Research Article

Smoke detection from foggy environment based on color spaces

Mehmet Erdal Özbek [a],* iD *, Uğur Emre Yıldız [a],* iD

[a]Department of Electrical and Electronics Engineering, İzmir Kâtip Çelebi University, İzmir, 35620, Turkey

ARTICLE INFO	ABSTRACT
	Detection of smoke from videos captured by surveillance cameras in outdoor environments is one of the useful outcome of Internet of Things (IoT) applications. The potential benefit increases when deep learning (DL) architectures are involved. However, an inherent difficulty is to detect smoke while natural events like fog exists. The effectiveness of color spaces in detection performance has not yet fully evaluated in those architectures. Moreover, the energy and memory requirements of DL architectures may not be applicable for handling IoT implementation demands. Therefore, in this work, a DL architecture with a suitable color space model, applicable for IoT implementations is proposed to detect smoke from videos in foggy environment. By collecting several videos including smoke samples, the performance comparison of popular and the state-of-the-art DL architectures denoted the outperforming result according to both accuracy and memory usage.

1. Introduction

Detection of fire is one of the major part of early warning systems for the safety of environment and people. At the initial instant, fire emits a visible smoke which may come with or without a flame. Today, electronic smoke detector equipment is already in common use for indoor environments while cameras are becoming a major part of the detection systems. In the case of outdoor or wildlife, early detection systems make use of surveillance systems with cameras including automated and intelligent capabilities. Those systems are now being considered in Internet of Things (IoT) framework which requires energy-efficient approaches when interconnected with other larger systems, for example in a smart city concept [1].

During the last two decades, studies on computer vision for fire detection systems have been increased [2, 3]. Since the smoke is visible before the flame, a majority of the works have also included smoke detection approaches especially for the detection of wildfire [4-7]. Those detection systems depend on extracting reliable and effective features reflecting texture, shape, color, movement, energy, and

frequency [8] that will help to deal with the real-world conditions such as fog, rain, or snow. Particularly, color information has shown to be useful and easily applicable [9-15].

Recent studies have focused on the use of the deep learning (DL) techniques that does not necessarily require a handcrafted feature extraction step. Therefore, it is more convenient to build end-to-end systems directly using the image or video data without the need of extracting features from them. In this manner, models based on convolutional neural networks (CNNs) have been considered [16-18]. The well-known baseline models like VGG-16 [19], AlexNet [20], GoogLeNet [21], ResNet [22], DenseNet [23], and MobileNet [24] have been applied. The DL methods for object recognition such as YOu LOok only once (YOLO) based on detecting the interested regions on the images have been used in an embedded system [25]. Apart from those general models, several smoke detection studies made use of other specific CNN architectures. Among them, a Deep Normalization and Convolutional Neural Network (DNCNN) has been proposed for smoke detection which handles both feature extraction and smoke recognition at the

same time [26]. The Faster R-CNN [27] model has been used to detect smoke in forest fire with augmented data of synthetic images [28]. Stacking basic blocks as a deep multi-scale CNN (DMCNN) for smoke recognition has been proposed [29] as a lightweight model. Likewise, a combination of VGG-16 and Resnet50 network architectures has been fused as a deep network to improve feature expression ability while increasing the depth of the whole network [30]. Specialized network models like FireNet [31] is another example of a lightweight model suitable for mobile and embedded applications. Energy efficient network models for similar intentions have been also given [32-36]. Other recent examples of CNN models for smoke detection include temporal evolution or combinations of networks such as the two-stage training of a Deep Convolutional Generative Adversarial Neural Network (DCGAN) [37], dilated CNN [38], deep saliency network [39], and deep dual-channel CNN based solutions [40, 41].

On the other hand, IoT is becoming more and more associated with the digitalization of environments for ease of control and respond to events. The IoT technology helps to detect fire incidents in forests, or in other areas by measuring real-world information such as temperature, gas levels, humidity, wind direction and speed. Today, computer vision-based techniques replace the conventional fire detection by overcoming the shortcomings of sensor-based methods [14]. Therefore, for the case of image/video, some of the aforementioned studies have presented their work [25, 31-34] considering the minimization of resources to be applicable for IoT implementations.

However, most of those works do not consider the effect of color spaces in camera recordings. Based on this motivation, we propose an energy-efficient smoke detection architecture of a color space based on CNNs. The novelty and the difference of this work lies in incorporating the color space models with the DL architectures in order to determine the best performance in detection of smoke. This is achieved by modifying the DL architectures and using the artificially generated foggy images to determine smoke from a foggy environment. Another aim is to determine the best structure requiring lower resources for IoT applications. A large set of videos is collected for this purpose and evaluations are performed for validation.

In the next section, the color space models are reviewed and the video collections gathered to be used with the DL architectures are summarized. Section 3 presents the proposed structure including the data preparation and modification of the DL architectures. Section 4 displays the evaluations and performance comparison with the state-of-the-art results. Final section summarizes the results and concludes the paper.

2. Smoke Detection

Smoke detection from image or video features vary based on the properties of the image texture, the segmentation applied for certain shape information, the representation captured as color spaces, the consecutive changes because of the movement, and other fundamental signal level features based on energy or frequency [8]. A brief information of the color spaces used in this work will be given in the sequel. Then the deep learning architectures and the sources of videos used in this work will be listed.

2.1. Color spaces

In computer graphics, based on tristimulus representation theory, color spaces are simply representations of color in three dimensional linear spaces or intensity channels of the red, green, and blue colors known as RGB. However, there are different interpretations which the selected components might be transformed into different color models. Perceptual color models use hue and saturation referring to chromaticity and additionally brightness information. The most common is the HSV (hue, saturation, value) which is also referred to as HSI (hue, saturation, intensity) or as the HSB (hue, saturation, brightness). A convenient model to represent the brightness in videos is the YUV that uses luminance and chrominance components. The Y component is called as the luma and the remaining components are referred to as the chrominance and specifically blue-based chrominance and red-based chrominance as in the YCbCr color space. A uniform color space is obtained by transformation of the reference points and the lightness value in L*a*b* defined by the International Commission on Illumination (CIE), describing a color on the red-green chrominances (a*), and on the yellow-blue chrominances (b*) [42, 43].

A color in a color model is described by numbers indicating how much of color, brightness, or other components is included. In digital imaging done by computers, these component values are often in the range 0 to 255, for an 8-bit resolution. For each of the images taken from the sequence of the videos, the component values of each pixel in the image, represented by the values of RGB, can be expressed with other color spaces by suitable transformations. In the following, the transformation formulas converting from RGB to YUV, HSV, and L*a*b* are presented, respectively.

From RGB to YUV:

$$Y = 0.229R + 0.587G + 0.114B$$
$$U = 0.1687R + 0.3313G + 0.5B + 128 \quad (1)$$
$$V = 0.5R + 0.4187G + 0.813B + 128$$

From RGB to HSV:

$$V = \max(R, G, B)$$

$$H = \begin{cases} \dfrac{60(G - B)}{V - \min(R, G, B)} & if \ V = R \\ \dfrac{120 + (B - R)}{V - \min(R, G, B)} & if \ V = G \\ \dfrac{240 + (R - G)}{V - \min(R, G, B)} & if \ V = B \end{cases} \quad (2)$$

$$S = \begin{cases} \frac{V - \min(R,G,B)}{V} & if\ V \neq 0 \\ 0 & otherwise \end{cases}$$

From RGB to L*a*b*:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$L^* = 116 f\left(\frac{Y}{Y_n}\right) - 16 \tag{3}$$

$$a^* = 500 \left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right) \right]$$

$$b^* = 200 \left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right) \right]$$

where $X_n = 0.0950456$, $Y_n = 1$, $Z_n = 1.088754$ and

$$f(s) = \begin{cases} \frac{1}{3}\left(\frac{29}{6}\right)^2 s + \frac{16}{116} & if\ s \leq \left(\frac{6}{29}\right)^3 \\ s^{\frac{1}{3}} & otherwise \end{cases} \tag{4}$$

While most of the cameras work based on the RGB intensity values, by using these transformations it is straightforward to convert the image pixel information irrespective of the cameras.

2.2. Deep learning architectures

Recent studies in both machine learning and computer hardware have contributed to propose efficient methods for training deep neural networks. Instead of fully connected hidden layers, the CNN typically has alternating convolution and pooling layers. Following the record-breaking success of AlexNet at 2010 ImageNet Large-Scale Visual Recognition Challenge several research groups have achieved lower error rates with higher number of layers as in GoogLeNet, ResNet and DenseNet.

Furthermore, advanced modules combine different deep learning architectures aiming to increase the performances. In Inception, many mini-network modules are built, multiple convolution filters of different sizes are then concatenated. On the other hand, it is known that increasing the number of layers are prone to the problem of vanishing gradients. The ResNet structure offers a solution for this problem. By combining the Inception with the ResNet lower error rates have been obtained. Another solution came up by making short-cuts between the input and output layers through the transition layers as in the DenseNet architecture. Further achievements have been proposed in Xception when depth-wise and point-wise convolutions are involved instead of conventional convolutional layers. An important achievement is obtained by using the MobileNet architecture, where mobile models of inverted residual structure is built with shortcut connections between the thin bottleneck layers. This is resulted with the minimization of requirements while improving the performance.

2.3. Video sources

The increasing use of cameras for wildfire detection led to many video sources available to be used for forest fire/smoke detection. However, most of the studies presented their results with a limited number of available videos. Today, with the increasing number of available videos and DL architectures, the problem can be investigated with more number of image data and with recent modules. As deeper learning structures may obtain better accuracy performances while handling more data, in this work an important number of videos is collected. The list of the video sources used in this study is presented in Table 1.

Table 1. List of the collected video sources

Source	Address
Wildfire smoke detection [5, 6]	https://homes.di.unimi.it/genovese/wild/wildre.htm
Center for Wildfire Research [44]	http://wildre.fesb.hr
VisiFire [4]	http://signal.ee.bilkent.edu.tr/VisiFire
MIVIA Smoke Dataset [12]	https://mivia.unisa.it/datasets/video-analysisdatasets/smoke-detection-dataset
Firesense [45]	https://zenodo.org/record/836749#.XrFiavkzZEY
Pond5	https://www.pond5.com
Pronghorn Productions	http://www.hdnaturefootage.net

3. Proposed Structure for Smoke Detection

There are many DL architectures used for fire/smoke detection from videos as mentioned in the previous sections. However, the real world weather conditions may degrade the performance. In the case of fog, our proposal depends on the effectiveness of the color spaces that has not yet presented for some of the DL architectures. On the other hand, the performance of DL structures highly depends on the massive amount of data. As there is a lack of smoke images in a foggy environment, our proposal includes augmenting artificial smoke to the images. Therefore, the structure is defined not only to decide whether the image contains smoke or not, but extended to determine the four possible outputs as smoky, foggy smoky, normal, and foggy.

The artificially generated foggy images use the 8-bit RGB image $(I_{R,G,B})$ for every channel. The image is first brightened by adding the 100 brightness value to the pixel values of the image. The highest brightness value of each color channel of RGB is then scaled with the highest brightness value of brightened image.

$$I_{foggy(R,G,B)} = (I_{R,G,B} + 100) \times \frac{argmax\{I_{R,G,B}\}}{argmax\{I_{R,G,B} + 100\}} \tag{5}$$

An example of a real and its modified version with the addition of synthetically generated fog are presented in Figure 1. Note that the increase in brightness may add further difficulty in detection of smoke.

Figure 1. An example of a real [46] (left) and its synthetically generated foggy image (right)

Irrespective of the images being real or includes artificial fog, the RGB images are converted to other color spaces using the equations (1) to (4). Our experiments denoted that the HSV color space is effective in detection of smoke regions in the image. In order to reveal this effectiveness a comparison is given with an example image of RGB, YUV, L*a*b*, and HSV, as presented in Figure 2. Note that the images presented for the other color spaces are the corresponding representations in RGB color space. In brief, the proposed structure depends on the use of HSV color space.


Figure 2. The comparison of color spaces for an example image [14]. RGB (top left), YUV (top right),

L*a*b* (bottom left), and HSV (bottom right)

4. Experimental Results

We performed experiments using the images collected from the video sources given in Table 1. A total of 188 video files have been converted into the image sequences of a total of 72220 images. As those videos have varying duration and frame per second rates, a selection is performed based on a similar work [32]. The whole data is separated into three groups with 20% for training, 30% for validation, and the rest 50% for testing in order to compare with the existing studies [32, 33]. The number of images for each group is displayed in Figure 3.
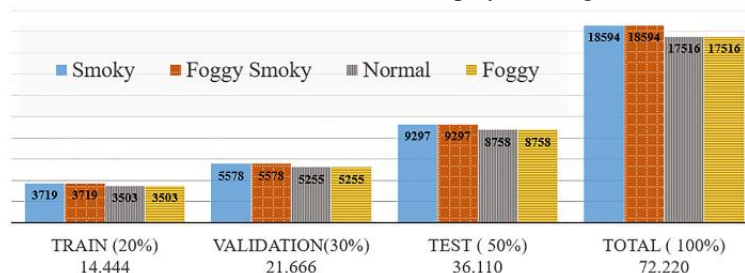
Figure 3. The number of images used for training, validation and testing

To maintain the structure of four outputs, the output softmax layer of the DL architectures is replaced with the proposed scheme accordingly as graphed in Figure 4. Then the performance of the proposed structure is compared with the selected and correspondingly modified architectures of VGG-16, VGG-19, InceptionV3, InceptionResNetV2, Xception, DenseNet169, DenseNet201, and MobileNetV2. In all of the architectures, the stochastic gradient descent (SGD) optimizer was used with a learning rate of 0.001. The batch size was selected as 16 with a number of epoch as 30.
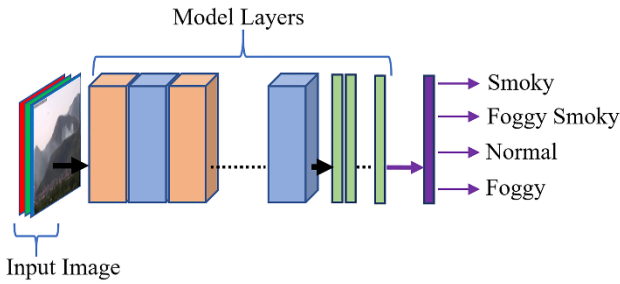

Figure 4. The general modified scheme of architectures

The performance of the systems is evaluated based on the accuracy of the test set in percentages and the memory space in megabytes (MB) required for storing the weight coefficients of the DL architectures. The average accuracy values for the corresponding architectures and color spaces is presented in Table 2. Notice that the RGB has the poorest performance with no highest score in any of the architectures, while HSV has the highest score in most of the DL architectures.

Table 2. Average accuracy values (%)

Architecture	Color spaces			
	RGB	YUV	L*a*b*	HSV
VGG-16	94.99	94.47	96.45	**97.50**
VGG-19	93.83	**96.36**	94.43	96.08
InceptionV3	95.46	97.50	95.20	**97.73**
InceptionResNetV2	95.15	95.79	94.99	**97.62**
Xception	96.44	96.72	**97.24**	97.19
DenseNet169	95.23	96.97	97.56	**98.14**
DenseNet201	95.27	94.18	**97.13**	95.21
MobileNetV2	96.72	98.19	97.45	**98.95**

The confusion matrix corresponding to the best performance of HSV with MobileNetV2 is shown in Table 3. The misclassifications of foggy and smoky is similar to the one of misclassifying a normal image with a smoky image.

Table 3. Confusion matrix of MobileNetV2 with HSV color space

	True			
Predicted	Smoky	Foggy smoky	Normal	Foggy
Smoky	9234	15	50	0
Foggy smoky	48	9236	0	22
Normal	9	0	8704	0
Foggy	6	46	4	8758

A comparison of memory usage for the DL architectures used in this study is given in Table 4. It is seen that the MobileNetV2 has considerably much lower memory space requirement.

Table 4. Memory usage

Architecture	Size (MB)
VGG-19	545.35
VGG-16	524.60
InceptionResNetV2	214.27
InceptionV3	86.03
Xception	81.90
DenseNet201	73.44
DenseNet169	50.96
MobileNetV2	9.29

The performance of the proposed method is compared with the state-of-the-art architectures using similar parameters, where a summary is presented in Table 5. Remark that the other studies not listed here utilizing any one of the color spaces make a classification between smoke and non-smoke.

Table 5. A comparison of studies based on color space, accuracy and memory size

Reference	Architecture	Color space	Accuracy (%)	Size (MB)
[32]	VGG-16	RGB	97.72	930
[33]	MobileNetV2	RGB	98.17	13.23
[46]	DenseNet169	YUV	97.80	50.9
Proposed	MobileNetV2	HSV	98.95	9.29

5. Conclusions

The detection of smoke from captured video images is an important step to prevent fire and its outcomes. While a binary detection scheme gives satisfactory results, real world situations necessitate detection systems to work in harsh conditions. One of the major condition in detection of smoke is the fog which share almost the same visible color information. Thus, our proposal is to benefit from color spaces in detection of smoke in foggy environments. This is accomplished with the DL architectures using a big number of image data collected and combined for this purpose. As the data is sparse for a fair comparison of smoke images in a foggy environment, artificially generated images are used.

Results demonstrated the efficiency of the proposed structure of the modified MobileNetV2 architecture with HSV color space. When compared with the well-known

DL architectures and similar works, it performed with a best accuracy score, while requiring lower memory space. This is foreseen to be an important issue especially in the IoT applications, as the DL architectures are becoming saturated in performance accuracy but they might still be compacted for further incoming lightweight implementations.

References

[1] A. Sharma, P. K. Singh, and Y. Kumar, "An integrated fire detection system using IoT and image processing technique for smart cities," *Sustainable Cities and Society*, vol. 61, 102332, 2020.

[2] A. E. Çetin, K. Dimitropoulos, B. Gouverneur, N. Grammalidis, O. Günay, Y. H. Habiboğlu, B. U. Töreyin, and S. Verstockt, "Video fire detection - Review," *Digital Signal Processing*, vol. 23, no. 6, pp. 1827-1843, 2013.

[3] P. Li and W. Zhao, "Image fire detection algorithms based on convolutional neural networks," *Case Studies in Thermal Engineering*, vol. 19, 100625, 2020.

[4] B. U. Töreyin, Y. Dedeoğlu, and A. E. Çetin, "Wavelet based real-time smoke detection in video," in *13th European Signal Processing Conference*, Antalya, Turkey, 2005, pp. 1-4.

[5] A. Genovese, R. D. Labati, V. Piuri, and F. Scotti, "Wildfire smoke detection using computational intelligence techniques," in *2011 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA) Proceedings*, Ottawa, ON, Canada, 2011, pp. 1-6.

[6] R. D. Labati, A. Genovese, V. Piuri, and F. Scotti, "Wildfire smoke detection using computational intelligence techniques enhanced with synthetic smoke plume generation," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 43, no. 4, pp. 1003-1012, July 2013.

[7] K. Zhou and X. Zhang, "Design of outdoor fire intelligent alarm system based on image recognition," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 34, no. 07, 2050018, 2020.

[8] X. Wu, Y. Cao, X. Lu, and H. Leung, "Patchwise dictionary learning for video forest fire smoke detection in wavelet domain," *Neural Computing and Applications*, vol. 33, pp. 7965-7977, 2021.

[9] T.-H. Chen, P.-H. Wu, and Y.-C. Chiou, "An early fire-detection method based on image processing," in *International Conference on Image Processing*, Singapore, 2004, pp. 1707-1710, vol. 3.

[10] T. Çelik and H. Demirel, "Fire detection in video sequences using a generic color model," *Fire Safety Journal*, vol. 44, no. 2, pp. 147-158, 2009.

[11] Y. Chunyu, F. Jun, W. Jinjun, and Z. Yongming, "Video fire smoke detection using motion and color features," *Fire Technology*, vol. 46, pp. 651-663, 2010.

[12] P. Foggia, A. Saggese, and M. Vento, "Real-time fire detection for video-surveillance applications using a combination of experts based on color, shape, and motion," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 9, pp. 1545-1556, Sept. 2015.

[13] C. E. Prema, S. S. Vinsley, and S. Suresh, "Multi feature analysis of smoke in YUV color space for early forest fire detection," *Fire Technology*, vol. 52, pp. 1319-1342, 2016.

[14] A. Khalil, S. U. Rahman, F. Alam, I. Ahmad, and I. Khalil, "Fire detection using multi color space and background modeling," *Fire Technology*, vol. 57, pp. 1221-1239, 2021.

[15] K.-M. Park and C.-O. Bae, "Smoke detection in ship engine rooms based on video images," *IET Image Processing*, vol. 14, no. 6, pp. 1141-1149, 2020.

[16] S. Frizzi, R. Kaabi, M. Bouchouicha, J. Ginoux, E. Moreau, and F. Fnaiech, "Convolutional neural network for video fire and smoke detection," in *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*, Florence, Italy, 2016, pp. 877-882.

[17] C. Tao, J. Zhang, and P. Wang, "Smoke detection based on deep convolutional neural networks," in *International Conference on Industrial Informatics - Computing Technology, Intelligent Technology, Industrial Information Integration (ICIICII)*, Wuhan, China, 2016, pp. 150-153.

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

[19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014. Online available: arXiv:1409.1556.

[20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems*, 2012, pp. 1097-1105.

[21] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1-9.

[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770-778.

[23] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017, pp. 2261-2269.

[24] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 2018, pp. 4510-4520.

[25] S. Saponara, A. Elhanashi, and A. Gagliardi, "Real-time video fire/smoke detection based on CNN in antifire surveillance systems," *Journal of Real-Time Image Processing*, vol.18, pp. 889-900, 2021.

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

[27] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *Advances in Neural Information Processing Systems*, 28, 2015.

[28] Q.-x. Zhang, G.-h. Lin, Y.-m. Zhang, G. Xu, and J.-j. Wang, "Wildland forest fire smoke detection based on faster R-CNN using synthetic smoke images," *Procedia Engineering*, vol. 211, pp. 441-446, 2018.

[29] F. Yuan, L. Zhang, B. Wan, X. Xia, and J. Shi, "Convolutional neural networks based on multi-scale additive merging layers for visual smoke recognition," *Machine Vision and Applications*, vol. 30, pp. 345-358, 2019.

[30] T. Liu, J. Cheng, X. Du, X. Luo, L. Zhang, B. Cheng, and Y. Wang, "Video smoke detection method based on change-cumulative image and fusion deep network," *Sensors*, vol. 19, no. 23, 5060, 2019.

[31] A. Jadon, M. Omama, A. Varshney, M. S. Ansari, and R. Sharma, "FireNet: a specialized lightweight fire & smoke detection model for real-time IoT applications," arXiv:190511922, 2019.

[32] S. Khan, K. Muhammad, S. Mumtaz, S. W. Baik, and V. H. C. de Albuquerque, "Energy-efficient deep CNN for smoke detection in foggy IoT environment," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 9237-9245, 2019.

[33] K. Muhammad, S. Khan, M. Elhoseny, S. H. Ahmed, and S. W. Baik, "Efficient fire detection for uncertain surveillance environment," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 5, pp. 3113-3122, May 2019.

[34] K. Muhammad, S. Khan, V. Palade, I. Mehmood, and V. H. C. de Albuquerque, "Edge intelligence-assisted smoke detection in foggy surveillance environments," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 1067-1075, February 2020.

[35] L. He, X. Gong, S. Zhang, L. Wang, F. Li, "Efficient attention based deep fusion CNN for smoke detection in fog environment," *Neurocomputing*, vol. 434, pp. 224-238, 2021.

[36] S. Khan, K. Muhammad, T. Hussain, J. der Ser, F. Cuzzolin, S. Bhattacharyya, Z. Akhtar, and V. H. C. de Albuquerque, "DeepSmoke: Deep learning model for smoke detection and segmentation in outdoor environments," *Expert Systems with Applications*, 115125, 2021.

[37] S. Aslan, U. Güdükbay B. U. Töreyin, and A. E. Çetin, "Early wildfire smoke detection based on motion-based geometric image transformation and deep convolutional generative adversarial networks," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, UK, 2019, pp. 8315-8319.

[38] T. Li, E. Zhao, J. Zhang, and C. Hu, "Detection of wildfire smoke images based on a densely dilated convolutional network," *Electronics*, vol. 8, no. 10, 1131, Oct. 2019.

[39] G. Xu, Y. Zhang, Q. Zhang, G. Lin, Z. Wang, Y. Jia, and J. Wang, "Video smoke detection based on deep saliency network," *Fire Safety Journal*, vol. 105, pp. 277-285, 2019.

[40] K. Gu, Z. Xia, J. Qiao, and W. Lin, "Deep dual-channel neural network for image-based smoke detection," *IEEE Transactions on Multimedia*, vol. 22, no. 2, pp. 311-323, 2020.

[41] F. Zhang, W. Qin, Y. Liu, Z. Xiao, J. Liu, Q. Wang, and K. Liu, "A dual-channel convolution neural network for image smoke detection," *Multimedia Tools and Applications*, vol. 79, pp. 34587-34603, 2020.

[42] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd edition, Prentice Hall, 2007.

[43] M. S. Nixon and A. S. Aguado, *Feature Extraction & Image Processing for Computer Vision*, 3rd edition, Academic Press, 2012.

[44] M. Bugaric, T. Jakovcevic, D. Stipanicev, "Adaptive estimation of visual smoke detection parameters based on spatial data and fire risk index," *Computer Vision and Image Understanding*, vol. 118, pp. 184-196, 2014.

[45] K. Dimitropoulos, P. Barmpoutis and N. Grammalidis, "Spatio-temporal flame modeling and dynamic texture analysis for automatic video-based fire detection," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 2, pp. 339-351, Feb. 2015.

[46] U. E. Yıldız and M. E. Özbek, "Deep learning based smoke detection for foggy environments," in *12th International Conference on Electrical and Electronics Engineering (ELECO)*, Bursa, Turkey, 2020, pp. 237-240.