Edge Intelligence-Assisted Smoke Detection in Foggy Surveillance Environments

Muhammad, K., Khan, S., Palade, V., Mehmood, I. & De Albuquerque, V. H. C.

Author post-print (accepted) deposited by Coventry University's Repository

Original citation & hyperlink:
Muhammad, K, Khan, S, Palade, V, Mehmood, I & De Albuquerque, VHC 2019, 'Edge Intelligence-Assisted Smoke Detection in Foggy Surveillance Environments' IEEE Transactions on Industrial Informatics, vol. (In-press), pp. (In-press).
https://dx.doi.org/10.1109/TII.2019.2915592

DOI 10.1109/TII.2019.2915592
ISSN 1551-3203
ESSN 1941-0050

Publisher: IEEE

© 2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

This document is the author's post-print version, incorporating any revisions agreed during the peer-review process. Some differences between the published version and this version may remain and you are advised to consult the published version if you wish to cite from it.

Edge Intelligence-Assisted Smoke Detection in Foggy Surveillance Environments

***Abstract*— Smoke detection in foggy surveillance environments is a challenging task and plays a key role in disaster management for industrial systems. The current smoke detection methods are applicable to only normal surveillance videos, providing unsatisfactory results for video streams captured from foggy environments, due to challenges related to clutter and unclear contents. In this paper, an energy-friendly edge intelligence-assisted smoke detection method is proposed using deep convolutional neural networks (CNN) for foggy surveillance environments. Our method uses a light-weight architecture, considering all necessary requirements regarding accuracy, running time, and deployment feasibility for smoke detection in industrial setting, compared to other complex and computationally expensive architectures including AlexNet, GoogleNet, and VGG. Experiments are conducted on available benchmark smoke detection datasets, and the obtained results show good performance of the proposed method over state-of-the-art for early smoke detection in foggy surveillance.**

***Index Terms*—Artificial Intelligence, CNN, Edge Intelligence, Smoke Detection, Foggy Surveillance Environment**

I. Introduction

THE recently deployed surveillance networks have rich processing capabilities, where video streams can be processed in nearly real-time to monitor ongoing activities through object tracking and detection [1], action and activity recognition, event detection, and scene understanding [2-4]. Among the events occurring in surveillance, fire disasters are comparatively dangerous, leading to both economic and social damage. Due to this reason, several fire detection systems are recently developed and significant research efforts are spent for further improvement [5, 6]. Both the literature and human observations show that smoke can be seen from far away distance, due to its faster movement in the upward direction compared to fire, and thus its early detection may help detect fire, which is helpful to disaster management systems. Despite these clues, AI-assisted detection of smoke is a difficult task, due to numerous challenges that restrict the performance of smoke detection methods [7].

Smoke detection methods are broadly classified into color-based, motion-based, and hybrid methods. For example, the methods in [8-10] use color features for smoke detection. In [8], color information is combined with motion using optical flow and back propagation neural networks for smoke detection and classification. In another work [9], color features are combined with the image's energy in order to perform smoke detection. The work in [10] uses color information for smoke detection by employing fuzzy C-mean and back propegation neural networks. In addition to color and motion, other properties of smoke, such as its shape and other spectral, spatial, and temporal characteristics are investigated for its detection and classification, as given in [11-13]. Other methods exploring different aspects of motion for smoke detection are presented in [14-16]. In addition to these approaches, several other methods explored texture features for smoke detection, as investigated in [17-19]. The previously mentioned smoke detection methods have several issues, such as limited accuracy, higher false alarms, and a lack of ability to detect smoke at a greater distance.

Recently, several intelligent methods were presented for improving the previous smoke detection methods. For instance, [20] represented video subsequences as histograms of high order dynamical system descriptors, and the classification accuracy was improved by combining spatio-temporal modeling with a multidimensional dynamic texture analysis of smoke via particle swarm optimization. However, the approach is computationally expensive, achieving limited detection rate. Furthermore, several deep CNN based smoke detection methods have been reported in the recent literature. For instance, Frizzi et al. [21] presented a nine-layer CNN architecture for fire and smoke detection in videos. Yin et al. [22] proposed deep normalization with CNN for smoke detection using 14 different layers for features extraction and classification. Another smoke detection method for surveillance networks is presented in [23], which uses color features with shape, and its performance is tested on both CPU and GPU on a CUDA platform. The most recent method is presented in [24] based on VGG-16, with a focus on uncertain surveillance videos.

Summarizing the aforementioned literature on smoke detection methods, it can be seen that these methods were mainly presented for regular surveillance scenes, and obtained unsatisfactory performance in foggy surveillance environments. Furthermore, certain methods were good, but at the cost of a huge running time, limiting its applicability for real-time video stream processing. In addition, considering the nature of disaster management, the accuracy of smoke detection needs further improvement, while a significant reduction is required in false alarms. These problems are addressed in this work by making the following major original contributions:

1. CNNs are extensively investigated for smoke detection, and an energy friendly CNN-based method is proposed for smoke detection. The light-weight architecture, excellent accuracy, and a minimum model size of our method increase its feasibility of deployment in smart cities, and especially in foggy surveillance networks in industrial setting, compared to state-of-the-art.
2. Compared to existing smoke detection methods that work well for only regular surveillance environments, we propose a framework for smoke detection in foggy surveillance scenes, which is inherently more challenging. The proposed method achieves better results for regular surveillance in general, and in foggy industrial surveillance video, in particular, as evident from experimental results.
3. Detailed experiments are conducted on existing benchmark datasets and another recently created smoke dataset by our team, in order to filter out the performance of all methods

Manuscript received: November 13, 2018; Accepted: XX, Published: XXX.

under consideration from different perspectives. Results on accuracy, false alarms, and other metrics suggest that the proposed method is an excellent candidate for smoke detection in foggy surveillance environments compared to state-of-the-art.

The rest of this work is structured as follows. The proposed system is discussed in detail with its main components in Section 2. The details about experimental setup, datasets, and evaluations are given in Section 3. This work is concluded in Section 4, with a list of future directions for further research.

II. The Proposed System

Recently, significant improvements have been reported in advancing visual sensors in terms of memory storage, processing, and intelligence. Through edge intelligence, several activities can be monitored by processing the video stream captured by cameras such as action and activity recognition, fire/smoke detection, and prioritization [25]. Currently, the majority of the hand-crafted and learned representations-based smoke detection approaches target surveillance videos of certain environment, which is relatively easy. Also, certain methods encounter a large number of false alarms with unsatisfactory accuracy. To cope up with these limitations, in this work, we investigate CNNs for smoke detection problem in uncertain surveillance videos, having fog, snow, and/or their combination. Next, considering the challenges of such videos and the requirements of disaster management, we propose an efficient edge intelligence-assisted CNN based system for detecting smoke, fog, and their variants. Unlike existing systems that either perform binary classification into smoke and non-smoke, our system classifies each video frame into one of the four classes: smoke, non-smoke, smoke with fog, and non-smoke with fog. The overall architecture of our system along with a sample image having four predictions is shown in Fig. **1**. The maximum probability represents the final label of the input image, as given in Fig. **2** for a set of sample images.

A. Architectural Details

This section describes the technical details of the architecture employed in the proposed system for smoke detection. We first studied and experimentally tested the famous CNN models, including AlexNet [26], GoogleNet [27], VGG-16 [28], and MobileNet (MNet) [29], for detecting smoke in video streams. The given models were tested with different sets of parameters, focusing on accuracy, false alarm rate, and other metrics. After extensive experiments, we found the MobileNet V2 as a suitable choice compared to other CNN popular models. The basic building block of the standard MobileNet V2 architecture is a bottleneck with residuals. In residuals bottleneck, the start and the end of a convolutional block is connected to each other with a skip connection. Using these states, the model has the opportunity of retrieving previous activations that are not updated in the convolutional block. The architecture of the MobileNet V2 initiates a convolutional layer followed by 19 residuals bottlenecks. After bottlenecks, there is a convolutional and pooling layer, followed further by another convolutional layer. The complete architectural details are given in Table **1**. MobileNet V2 was primarily trained on the ImageNet dataset, for classification into 1000 classes. We changed its last fully connected layer from 1000 classes to four classes: “smoke”, “non-smoke”, “smoke with fog”, and “non-smoke with fog”. We also applied the fine-tuning strategy to increase the smoke detection accuracy and minimize false alarm rate. After all necessary changes, our system was able to efficiently differentiate among the given four classes with significant accuracy, compared to state-of-the-art. The basic block of the employed architecture is given in Fig. **3**.

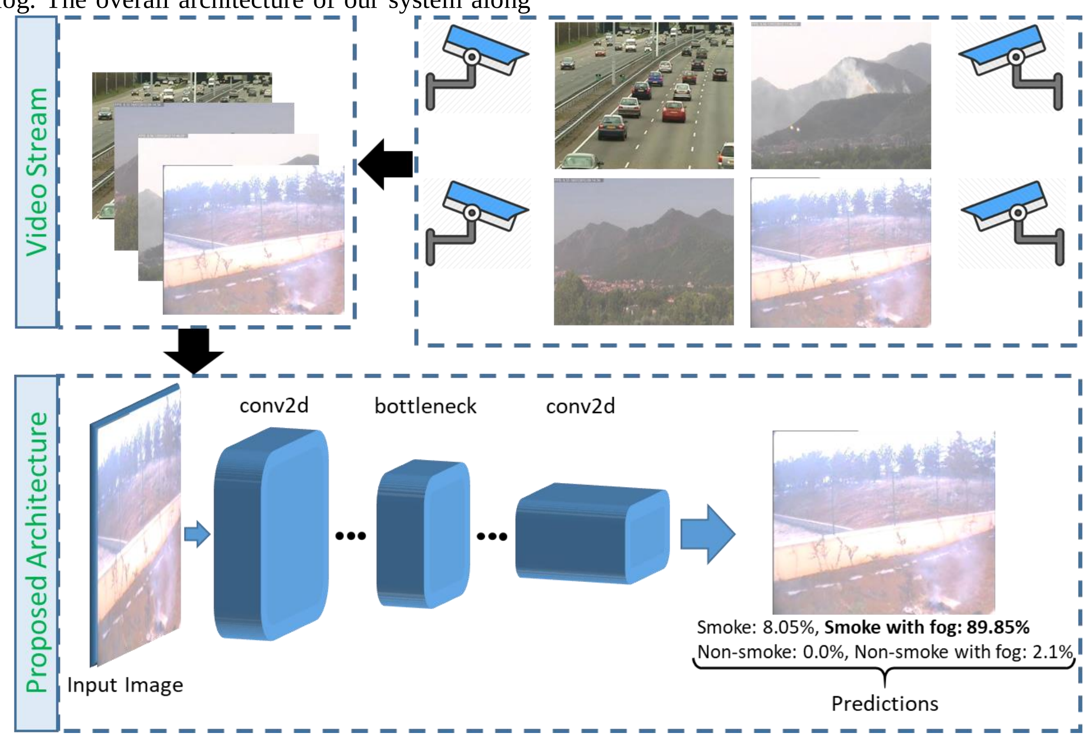

Fig. 1. Edge intelligence-assisted smoke detection system for foggy surveillance environments. The maximum prediction score, which is marked as bold, shows the final label of the input image.

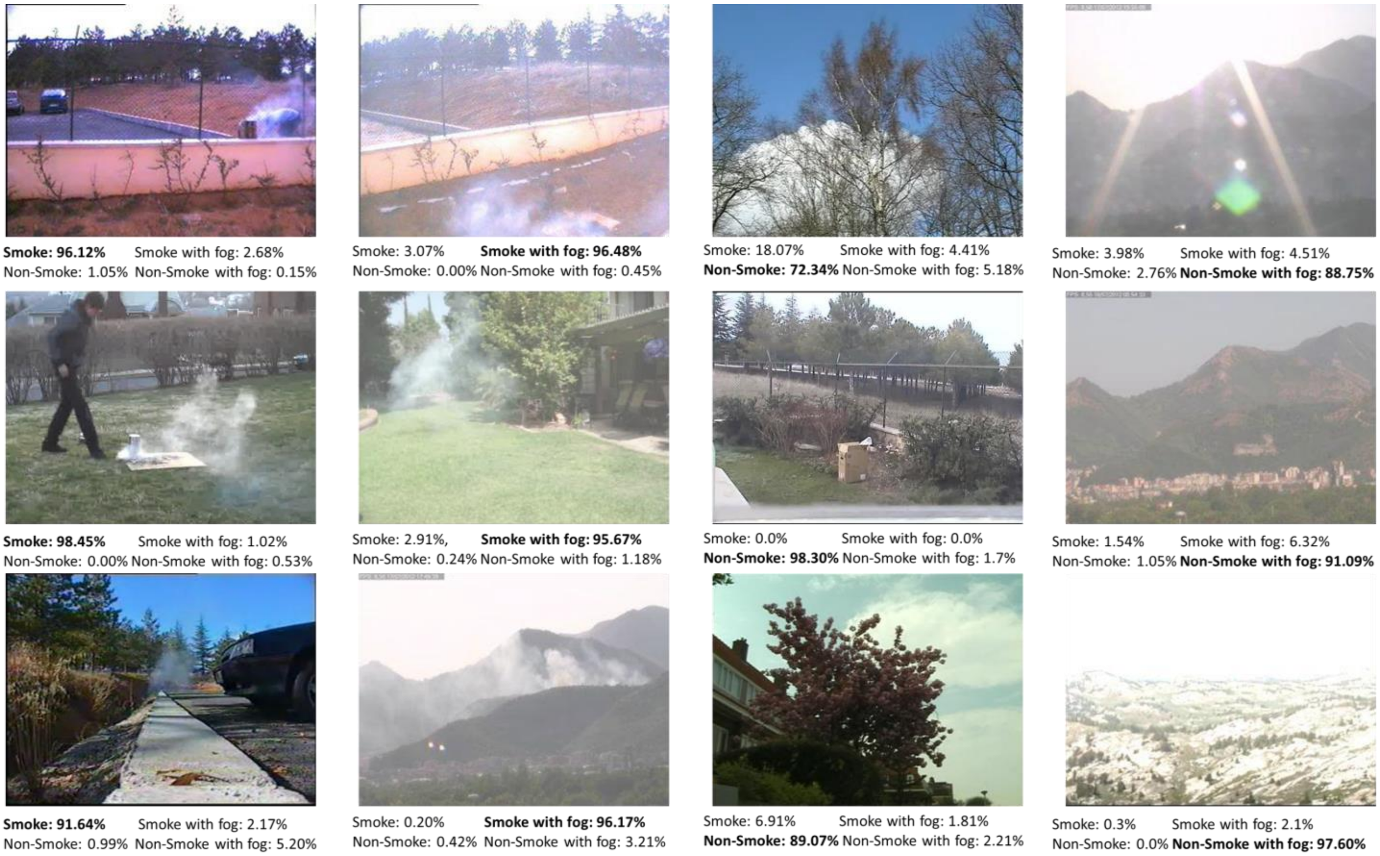

Fig. 2: Prediction probabilities of our proposed system for video frames captured in certain and uncertain environment. Predictions marked as bold refer to the final label of each image.

Table I
Architectural details of MobileNet V2.

Layer type	Layer	Number of repetition	Stride size
Convolution	conv2d 3×3	1	2
Bottleneck	bottleneck1	1	1
Bottleneck	bottleneck2	2	2
Bottleneck	bottleneck3	3	2
Bottleneck	bottleneck4	4	2
Bottleneck	bottleneck5	3	1
Bottleneck	bottleneck6	3	2
Bottleneck	bottleneck7	1	1
Convolution	conv2d 1×1	1	1
Pooling	avgpool1 7×7	1	-
Convolution	conv2d 1×1	1	-

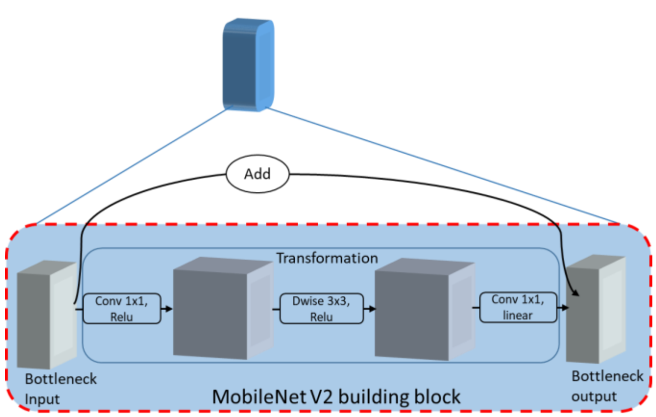

Fig. 3. Details of a single block of the employed architecture in the proposed system.

B. *Model Selection (MobileNet V1 vs MobileNet V2)*

This section describes the reasons for selecting the employed architecture compared to other CNNs, in general, and MobileNet V1, in particular. Several points are considered when dealing with video streams in uncertain surveillance environments, especially for disaster management and resource-constrained devices, such as delay, response time, accuracy, and false alarm rate. After an in-depth analysis of the investigated CNN models in the light of the aforementioned criteria, we chose MobileNet V2. The selected model is also highly feasible for devices with restricted memory and resources, such as Pi and FFGA. Other comparative statistics of the selected and related models are given in Table **II**. The given metrics clearly show the excellence of the employed architecture compared to other CNNs for our proposed system on large scale ImageNet dataset [30].

Table II
Statistics of MobileNet V2 against other models

Metrics	GNet	ANet	VGG-19	VGG-16	MNet V1	MNet V2
Parameters (millions)	7	60	144	138	4.24	3.47
Top-1 Accuracy (%)	69.8	57.1	70.5	70.5	70.9	71.8
Top-5 Accuracy (%)	89.3	80.2	91.2	91.0	89.9	91.0
Top-5 test error (%)	7.9	16.4	6.8	7.0	10.4	9.8

III. Experiments, Results, and Discussion

The detailed experiments are conducted in this section to evaluate and compare the performance of our method with other state-of-the-art methods. Firstly, we describe the details about the datasets used for the evaluation. Next, we compare our employed architecture with existing architectures using different evaluation strategies. Following this, the results of our method are compared with recent smoke detection methods. Finally, the running time and feasibility of our method are discussed in detail.

A. Datasets Description

The experiments mainly focus on a recently created dataset [24] and a set of seven videos. The new dataset is made of three existing datasets [31-33] and consists of four classes, i.e., “smoke”, “non-smoke”, “smoke with fog”, and “non-smoke with fog”. We divided the total number of 72,012 images into 20%, 30%, and 50% for training, validation, and testing, respectively. The overall statistics of these three sets are visualized in Fig. **4**, while the representative images of each class are presented in Fig. **5**. To extend the comparison analysis, we considered the seven publicly available videos [34, 35] as a second dataset for testing. It is worth notable that none of the image from these videos was used in the training process. The overall description of these videos with name, duration, and frame rate is given in Table **III**. The representative frames of these videos are visualized in Fig. **6** with labels from V1 to V7.

TABLE III
Description of the seven test videos from state-of-the-art

Video Number	Name	Duration (secs)	Frame rate	Description
V1	Cotton_rope_smoke_04.avi	115	25	Smoke originating from a cotton rope with a person standing nearby.
V2	Dry_leaf_smoke_02.avi	48	25	Smoke originating from dry leaves
V3	sBtFence2.avi	140	10	Persons moving in the scene with background similar to smoke. Smoke is at larger distance
V4	sMoky.avi	60	15	A video having contents with smoke color background
V5	sParkingLot.avi	69	25	Smoke produced in a parking lot. Objects’ movement and tree shaking is covered in the scene.
V6	sWasteBasket.avi	90	10	Smoke with a nearby red-color waste basket
V7	sWindow.avi	16	15	Smoke produced in a bucket and recorded from a window at larger distance

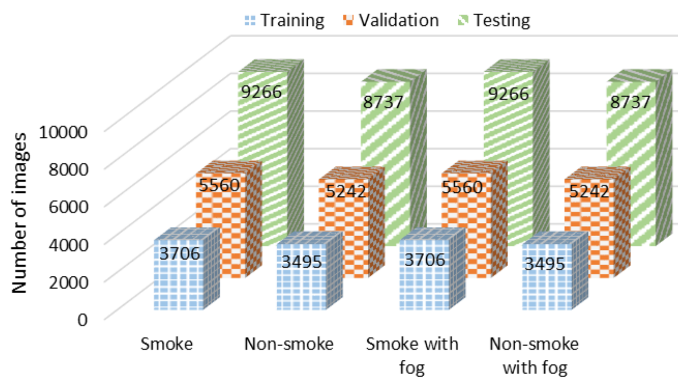

Fig. 4. Class-wise details of the employed dataset for our system.

Fig. 5. Sample frames from the dataset belonging to our target four classes: “smoke”, “non-smoke”, “smoke with fog”, and “non-smoke with fog”.

Fig. 6. Sample images from the seven testing smoke videos.

B. Comparison with CNN-based Smoke Detection Methods

We compare our proposed method with other state-of-the-art CNN models using the overall integrated dataset based on two evaluation strategies. In the first evaluation strategy, we use three metrics including false positive (FP), false negative (FN), and accuracy (A), as used by Foggia et al. [31]. FP is defined as a false alarm rate of the system, FN is the wrong prediction of positive class, and accuracy is described as the ratio of correctly predicted samples in the dataset, as given in Eq. **1**. The second evaluation strategy uses precision (P), recall (R), and F-measure (F). P is described as the ratio of correctly classified positive samples to the total predicted positive samples for a system, as shown in Eq. **2**. True positive (TP) rate is the correctly predicted positive samples, while TP + FP is the total positive samples. R is considered as the ratio of correctly classified positive samples to the total samples present in the class, as given in Eq. **3**. R refers to the sensitivity, or true positive rate of a system. In addition to P and R, F is calculated using the weighted average of P and R, as shown in Eq. **4**.

$$A = \frac{\mathrm{TP+TN}}{\mathrm{TP + FP+FN+TN}} \tag{1}$$

$$P = \frac{\mathrm{TP}}{\mathrm{TP + FP}} \tag{2}$$

$$R = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \tag{3}$$

$$F = 2 \times \left(\frac{\mathrm{P \times R}}{\mathrm{P+R}}\right) \tag{4}$$

Using the metrics of the first evaluation strategy, our model is compared with AlexNet (ANet), GoogleNet (GNet), and VGG-19 (VGGNet), and results are shown in Table **IV**.

Table IV
Comparison of our system with other CNN models on the test data with evaluation strategy 1

Model	FP (%)	FN (%)	A (%)
ANet [26]	3.39	4.16	95.87
GNet [27]	3.17	2.01	96.11
VGGNet [28]	2.30	2.01	97.72
Our method	**2.06**	**1.18**	**98.17**

It is evident from Table **IV** that ANet attained the worst accuracy, false-positive, and false-negative values, as compared to other models. GNet and VGGNet achieved a similar false-negative value, but in terms of accuracy and false-positive score VGGNet performed better than GNet. Our proposed system performed best compared to previous state-of-the-art methods by achieving the highest accuracy of 98.17%, minimum false-negative value of 2.06%, and minimum false alarm rate of 1.18%.

Table V
Comparison of our system with other CNN models on the test data with evaluation strategy 2

Model	P	R	F
ANet [26]	0.96	0.95	0.96
GNet [27]	0.96	0.96	0.96
VGGNet [28]	0.98	0.97	0.97
Our method	**0.98**	**0.97**	**0.98**

Furthermore, the second evaluation strategy is also employed to evaluate our proposed method in contrast to other state-of-the-art methods. To this end, results on the test set of the overall dataset are given in Table **V**. From these results, it can be observed that ANet and GNet have similar P and F values, while in terms of R, GNet performed better than ANet. VGGNet resulted in an R value similar to GNet but greater P and F values than ANet and GNet. Our proposed method outperformed all the three architectures using R and F values, and attained a similar P value to VGGNet. To sum up, our proposed method successfully dominated the state-of-the-art CNN architectures using both evaluation strategies, showing its superiority on smoke detection in foggy environments.

C. Comparison with other Smoke Detection Methods

This section describes the performance of our proposed system and other state-of-the-art smoke detection methods. The results are evaluated using seven test videos, as described in **Section III (A)**. The proposed method is compared with several image processing and learning based smoke detection methods, with comparative results given in Table **VI**. The evaluation metrics include accuracy, false alarm rate, and the processing time in frames per second (fps). Results show that the method described in [36] performed worst among all the methods under observation, due to its very low accuracy of 47% and high false alarm rate, but its fps is still better than ANet and GNet. The method [23] achieved best fps as compared to all other methods, but its accuracy and false alarm rate are still worse than the proposed and other models. The next three models, ANet, GNet, and VGGNet attained comparatively better accuracy and false alarm rate than the previous two methods, but their fps is still low. Our proposed system outperformed all the existing methods and models in terms of accuracy and false alarm rate and achieved the best combination of these evaluation metrics.

TABLE VI
Comparison with different smoke detection methods

Method	False alarm rate	fps	Accuracy (%)
Tian et al. [37]	4.1	2	-
Yuan et al. [36]	5.0	25	47.71
Yuan et al. [38]	4.57	2	81.3
Dimitropoulos et al. [20]	-	5.2	91.94
Yuan et al. [39]	3.92	-	-
Yin et al. [22]	2.44	30.73	-
Filonenko et al. [23]	4.29	**61**	84.85
Tian et al. [40]	-	3.25	84.47
ANet [26]	4.21	17	89.32
GNet [27]	3.57	23	92.24
VGGNet [28]	3.11	31.33	92.31
Our	**2.06**	39.78	**94.76**

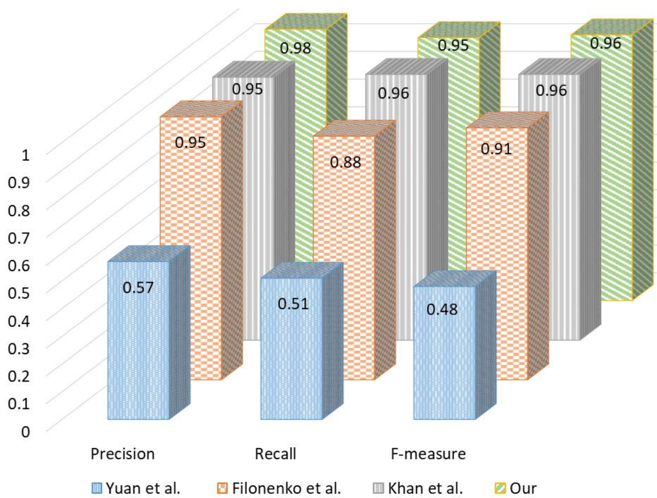

Fig. 7. Comparison of our proposed method with existing smoke detection methods including Yuan et al. [36], Filonenko et al. [23], and Khan et al. [24] using precision, recall, and F-measure. A single score represents the average for the seven test videos for each concerned metric.

Besides the above comparisons, we compared our method with the latest smoke detection methods in [36], [23], and [24] to show its effectiveness in normal environments. Seven test videos were used in this experiment and their average precision, recall, and F-measure are shown in **Fig. 7**. Results show that [36] achieved worst performance in terms of all three evaluation metrics. The average precision, recall, and F-measure values were 0.57, 0.51, and 0.48, respectively. The method [23] relatively performed better than [36], with an average precision 0.95, recall 0.88, and F-measure 0.91. Finally, our proposed method attained 0.98, 0.95, and 0.96 scores for precision, recall, and F-measure, respectively. Based on these results, our method outperformed the given benchmark recent methods, showing its superiority.

D. *Computational Complexity and Feasibility based Analysis*

This section provides the running time performance of our system and its feasibility for deployment in real-world scenarios. For this purpose, we performed experiments using a computer equipped with a GPU of NVidia GetForce TITAN X (Pascal). Further, our system have 12 GB onboard memory with a deep learning framework Caffe [41], running over a hardware of Intel Core i5 CPU with Ubuntu OS and 64 GB RAM. Using this setup, our system can process up to 40 frames per seconds, which is faster enough for real-time processing, because a normal camera can capture 25 to 30 frames per second. A detailed running-time based comparison of our proposed system with other state-of-the-art methods using seven test videos is given in Fig. **8**.

From these results, we can see that Filonenko et al. [23] achieved the best processing time, with an average of 20 ms per frame. In this case, the processing time for each video varies from minimum 16.30 ms to maximum 22.02 ms per frame. The limitation of this approach is its lower accuracy of 85% and higher false alarm rate of 4.29%, restricting its applicability for disaster management systems. The other competing method Yuan et al. [36] attained the worst processing time of 67.16 ms per frame on average. Khan et al. [24] performs better than Yuan et al. [36], but their running time is higher than our method's and of Filonenko et al. [23]. Our proposed method achieved an average processing time of 37.25 ms per frame, showing better performance than [36] and [24].

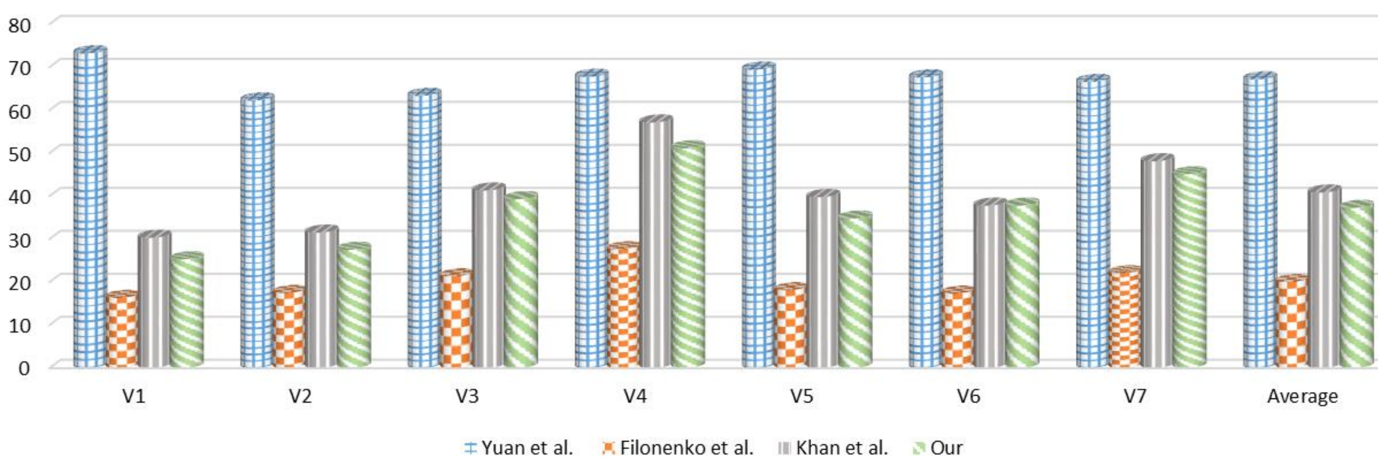

Fig. 8. Mean execution time (in milliseconds) for processing a frame by our system and existing methods (Yuan et al. [36], Filonenko et al. [23], and Khan et al. [24]) for the seven test videos.

Finally, we further compared our system with other state-of-the-art architectures, as given in Table **VII**. The goal of this comparison is to highlight the feasibility and deployment of our architecture over smart cameras and embedded devices in industrial surveillance. The parameters used for comparison are the MFLOPS/image and the size of the architecture in MB. From Table VII, it can be observed that ANet has better MFLOPS/image than GNet and VGGNet, however, its size is greater than that of GNet and our employed architecture. GNet is smaller in size than ANet and VGGNet, but its MFLOPS/image is higher than that of the ANet and the proposed architecture. VGGNet has a large number of parameters, yielding to highest MFLOPS/image and size compared to all other architectures. In contrast, our employed architecture consists of minimum MFLOPS/image and size, making it a more suitable choice for deployment over embedded devices, and allowing it to process industrial surveillance streams over edge in real-time.

TABLE VII
Models comparison in terms of mega floating point operations (MFLOPS)/image and size

Method Name	MFLOPS/image	Size (MB)
ANet [26]	720	219
GNet [27]	1500	39.66
VGGNet [28]	20000	930
Our method	**300**	**13.23**

IV. Conclusions and Future Work

With the available smart cameras, different abnormal events such as fire, flood, violence, etc., can be detected at early stages, and appropriate actions can be performed accordingly. Detecting these activities in regular surveillance videos is comparatively easy, however, it becomes significantly challenging when the environment is uncertain and the captured video stream is contaminated by fog, snow, or rain. For such scenarios, the current smoke detection systems result in limited performance, needing urgent attention. With this motivation in mind, we proposed an energy-friendly edge intelligence-assisted smoke detection method in this work, based on deep CNNs, in order to be used in foggy surveillance environments. Our method uses a light-weight architecture, considering all necessary requirements on accuracy, execution time, and deployment feasibility. Detailed experiments are conducted on benchmark smoke detection datasets, and the obtained results show encouraging performance of the proposed method for early smoke detection in foggy surveillance over state-of-the-art. In the light of the aforementioned characteristics of our system, we believe that it can efficiently monitor both certain and uncertain environments for smoke detection, which can be helpful in industrial scenarios, saving valuable resources from destruction.

In future, this work will be extended to monitor both fire [42] and smoke in video streams in both certain and uncertain environments [43], by further investigating edge as well as fog computing technologies [44]. The current smoke detection work will be further extended to include smoke segmentation/localization and contextual information extraction, in order to develop a collaborative intelligent scene analysis system. Furthermore, the current work can be merged with other abnormal event detection systems in smart cities, for smarter industrial surveillance [45].

References

[1] G. Liu, S. Liu, K. Muhammad, A. K. Sangaiah, and F. Doctor, "Object Tracking in Vary Lighting Conditions for Fog Based Intelligent Surveillance of Public Spaces," *IEEE Access*, vol. 6, pp. 29283-29296, 2018.

[2] K. Muhammad, R. Hamza, J. Ahmad, J. Lloret, H. Wang, and S. W. Baik, "Secure Surveillance Framework for IoT Systems Using Probabilistic Image Encryption," *IEEE Transactions on Industrial Informatics*, vol. 14, pp. 3679-3689, 2018.

[3] G. Han, L. Liu, J. Jiang, L. Shu, and G. Hancke, "Analysis of energy-efficient connected target coverage algorithms for industrial wireless sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 13, pp. 135-143, 2017.

[4] C. Zhu, L. Shu, V. C. Leung, S. Guo, Y. Zhang, and L. T. Yang, "Secure multimedia big data in trust-assisted sensor-cloud for smart city," *IEEE Communications Magazine*, vol. 55, pp. 24-30, 2017.

[5] K. Muhammad, J. Ahmad, Z. Lv, P. Bellavista, P. Yang, and S. W. Baik, "Efficient Deep CNN-Based Fire Detection and Localization in Video Surveillance Applications," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1-16, 2018.

[6] K. Muhammad, J. Ahmad, and S. W. Baik, "Early fire detection using convolutional neural networks during surveillance for effective disaster management," *Neurocomputing*, vol. 288, pp. 30-42, 2018.

[7] C. Brunner, T. Peynot, T. Vidal - Calleja, and J. Underwood, "Selective combination of visual and thermal imaging for resilient localization in adverse conditions: Day and night, smoke and fire," *Journal of Field Robotics*, vol. 30, pp. 641-666, 2013.

[8] 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.

[9] S. Calderara, P. Piccinini, and R. Cucchiara, "Vision based smoke detection system using image energy and color information," *Machine Vision and Applications*, vol. 22, pp. 705-719, 2011.

[10] T. K. T. Nguyen and J.-M. Kim, "Multistage optical smoke detection approach for smoke alarm systems," *Optical Engineering*, vol. 52, p. 057001, 2013.

[11] B. U. Toreyin, Y. Dedeoglu, and A. E. Cetin, "Contour based smoke detection in video using wavelets," in *European signal processing conference*, 2006.

[12] F. Yuan, "A double mapping framework for extraction of shape-invariant features based on multi-scale partitions with AdaBoost for video smoke detection," *Pattern Recognition*, vol. 45, pp. 4326-4336, 2012.

[13] C.-C. Ho, "Machine vision-based real-time early flame and smoke detection," *Measurement Science and Technology*, vol. 20, p. 045502, 2009.

[14] F. Yuan, "A fast accumulative motion orientation model based on integral image for video smoke detection," *Pattern Recognition Letters*, vol. 29, pp. 925-932, 2008.

[15] T. X. Tung and J.-M. Kim, "An effective four-stage smoke-detection algorithm using video images for early fire-alarm systems," *Fire Safety Journal*, vol. 46, pp. 276-282, 2011.

[16] C.-Y. Lee, C.-T. Lin, C.-T. Hong, and M.-T. Su, "Smoke detection using spatial and temporal analyses," *International Journal of Innovative Computing, Information and Control*, vol. 8, pp. 4749-4770, 2012.

[17] Y. Chunyu, Z. Yongming, F. Jun, and W. Jinjun, "Texture analysis of smoke for real-time fire detection," in *Computer Science and Engineering, 2009. WCSE'09. Second International Workshop on*, 2009, pp. 511-515.

[18] H. Maruta, A. Nakamura, and F. Kurokawa, "A new approach for smoke detection with texture analysis and support vector machine," in *Industrial Electronics (ISIE), 2010 IEEE International Symposium on*, 2010, pp. 1550-1555.

[19] F. Yuan, "Video-based smoke detection with histogram sequence of LBP and LBPV pyramids," *Fire safety journal*, vol. 46, pp. 132-139, 2011.

[20] K. Dimitropoulos, P. Barmpoutis, and N. Grammalidis, "Higher Order Linear Dynamical Systems for Smoke Detection in Video

Surveillance Applications," *IEEE Trans. Circuits Syst. Video Techn.*, vol. 27, pp. 1143-1154, 2017.

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

[22] 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.

[23] A. Filonenko, D. C. Hernández, and K.-H. Jo, "Fast Smoke Detection for Video Surveillance Using CUDA," *IEEE Transactions on Industrial Informatics,* vol. 14, pp. 725-733, 2018.

[24] S. Khan, K. Muhammad, S. Mumtaz, S. W. Baik, and V. H. C. d. Albuquerque, "Energy-Efficient Deep CNN for Smoke Detection in Foggy IoT Environment," *IEEE Internet of Things Journal,* pp. 1-1, 2019.

[25] K. Muhammad, J. Lloret, and S. W. Baik, "Intelligent and Energy-Efficient Data Prioritization in Green Smart Cities: Current Challenges and Future Directions," *IEEE Communications Magazine,* vol. 57, pp. 60-65, 2019.

[26] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.

[27] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov*, et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1-9.

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

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

[30] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma*, et al.*, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision,* vol. 115, pp. 211-252, 2015.

[31] 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, pp. 1545-1556, 2015.

[32] 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, pp. 339-351, 2015.

[33] 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, pp. 1003-1012, 2013.

[34] "http://staff.ustc.edu.cn/~yfn/vsd.html (cited on 12 August 2018) ".

[35] "http://signal.ee.bilkent.edu.tr/VisiFire/Demo/SmokeClips/ (cited on 12 August 2018)."

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

[37] H. Tian, W. Li, L. Wang, and P. Ogunbona, "Smoke detection in video: an image separation approach," *International journal of computer vision,* vol. 106, pp. 192-209, 2014.

[38] F. Yuan, J. Shi, X. Xia, Y. Fang, Z. Fang, and T. Mei, "High-order local ternary patterns with locality preserving projection for smoke detection and image classification," *Information Sciences,* vol. 372, pp. 225-240, 2016.

[39] F. Yuan, X. Xia, J. Shi, H. Li, and G. Li, "Non-linear dimensionality reduction and gaussian process based classification method for smoke detection," *IEEE Access,* vol. 5, pp. 6833-6841, 2017.

[40] H. Tian, W. Li, P. O. Ogunbona, and L. Wang, "Detection and Separation of Smoke From Single Image Frames," *IEEE Transactions on Image Processing,* vol. 27, pp. 1164-1177, 2018.

[41] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick*, et al.*, "Caffe: Convolutional architecture for fast feature embedding," in *Proceedings of the 22nd ACM international conference on Multimedia*, 2014, pp. 675-678.

[42] K. Muhammad, J. Ahmad, I. Mehmood, S. Rho, and S. W. Baik, "Convolutional Neural Networks Based Fire Detection in Surveillance Videos," *IEEE Access,* vol. 6, pp. 18174-18183, 2018.

[43] V. Petrov, S. Andreev, M. Gerla, and Y. Koucheryavy, "Breaking the Limits in Urban Video Monitoring: Massive Crowd Sourced Surveillance over Vehicles," *IEEE Wireless Communications,* 2018.

[44] L. Shu, G. P. Hancke, D.-J. Deng, C. Zhu, and M. Mukherjee, "Guest Editorial Fog Computing for Industrial Applications," *IEEE Transactions on Industrial Informatics,* vol. 14, pp. 4481-4486, 2018.

[45] H. Khelifi, S. Luo, B. Nour, A. Sellami, H. Moungla, S. H. Ahmed*, et al.*, "Bringing Deep Learning at The Edge of Information-Centric Internet of Things," *IEEE Communications Letters,* 2018.