are not immediately available, limiting their potential eventual widespread use by local authorities, forest agencies and experts.

Thus, in this paper, taking into account the advantage of the cost of UAVs having significantly decreased and the fact that recently, many high-resolution omnidirectional cameras have been developed, we propose a new remote sensing system in order to achieve fire detection at an early stage. Compared to conventional cameras, omnidirectional cameras can cover a wider field with only a single camera, showing their great potential on surveillance applications [43]. Thus, given the urgent priority around the protection of the value and potential of forest ecosystems and global forest future, we aim to extend our preliminary work [44] and the already well-established fire detection systems [13,14,26,29,33] by proposing a novel 360-degree remote sensing system, incorporating the recently introduced 360-degree sensors, DeepLab v3+ models and a validation approach, taking into account the environmental appearance of the examined instant captures for early forest fire detection. More specifically, this paper makes two main contributions:

- We propose a novel early fire detection remote sensing system using aerial 360-degree digital cameras in an operationally and time efficient manner, aiming to overcome the limited field of view of state-of-the-art systems and human-controlled specified data capturing.
- A novel method is proposed for fire detection based on the extraction of stereographic projections and aiming to detect both flame and smoke through two deep convolutional neural networks. Specifically, we initially perform flame and smoke segmentation, identifying candidate fire regions through the use of two Deeplab V3+ models. Then, the detected regions are combined and validated taking into account the environmental appearance of the examined instant capture test image.

Finally, in order to evaluate the efficiency of the proposed methodology, we used seven datasets, including a created 360-degree dataset, namely the “Fire detection 360-degree dataset”, that consists of 150 images of forest and urban areas that contain synthetic and real fire events.

The remainder of this paper is organized as follows: Section 2 describes the methodology proposed in this work. Section 3 describes the protocol followed for the experimental analysis and presents and discusses the experimental results. Finally, Section 4 summarizes the main conclusions of this work and points to future directions.

2. Materials and Methods

The framework of the proposed methodology is shown in Figure 1. In this figure, an aerial 360-degree remote sensing system is used in order to capture unlimited field of view images. Initially, two DeepLab v3+ networks are trained towards the aim of the identification of candidate fire (flame and smoke) regions (Figure 1a). Then, once equirectangular raw images are acquired, they are converted to stereographic projection format images (Figure 1b). Subsequently, these images are fed to the trained networks and the candidate detected fire regions are validated taking into account the environmental appearance of the examined test images (Figure 1c).

Specifically, exploiting the 360-degree sensors, in order to reduce false-positive fire alarms after the DeepLab V3+-based segmentation, the detected regions are divided into blocks and these are compared with parts of the test image that potentially cause false alarms. To this end, to filter out the possibly false-positives, multidimensional texture features are estimated, and a clustering approach is used in order to accurately identify the fire regions.

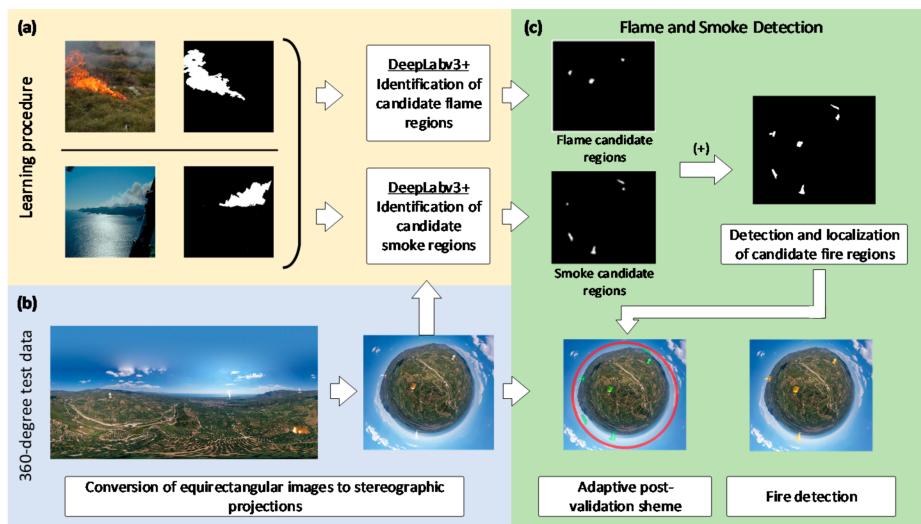


Figure 1. The proposed methodology: (a) Learning procedure of two DeepLab v3+ networks, (b) conversion of equirectangular raw images to stereographic projection format images, (c) post-validation of the fire detected images.

2.1. Data Description

For the evaluation of the proposed method, we used 7 datasets consisting of more than 4000 images that contain both flame and smoke events. More specifically, for the training of the proposed methodology, we used the Corsican Fire Database (CFDB) [45,46], the FireNet dataset [47], two datasets created by the Center for Wildfire Research [48,49] and two more publicly available smoke datasets [50,51]. Furthermore, we created a test 360-dataset (Figure 2), namely the “Fire detection 360-degree dataset” [52], consisting of 150 360-degree equirectangular images of forest and urban areas that contain synthetic and real fire events.



Figure 2. Sample testing images of the “Fire detection 360-degree dataset”.

More specifically, in order to capture the equirectangular images, we used a 360-degree camera (sensor type: CMOS, sensor size 1/2.3") mounted on an UAV equipped with GPS. Subsequently, in order to create the synthetic fire events, we reproduced synthetic video frames [13] solving linear system equations and estimating the tensor generated at time k when the system is driven by random

noise, V . For more details regarding dynamic texture synthesis, we refer the reader to reference [53]. Then, the synthesized data was adapted to the 360-degree images and the size of fires was suitably manually adjusted with regards to the distance and the assumed start time of fire. To the best of our knowledge, and as 360-degree digital camera sensors are a newly introduced type of camera, there is not any dataset consisting of 360-degree images that contain fire.

2.2. Stereographic Projection of Equirectangular Raw Projections

Nowadays, as 360-degree cameras become widely available, 360-degree data has become more popular. The 360-degree images bring unlimited view and can be represented into various projections. The most widely used of them are the equirectangular projection, stereographic projection, rectilinear projection and the cubemap projection [54]. Each type of 360-degree image projections has its own unique features.

Equirectangular projection is the most popular projection which maps a full 3D scene onto a 2D surface. The equirectangular raw images are captured by 360-degree cameras, with the aspect ratio of height (h): width (w) equal to 1:2. For the task of fire detection and in order to avoid false alarms due to the existence of distortions in the equirectangular images, we convert these images to stereographic projections (Figure 3). Stereographic projection performs well at minimizing distortions [54], and such type of projection concentrates the target object into the center, making the response for the task of early fire detection easier, as the sky region, including clouds and sunlight reflections, are arranged in the peripheral area, which can easily be used for the rejection of false-positives (Figure 3b). Furthermore, it is worth mentioning that if the parameter zf and UAV flight altitude remain constant, the object is always in the same position and has the same size proportion. The stereographic projection is shown in Figure 4 and the coordinates $p'(x', y')$ are calculated as follows:

$$x' = \frac{w}{2} - \left(-\frac{\text{longitude}}{2\frac{\pi}{w}} \right) \quad (1)$$

$$y' = \frac{h}{2} - \left(\frac{\text{latitude}}{2\frac{\pi}{w}} \right) \quad (2)$$

where *longitude* and *latitude* are the polar coordinates:

$$\text{latitude} = 2 \tan^{-1} \left(\frac{\sqrt{\left(x - \frac{w}{2}\right)^2 + \left(y - \frac{h}{2}\right)^2}}{zf} \right) \quad (3)$$

$$\text{longitude} = \tan^{-1} \left(\frac{y - \frac{h}{2}}{x - \frac{w}{2}} \right) - \frac{\pi}{4} \quad (4)$$

and the zf is a fitting parameter of the stereographic projection to equirectangular image dimensions.



Figure 3. Projections of 360-degree images: (a) equirectangular, (b) stereographic projection.

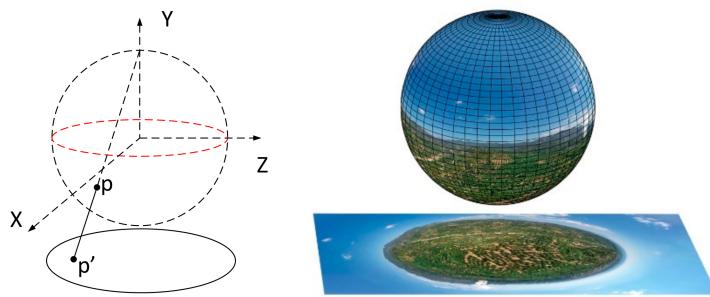


Figure 4. Stereographic projection of 360-degree images.

2.3. Detection and Localization of Candidate Fire Regions

A popular DeepLab V3+ model [55] for semantic segmentation is employed in this work for the detection and localization of candidate fire regions. The DeepLab models have been extensively used in the task of semantic image segmentation and tested on large volumes of image datasets. This model provides a capability in learning multi-scale contextual features through atrous spatial pyramid pooling (ASPP) and uses a decoder module for the refinement of the segmentation results, especially along object boundaries. The applied model utilizes an ImageNet pretrained InceptionResNet v2 as the main feature extractor, which has been proven to be a robust network in the image recognition field. It is worth mentioning that the InceptionResNet v2 outperforms ResNet-101 on the Pascal VOC2012 validation set and it has been identified as a strong baseline [56]. In this study, two DeepLab V3+ networks are trained, and a modified loss function is used in order to adjust the model to better deal with candidate fire regions' detection task. The modified loss function of the DeepLab model is as below:

$$\text{Loss} = - \sum_{p=1}^N w_p r_p \log(t_p) \quad (5)$$

where w_p , r_p and t_p denote the weighting factors, the reference values and the predicted values at pixel p respectively, and N is the total number of pixels. Regarding the weighting factors, we set the $w_p = 3$ when p is a fire pixel and otherwise, we set $w_p = 1$. Thus, we aim to force the model to effectively detect the total number of fire events in each image. Finally, a median value in a 7-by-7 neighborhood was applied to the output of the DeepLab v3+ models.

2.4. Adaptive Post-Validation Scheme

For the validation of the identified candidate fire regions and aiming to decrease the high false alarm rates caused by natural objects which present similar characteristics with fire, we propose a new adaptive scheme through the division of the candidate regions into blocks with size $n \times n$ and the comparison of them with the training dataset and with specific regions of each test image. More specifically, we specified that around the horizon level, there is a number of natural objects like clouds, increased humidity and sunlight reflections that have similar appearance to flame or smoke. Therefore, we divided training images into blocks ($n \times n$), and in the stereographic projection images, we divided a circular formation around the horizon level into blocks ($n \times n$ —we set $n = 16$ based on our previous research [26]) (Figure 5). In this, we excluded the blocks that were identified as fire events from the fire segmentation part.

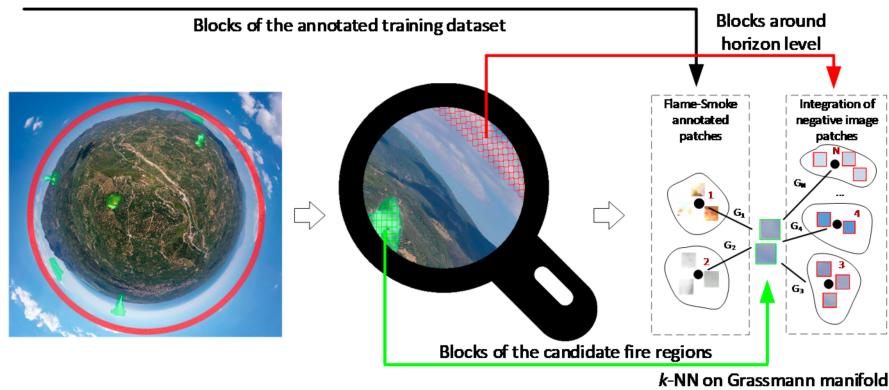


Figure 5. The proposed adaptive validation scheme.

Subsequently, we consider that candidate flame and smoke regions and the created blocks of the training dataset and of the region around the horizon level contain spatially varying visual patterns that exhibit certain stationarity properties [26]. These are modelled through the following linear dynamical system considering them as a multidimensional signal evolving in the spatial domain:

$$x(t+1) = Ax(t) + Bv(t) \quad (6)$$

$$y(t) = \bar{y} + Cx(t) + w(t) \quad (7)$$

where $x \in \mathbb{R}^n$ is the hidden state process, $y \in \mathbb{R}^d$ is the observed data, $A \in \mathbb{R}^{n \times n}$ is the transition matrix of the hidden state and $C \in \mathbb{R}^{d \times n}$ is the mapping matrix of the hidden state to the output of the system. The quantities $w(t)$ and $Bv(t)$ are the measurement and process noise respectively, with $w(t) \sim N(0, R)$ and $Bv(t) \sim N(0, Q)$, while $\bar{y} \in \mathbb{R}^d$ is the mean value of the observation data [26]. Then, assuming that the tuple $M = (A, C)$ describe each block (Figure 6), we estimate the finite observability matrix of each dynamical system, $O_m^T(M) = [C^T, (CA)^T, (CA^2)^T, \dots, (CA^{m-1})^T]$ and we apply a Gram-Schmidt orthonormalization procedure [57], i.e., $O_m^T = GR$, in order to represent each block descriptor as a Grassmannian point, $G \in \mathbb{R}^{mT \times 3}$ [26].

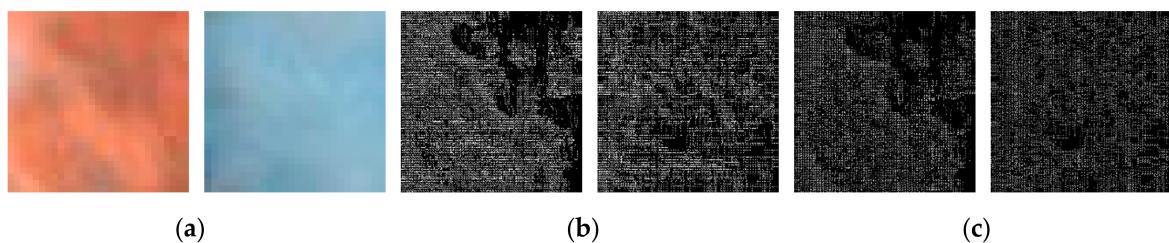


Figure 6. Visualization of: (a) candidate fire and smoke blocks, (b) transition matrices A and (c) mapping matrices C.

For the comparison of candidate fire regions' blocks with the training dataset and around the horizon-level blocks, the distance between them needs to be estimated. Furthermore, in order to solve the non-linearity of the problem, each candidate fire region, through its division into blocks, is considered as a cloud of points on the Grassmann manifold. To this end, we estimated k clusters of blocks of the training dataset and blocks around the horizon level of each test image by the Karcher mean algorithm [58,59] and we applied a k -NN (k -Nearest Neighbors) approach, with each candidate fire block being assigned to the nearest neighbor cluster. To address the k -NN problem, we apply the inverse exponential map between two points on the manifold, e.g., G_1 and G_2 , to map the first Grassmannian point to a tangent space of the second one, while preserving the distance between the

points [58]. Thus, using the inverse exponential map, we move from the Grassmann manifold to the Euclidean space. Hence, the dissimilarity metric between G_1 and G_2 can be defined as follows:

$$d(G_1, G_2) = \|\exp_{G_2}^{-1} G_1\|_F \quad (8)$$

where the inverse exponential map, \exp^{-1} , defines a vector in the tangent space of a manifold's point, i.e., the mapping of G_1 to the tangent space of G_2 , and $\|\cdot\|_F$ indicates the Frobenius norm. Then, the candidate flame and smoke regions are assigned to the class most common amongst its k nearest neighbors.

3. Experimental Results

Through the experimental evaluation, we want to demonstrate the superiority of the proposed system using 360-degree cameras. Furthermore, we want to show that the proposed methodology improves the detection accuracy of fires against other state-of-the-art approaches. To this end, initially, we carried out an ablation analysis and then we compared the proposed methodology with a number of widely used methods in order to show the great potential of the proposed methodology. Both analyses were conducted on the developed "Fire 360-degree dataset".

The code of the proposed structure was implemented in Matlab and all calculations were performed on a 12 GB GPU. Furthermore, it is worth mentioning that in order to have a fair comparison, we used the same training and testing set in our experiments.

The performance of early fire detection (e.g., detection of small fires) in 360-degree stereographic projection images was evaluated using two measures, namely, F-score and mean Intersection over Union (mIoU). The IoU, also known as the Jaccard Index, is an effective metric and is defined as the area of overlap between the detected fire region and the ground truth divided by the area of union between the detected fire region and the ground truth:

$$\text{IoU} = \frac{\text{groudTruth} \cap \text{prediction}}{\text{groudTruth} \cup \text{prediction}} \quad (9)$$

F-score is the most used evaluation metric and is defined as the harmonic mean of precision and recall, taking into account the true-positives, the false-positives and the false-negatives detected regions, and it is defined as follows:

$$\text{F score} = 2 \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (10)$$

IoU and F-score values' range is between 0 and 1, where these metric scores reach their best values at 1. It is worth mentioning that for the evaluation of the proposed methodology, we considered IoU as a pixel-based metric, while we considered F-score as a region-based metric. For example, for the calculation of F-score, a flame or smoke region is considered as a true-positive if it contains at least one pixel with actual flame or smoke, respectively. In our analysis, the mean IoU score and F-score were calculated for flame and smoke classes separately and then taking into account the flame and smoke of a fire region as one fire event. More specifically, in order to calculate the F-score, we count the positive detected fire events if only one of flame or smoke are detected.

3.1. Ablation Analysis

The presented fire detection method comprised two main components of flame/smoke segmentation and post-validation processing. Thus, in this section, we elaborate on a more detailed analysis in terms of the contribution of each feature to the fire identification process. The applied DeepLab v3+ model was used to identify the candidate fire regions in a semantic

image segmentation task, while the adaptive post-validation scheme was used to reject the flame or smoke false-positive regions.

The results for early fire detection performance under different configurations, with or without the validation step, are shown in Table 1. It can be seen that the validation scheme improves the overall accuracy of fire detection. This is because the post-validation processing step rejects false-positives that have similar characteristics with flame and smoke. Furthermore, the proposed methodology works better for flame detection as the number of false-positive in smoke detection is higher, mainly due to the existence of clouds with similar characteristics with smoke. Finally, in the case of taking into account either the detection of flame or smoke in order to define an identified region as a fire event, the precision rate achieved is 90.3% and the recall rate achieved is 99.3%, while the F-score rate is 94.6%.

Table 1. Ablation analysis.

	Flame Detection		Smoke Detection		Flame or Smoke Detection			
	mIoU	F-score	mIoU	F-score	mIoU	F-score	Precision	Recall
DeepLab v3+	76.5%	81.3%	65.2%	80.1%	71.2%	81.4%	68.9%	99.3%
Proposed	78.2%	94.8%	70.4%	93.9%	77.1%	94.6%	90.3%	99.3%

mIoU: mean intersection Intersection over union Union.

3.2. Comparison Evaluation

In Table 2, we present the evaluation results of the proposed framework in comparison to six state-of-the-art methods. This analysis revealed improved robustness of the proposed methodology compared to other methods. More specifically, the proposed system towards flame, smoke and either flame or smoke detection achieves F-score rates 94.8%, 93.9% and 94.6% respectively, achieving F-score rate improvements up to 11.4% in fire detection through flame (Faster R-CNN/Grassmannian VLAD encoding: 83.4%, proposed: 94.8%), up to 6.5% in fire detection through smoke (Faster R-CNN/Grassmannian VLAD encoding: 87.4%, proposed: 93.9%) and up to 7.2% in fire detection through either flame or smoke (Faster R-CNN/Grassmannian VLAD encoding: 87.4%, proposed: 94.6%). Similarly, it achieves mIoU rates of 78.2%, 70.4% and 77.1%, achieving mIoU rate improvements up to 3.8% in fire detection through flame (Faster R-CNN/Grassmannian VLAD encoding: 74.4%, proposed: 78.2%), up to 0.5% in fire detection through smoke (Faster R-CNN/Grassmannian VLAD encoding: 69.9%, proposed: 70.4%) and up to 3.3% in fire detection through either flame or smoke (Faster R-CNN/Grassmannian VLAD encoding: 73.8%, proposed: 77.1%). It is worth mentioning that the method that combines Faster R-CNN and Grassmannian VLAD (Vector of Locally Aggregated Descriptors) encoding [26] produces the best fire detection results after the proposed methodology. It could be explained by the fact that the fire regions that are detected from the Faster R-CNN are validated through their projection to a Grassmannian space and their representation as a cloud of points on the manifold.

Table 2. Comparison results.

	Flame Detection		Smoke Detection		Flame or Smoke Detection	
	mIoU	F-Score	mIoU	F-Score	mIoU	F-Score
SSD	61.2%	69.7%	59.1%	67.3%	59.8%	67.6%
FireNet	62.9%	72.2%	60.5%	70.5%	61.4%	71.1%
YOLO v3	71.4%	80.6%	68.2%	78.3%	69.5%	78.8%
Faster R-CNN	65.8%	72.7%	64.1%	70.6%	65.0%	71.5%
Faster R-CNN/Grassmannian VLAD encoding	74.4%	83.4%	69.9%	87.4%	73.8%	87.4%
U-Net	68.4%	74.4%	64.8%	71.3%	67.4%	71.9%
Proposed	78.2%	94.8%	70.4%	93.9%	77.1%	94.6%

mIoU: mean intersection over union Union, SSD: Single Shot multibox Detector, YOLO: You Only Look Once, R-CNN: Region-based Convolutional Neural Network, VLAD: Vector of Locally Aggregated Descriptors.

In addition to the method that combines Faster R-CNN and Grassmannian VLAD encoding, the proposed methodology is compared against: SSD [60], FireNet [47], YOLO v3 [61], Faster R-CNN [62] and U-Net [63] architectures. Towards the aim of fire detection through either flame or smoke detection, these methods achieve F-score rates of 67.6%, 71.1%, 78.8%, 71.5% and 71.9% and mIoU rates of 59.8%, 61.4%, 69.5%, 65% and 67.4%, respectively. Also, it is notable that fire detection through flame achieves better rates than through smoke detection. This is due to the fact that in the images of the created dataset, there are more objects that have similar characteristics to smoke. So, the number of false-negatives is higher in smoke detection. Furthermore, as depicted in Table 2, the approaches that take into account the dynamics of flame and smoke improve the fire detection rates.

Figures 7 and 8 illustrate qualitative results of the proposed method in images that contain synthetic fire and real fire events, respectively. Figure 7 provides examples of fire detection in different environments. In the case of a sub-urban environment (Figure 7a), the proposed systems accurately detect the smoke (Figure 7c). Furthermore, in the case of a single fire event in a forest (Figure 7d) that consists from both flame and smoke, the fire is accurately detected but only through flame detection. On the other hand, in the case of multi-fire events (Figure 7g), the system detects the fire events either through flame or smoke detection and erroneously detects some clouds as a smoke event. The results of Figure 8 show that in good visibility conditions, the proposed system detects the fire event in a suburban environment even in the early stages of the fire. Finally, it is worth mentioning that fire detection through either flame or smoke increases true-positive rates at the maximum level, but at the same time, increases false-positive rates as the rates of non-overlapping false-positive flame and smoke detected regions are combined.

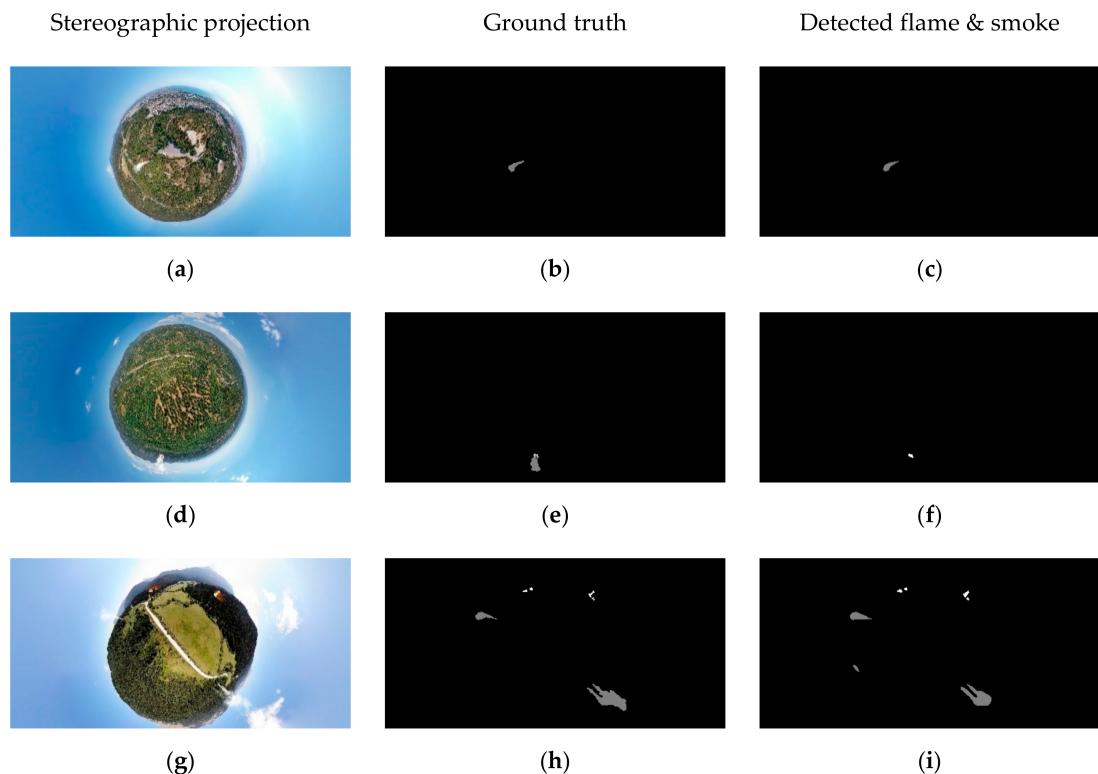


Figure 7. Results of the proposed methodology in 360-degree images containing synthetic fire events of the “Fire detection 360-degree dataset”. Equirectangular “360-FIRE-B” images containing flame (a–c), containing smoke (d–f) and containing both flame and smoke (g–i).

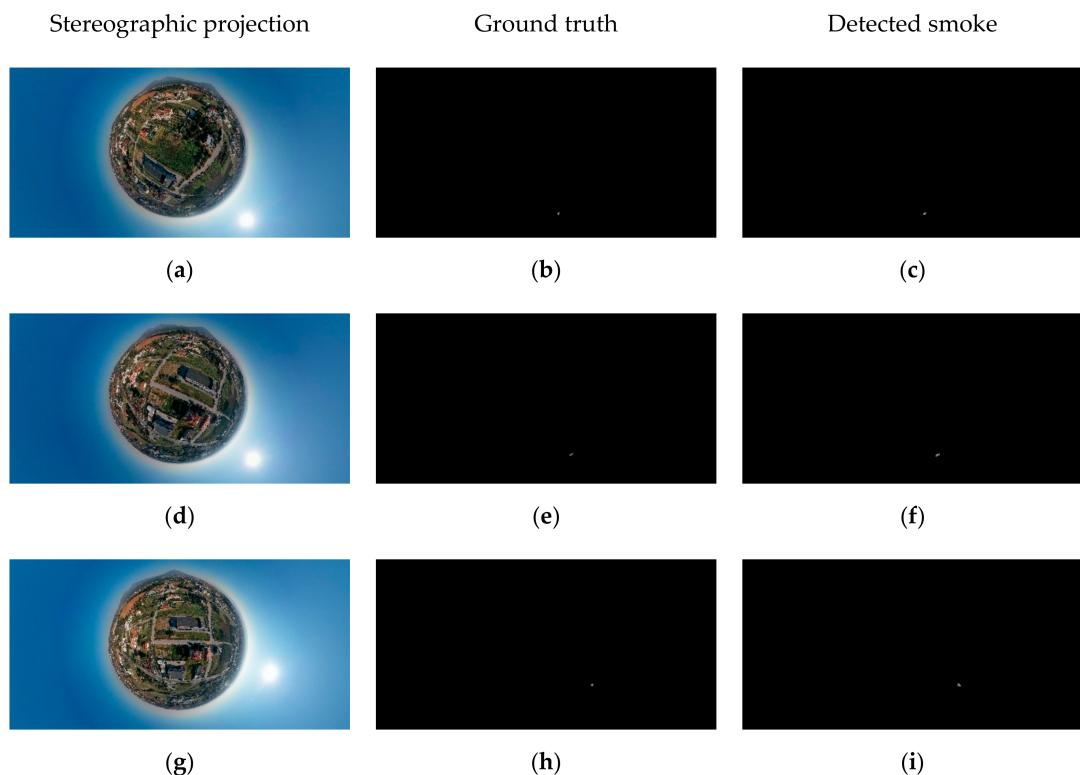


Figure 8. Results of the proposed methodology in 360-degree images containing real fire events of the “Fire detection 360-degree dataset”. Equirectangular “360-FIRE” images containing flame (a–c), containing smoke (d–f) and containing both flame and smoke (g–i).

4. Discussion

Remote sensors play an important role in vision-based surveillance systems. Thus, in combination with computer vision methods and the improvement of computation and storage ability, surveillance systems have been developed, achieving accurate object recognition and localization. The proposed novel remote sensing system for fire detection demonstrates that early fire detection can be accurately achieved using 360-degree cameras mounted on an UAV.

In this study, we used visible-range data captured by 360-degree optical cameras, aiming to propose a flexible and cost-effective remote sensing system for a fire detection system [64]. The use of 360-degree cameras and stereographic projections has the advantage that it significantly reduces the number of redundant sensors and overlapped information. A wide field of view is always advantageous in surveillance since it can collect richer information by reducing the number of blinding areas to zero. Compared to conventional sensors (optical, infrared and microwave), omnidirectional cameras can cover a 360-degree field of view with only a single camera reducing the number of required sensors. For example, if the field of view of a conventional camera is 90 degrees, then in order to capture the 360-degree or panoramic scene, at least 4 sensors will be required. It is worth mentioning that many UAVs take 24 single shots in order to reconstruct 360-degree images accurately [65]. In the past, rotating the cameras or increasing the number of available cameras were often used for enlarging the field viewing region. However, rotating cameras cannot record one 360-degree scene at the same time. In addition, using multiple sensors, the location of cameras needs to be delicately designed for fully exploring their field of view and reducing the overlapping viewing areas. Subsequently, in contrast to satellite-based systems, the proposed system offers better data availability but not global surveillance.

Furthermore, another advantage of the 360-degree cameras and the use of stereographic projections is the fact that the examined scene is always located in the center of the image. This makes the detection of the horizon line and of regions, that tend to cause plenty of false alarms due to the sunlight, clouds and other smoke-colored objects, easier. To this end, we proposed the adaptive post-validation scheme, with the aim of decreasing the high false alarm rates caused by natural objects which present similar characteristics to fire. This technique could easily be used in data retrieved from different remote sensing technologies and sensors in the case that the horizon line is known in advance.

There are, however, some limitations in the proposed system. UAVs are affected by weather conditions including wind or other extreme events. In addition, the flight time of UAVs is limited, so a network of UAVs is required in order for us to achieve continuous capture. Otherwise, a 360-degree camera can be placed at the top of a watchtower, achieving similar results with the proposed system and methodology. Moreover, the proposed system is not capable of wildfire detection at night. However, the proposed remote sensing system can be integrated with 360-degree data from other types of sensors, e.g., infrared or microwave, expanding the fire detection capabilities.

In the future, we aim to install an autonomous operation network of 360-degree sensors, performing periodic flights, for the surveillance of wide areas and to apply the proposed method for the detection of real fire events. Thus, we aim to extend the system, estimating the spread rate of fire and assisting in evacuation plans in case of hazardous events. With the frame sequences extracted from the videos and motion analysis techniques, we aim to calculate the spread rate and even predict the movement directions of fire, which will make great contributions to fire escape in real-world situations.

5. Conclusions

Nowadays, omnidirectional and 360-degree cameras are widely available and have become more and more popular, hence they can be used for surveillance, enabling wider field of view and avoiding the need to use many sensors or install complicated systems. Thus, in this paper, we proposed an intelligent system integrating an UAV and a 360-degree camera in order to achieve forest fire surveillance in a way that was never considered before. The proposed methodology combines deep convolutional neural networks and exploits dynamics of flame and smoke. From the experimental results, we conclude that the post-validation scheme and exploitation of fire dynamic texture significantly improves the

detection rates, reducing the false-positives. This, along with stereographic projections, enables us to discard false-positive regions that have similar characteristics with regions that are more likely to contain similar to false-positive features of each examined image.

Author Contributions: Conceptualization, P.B. and N.G.; methodology, P.B., T.S., K.D. and N.G.; software, P.B.; validation, P.B.; formal analysis, P.B.; data curation, P.B.; visualization, P.B.; writing—original draft preparation, P.B.; writing—review and editing, P.B., T.S., K.D. and N.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research is co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning» in the context of the project “Reinforcement of Postdoctoral Researchers—2nd Cycle” (MIS-5033021), implemented by the State Scholarships Foundation (IKY). Kosmas Dimitropoulos and Nikos Grammalidis have received funding from INTERREG V-A COOPERATION PROGRAMME Greece-Bulgaria 2014–2020 project “e-OUTLAND: Protecting biodiversity at NATURA 2000 sites and other protected areas from natural hazards through a certified framework for cross-border education, training and support of civil protection volunteers based on innovation and new technologies”.

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

References

1. HM Government. A Green Future: Our 25 Year Plan to Improve the Environment. Available online: <https://www.gov.uk/government/publications/25-year-environment-plan> (accessed on 24 October 2019).
2. Wieting, C.; Ebel, B.A.; Singha, K. Quantifying the effects of wildfire on changes in soil properties by surface burning of soils from the Boulder Creek Critical Zone Observatory. *J. Hydrol. Reg. Stud.* **2017**, *13*, 43–57. [[CrossRef](#)]
3. Allison, R.S.; Johnston, J.M.; Craig, G.; Jennings, S. Airborne optical and thermal remote sensing for wildfire detection and monitoring. *Sensors* **2016**, *16*, 1310. [[CrossRef](#)]
4. Govil, K.; Welch, M.L.; Ball, J.T.; Pennypacker, C.R. Preliminary Results from a Wildfire Detection System Using Deep Learning on Remote Camera Images. *Remote Sens.* **2020**, *12*, 166. [[CrossRef](#)]
5. NASA EARTHDATA. Remote Sensors. Available online: <https://earthdata.nasa.gov/learn/remote-sensors> (accessed on 10 September 2020).
6. Szpakowski, D.M.; Jensen, J.L. A review of the applications of remote sensing in fire ecology. *Remote Sens.* **2019**, *11*, 2638. [[CrossRef](#)]
7. Veraverbeke, S.; Dennison, P.; Gitas, I.; Hulley, G.; Kalashnikova, O.; Katagis, T.; Kuai, L.; Meng, R.; Roberts, D.; Stavros, N. Hyperspectral remote sensing of fire: State-of-the-art and future perspectives. *Remote Sens. Environ.* **2018**, *216*, 105–121. [[CrossRef](#)]
8. Yuan, C.; Liu, Z.; Zhang, Y. Aerial images-based forest fire detection for firefighting using optical remote sensing techniques and unmanned aerial vehicles. *J. Intell. Robot. Syst.* **2017**, *88*, 635–654. [[CrossRef](#)]
9. Hendel, I.G.; Ross, G.M. Efficacy of Remote Sensing in Early Forest Fire Detection: A Thermal Sensor Comparison. *Can. J. Remote Sens.* **2020**, *1*–15. [[CrossRef](#)]
10. Töreyin, B.U.; Dedeoğlu, Y.; Güdükbay, U.; Cetin, A.E. Computer vision based method for real-time fire and flame detection. *Pattern Recognit. Lett.* **2006**, *27*, 49–58. [[CrossRef](#)]
11. Dimitropoulos, K.; Tsalakanidou, F.; Grammalidis, N. Flame detection for video-based early fire warning systems and 3D visualization of fire propagation. In Proceedings of the 13th IASTED International Conference on Computer Graphics and Imaging, Crete, Greece, 18–20 June 2012.
12. Grammalidis, N.; Cetin, E.; Dimitropoulos, K.; Tsakalanidou, F.; Kose, K.; Gunay, O.; Gouverneur, B.; Torri, D.; Kuruoglu, E.; Tozzi, S.; et al. A Multi-Sensor Network for the Protection of Cultural Heritage. In Proceedings of the 19th European Signal Processing Conference, Barcelona, Spain, 29 August–2 September 2011.
13. Barmpoutis, P.; Dimitropoulos, K.; Grammalidis, N. Real time video fire detection using spatio-temporal consistency energy. In Proceedings of the 10th IEEE International Conference on Advanced Video and Signal Based Surveillance, Krakow, Poland, 27–30 August 2013; pp. 365–370.
14. Dimitropoulos, K.; Barmpoutis, P.; Grammalidis, N. Spatio-temporal flame modeling and dynamic texture analysis for automatic video-based fire detection. *IEEE Trans. Circuits Syst. Video Technol.* **2015**, *25*, 339–351. [[CrossRef](#)]

15. Lloret, J.; Garcia, M.; Bri, D.; Sendra, S. A wireless sensor network deployment for rural and forest fire detection and verification. *Sensors* **2009**, *9*, 8722–8747. [[CrossRef](#)]
16. Prema, C.E.; Vinsley, S.S.; Suresh, S. Efficient flame detection based on static and dynamic texture analysis in forest fire detection. *Fire Technol.* **2018**, *54*, 255–288. [[CrossRef](#)]
17. Gubin, N.A.; Zolotarev, N.S.; Poletaev, A.S.; Chensky, D.A.; Batzorig, Z.; Chensky, A.G. A microwave radiometer for detection of forest fire under conditions of insufficient visibility. *J. Phys. Conf. Ser.* **2019**, *1353*, 012092. [[CrossRef](#)]
18. Varotsos, C.A.; Krapivin, V.F.; Mkrtchyan, F.A. A New Passive Microwave Tool for Operational Forest Fires Detection: A Case Study of Siberia in 2019. *Remote Sens.* **2020**, *12*, 835. [[CrossRef](#)]
19. Koltunov, A.; Ustin, S.L.; Quayle, B.; Schwind, B.; Ambrosia, V.G.; Li, W. The development and first validation of the GOES Early Fire Detection (GOES-EFD) algorithm. *Remote Sens. Environ.* **2016**, *184*, 436–453. [[CrossRef](#)]
20. Vani, K. Deep Learning Based Forest Fire Classification and Detection in Satellite Images. In Proceedings of the 2019 11th International Conference on Advanced Computing (ICoAC), Chennai, India, 18–20 December 2019; pp. 61–65.
21. Jang, E.; Kang, Y.; Im, J.; Lee, D.W.; Yoon, J.; Kim, S.K. Detection and monitoring of forest fires using Himawari-8 geostationary satellite data in South Korea. *Remote Sens.* **2019**, *11*, 271. [[CrossRef](#)]
22. Polivka, T.N.; Wang, J.; Ellison, L.T.; Hyer, E.J.; Ichoku, C.M. Improving nocturnal fire detection with the VIIRS day–night band. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 5503–5519. [[CrossRef](#)]
23. Sharma, J.; Granmo, O.C.; Goodwin, M.; Fidje, J.T. Deep convolutional neural networks for fire detection in images. In Proceedings of the International Conference on Engineering Applications of Neural Networks, Athens, Greece, 25–27 August 2017; Springer: Berlin/Heidelberg, Germany, 2017; pp. 183–193.
24. Frizzi, S.; Kaabi, R.; Bouchouicha, M.; Ginoux, J.M.; Moreau, E.; Fnaiech, F. Convolutional neural network for video fire and smoke detection. In Proceedings of the IECON 2016–42nd Annual Conference of the IEEE Industrial Electronics Society, Florence, Italy, 24–27 October 2016; pp. 877–882.
25. Muhammad, K.; Ahmad, J.; Lv, Z.; Bellavista, P.; Yang, P.; Baik, S.W. Efficient deep CNN-based fire detection and localization in video surveillance applications. *IEEE Trans. Syst. Man Cybern. Syst.* **2018**, *49*, 1419–1434. [[CrossRef](#)]
26. Barmpoutis, P.; Dimitropoulos, K.; Kaza, K.; Grammalidis, N. Fire Detection from Images Using Faster R-CNN and Multidimensional Texture Analysis. In Proceedings of the ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing, Brighton, UK, 12–17 May 2019; pp. 8301–8305.
27. Chen, Y.; Zhang, Y.; Xin, J.; Yi, Y.; Liu, D.; Liu, H. A UAV-based Forest Fire Detection Algorithm Using Convolutional Neural Network. In Proceedings of the IEEE 37th Chinese Control Conference, Wuhan, China, 25–27 July 2018; pp. 10305–10310.
28. Zhao, Y.; Ma, J.; Li, X.; Zhang, J. Saliency detection and deep learning-based wildfire identification in UAV imagery. *Sensors* **2018**, *18*, 712. [[CrossRef](#)]
29. Dimitropoulos, K.; Barmpoutis, P.; Grammalidis, N. Higher order linear dynamical systems for smoke detection in video surveillance applications. *IEEE Trans. Circuits Syst. Video Technol.* **2016**, *27*, 1143–1154. [[CrossRef](#)]
30. Filonenko, A.; Hernández, D.C.; Jo, K.H. Fast smoke detection for video surveillance using CUDA. *IEEE Trans. Ind. Inform.* **2017**, *14*, 725–733. [[CrossRef](#)]
31. Toreyin, B.U.; Dedeoglu, Y.; Cetin, A.E. Contour based smoke detection in video using wavelets. In Proceedings of the IEEE 14th European Signal Processing Conference, Florence, Italy, 4–8 September 2006; pp. 1–5.
32. Chunyu, Y.; Jun, F.; Jinjun, W.; Yongming, Z. Video fire smoke detection using motion and color features. *Fire Technol.* **2010**, *46*, 651–663. [[CrossRef](#)]
33. Barmpoutis, P.; Dimitropoulos, K.; Grammalidis, N. Smoke detection using spatio-temporal analysis, motion modeling and dynamic texture recognition. In Proceedings of the 22nd European Signal Processing Conference, Lisbon, Portugal, 1–5 September 2014; pp. 1078–1082.
34. Sudhakar, S.; Vijayakumar, V.; Kumar, C.S.; Priya, V.; Ravi, L.; Subramaniyaswamy, V. Unmanned Aerial Vehicle (UAV) based Forest Fire Detection and monitoring for reducing false alarms in forest-fires. *Comput. Commun.* **2020**, *149*, 1–16. [[CrossRef](#)]
35. Yuan, C.; Liu, Z.; Zhang, Y. Learning-based smoke detection for unmanned aerial vehicles applied to forest fire surveillance. *J. Intell. Robot. Syst.* **2019**, *93*, 337–349. [[CrossRef](#)]

36. Dai, M.; Gao, P.; Sha, M.; Tian, J. Smoke detection in infrared images based on superpixel segmentation. In Proceedings of the MIPPR 2019: Remote Sensing Image Processing, Geographic Information Systems, and Other Applications, Wuhan, China, 2–3 November 2019.
37. Wang, J.; Roudini, S.; Hyer, E.J.; Xu, X.; Zhou, M.; Garcia, L.C.; Reid, J.S.; Peterson, D.; da Silva, A.M. Detecting nighttime fire combustion phase by hybrid application of visible and infrared radiation from Suomi NPP VIIRS. *Remote Sens. Environ.* **2020**, *237*, 111466. [CrossRef]
38. Yin, Z.; Wan, B.; Yuan, F.; Xia, X.; Shi, J. A deep normalization and convolutional neural network for image smoke detection. *IEEE Access* **2017**, *5*, 18429–18438. [CrossRef]
39. Tao, C.; Zhang, J.; Wang, P. Smoke detection based on deep convolutional neural networks. In Proceedings of the IEEE International Conference on Industrial Informatics-Computing Technology, Intelligent Technology, Industrial Information Integration, Wuhan, China, 3–4 December 2016; pp. 150–153.
40. Luo, Y.; Zhao, L.; Liu, P.; Huang, D. Fire smoke detection algorithm based on motion characteristic and convolutional neural networks. *Multimed. Tools Appl.* **2018**, *77*, 15075–15092. [CrossRef]
41. Xu, G.; Zhang, Y.; Zhang, Q.; Lin, G.; Wang, J. Deep domain adaptation based video smoke detection using synthetic smoke images. *Fire Saf. J.* **2017**, *93*, 53–59. [CrossRef]
42. Zhang, Q.X.; Lin, G.H.; Zhang, Y.M.; Xu, G.; Wang, J.J. Wildland forest fire smoke detection based on faster R-CNN using synthetic smoke images. *Procedia Eng.* **2018**, *211*, 441–446. [CrossRef]
43. Mi, T.W.; Yang, M.T. Comparison of Tracking Techniques on 360-Degree Videos. *Appl. Sci.* **2019**, *9*, 3336. [CrossRef]
44. Barmpoutis, P.; Stathaki, T. A Novel Framework for Early Fire Detection Using Terrestrial and Aerial 360-Degree Images. In Proceedings of the International Conference on Advanced Concepts for Intelligent Vision Systems, Auckland, New Zealand, 10–14 February 2020; Springer: Cham, Switzerland, 2020; pp. 63–74.
45. Corsican Fire Database. Available online: <http://cfdb.univ-corse.fr/modules.php?name=Sections&sop=viewarticle&artid=137&menu=3> (accessed on 7 September 2019).
46. Toulouse, T.; Rossi, L.; Campana, A.; Celik, T.; Akhloufi, M.A. Computer vision for wildfire research: An evolving image dataset for processing and analysis. *Fire Saf. J.* **2017**, *92*, 188–194. [CrossRef]
47. Jadon, A.; Omaha, M.; Varshney, A.; Ansari, M.S.; Sharma, R. Firenet: A specialized lightweight fire & smoke detection model for real-time iot applications. *arXiv* **2019**, arXiv:1905.11922.
48. Center for Wildfire Research. Available online: http://wildfire.fesb.hr/index.php?option=com_content&view=article&id=62&Itemid=72 (accessed on 20 January 2020).
49. Center for Wildfire Research. Available online: http://wildfire.fesb.hr/index.php?option=com_content&view=article&id=66&Itemid=76 (accessed on 24 April 2020).
50. Open Wildfire Smoke Datasets. Available online: <https://github.com/aiformankind/wildfire-smoke-dataset> (accessed on 18 July 2020).
51. Smoke Dataset. Available online: <https://github.com/jiyongma/Smoke-Data> (accessed on 29 May 2020).
52. Barmpoutis, P. Fire detection-360-degree Dataset. *Zenodo* **2020**. [CrossRef]
53. Costantini, R.; Sbaiz, L.; Susstrunk, S. Higher order SVD analysis for dynamic texture synthesis. *IEEE Trans. Image Process.* **2007**, *17*, 42–52. [CrossRef]
54. Chang, C.H.; Hu, M.C.; Cheng, W.H.; Chuang, Y.Y. Rectangling stereographic projection for wide-angle image visualization. In Proceedings of the 2013 IEEE International Conference on Computer Vision, Sydney, Australia, 1–8 December 2013; pp. 2824–2831.
55. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the 2018 European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 801–818.
56. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A.A. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017.
57. Arfken, G.B.; Weber, H.J. Mathematical methods for physicists. *Am. J. Phys.* **1999**, *67*, 165. [CrossRef]
58. Dimitropoulos, K.; Barmpoutis, P.; Kitsikidis, A.; Grammalidis, N. Classification of multidimensional time-evolving data using histograms of grassmannian points. *IEEE Trans. Circuits Syst. Video Technol.* **2016**, *28*, 892–905. [CrossRef]
59. Karcher, H. Riemannian center of mass and mollifier smoothing. *Commun. Pure Appl. Math.* **1977**, *30*, 509–541. [CrossRef]

60. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. SSD: Single shot multibox detector. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; Springer: Cham, Switzerland, 2016; pp. 21–37.
61. Redmon, J.; Farhadi, A. Yolov3: An incremental improvement. *arXiv* **2018**, arXiv:1804.02767.
62. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. In Proceedings of the Advances in Neural Information Processing Systems 28: 29th Annual Conference on Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; pp. 91–99.
63. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Springer: Cham, Switzerland, 2015; pp. 234–241.
64. Aslan, S.; Güdükbay, U.; Töreyin, B.U.; Çetin, A.E. Early wildfire smoke detection based on motion-based geometric image transformation and deep convolutional generative adversarial networks. In Proceedings of the ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019; pp. 8315–8319.
65. DJI. Mavic 2 Pro Specification. Available online: <https://www.dji.com/gr/mavic-air-2/specs> (accessed on 12 September 2020).



© 2020 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/>).