This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2897594, IEEE Transactions on Industrial Informatics

Efficient Fire Detection for Uncertain Surveillance Environment

Khan Muhammad, *Member, IEEE*, Salman Khan, *Student Member, IEEE*, Mohamed Elhoseny, Syed Hassan Ahmed, *Senior Member, IEEE*, Sung Wook Baik, *Member, IEEE*

***Abstract*—Tactile Internet can combine multiple technologies by enabling intelligence via mobile edge computing and data transmission over a 5G network. Recently, several convolutional neural networks (CNN) based methods via edge intelligence are utilized for fire detection in certain environment with reasonable accuracy and running time. However, these methods fail to detect fire in uncertain IoT environment having smoke, fog, and snow. Furthermore, achieving good accuracy with reduced running time and model size is challenging for resource constrained devices. Therefore, in this paper, we propose an efficient CNN based system for fire detection in videos captured in uncertain surveillance scenarios. Our approach uses light-weight deep neural networks with no dense fully connected layers, making it computationally inexpensive. Experiments are conducted on benchmark fire datasets and the results reveal the better performance of our approach compared to state-of-the-art. Considering the accuracy, false alarms, size, and running time of our system, we believe that it is a suitable candidate for fire detection in uncertain IoT environment for mobile and embedded vision applications during surveillance.**

***Index Terms*—CNNs, Embedded Vision, 5G, Fire Detection, MobileNet, Disaster Management, Tactile Internet, Image Classification, Surveillance, Uncertain IoT Environment**

I. Introduction

THE connectivity of billions of smart devices have resulted in internet of things (IoT) and the maturity of installed sensors is ready for the emergence of Tactile Internet (TI), which have several useful applications for e-health, smarter surveillance, law enforcement, and disaster management [1-7]. In smart surveillance, edge intelligence plays an important role in security and disaster management. The instant reporting of unusual situations such as disaster in surveillance is very necessary for quick actions. The recent employed approach for instant transmission of such alarming information is 5G TI networks. Disaster management is mainly based on smoke/fire detection which can be performed using mobile edge computing. The main causes of fire are human mistakes or systems failure, which endangers human lives and properties. The statistics presented in [8] shows that wildfire disaster alone made an overall damage of 3.1 billion USD in 2015. Furthermore, in Europe 10,000 km^2 of area of vegetation is affected by fire disasters every year. To detect fire, researchers have presented both traditional and learned representation based fire detection methods. In literature, the traditional methods use either color or motion characteristics for fire detection. For instance, [9-16] used color features for fire detection by exploring different color models including HSI [12], YUV [13], YCbCr [14], RGB [15], and YUC [9]. The major issue with these methods is their high rate of false alarms. Several attempts have been made to solve this issue by combing the color information with motion and analyses of fire's shape and other characteristics [17-20]. However, maintaining a well-agreed trade-off between the accuracy, false alarms, and computational efficiency still remained a challenge. In addition, several methods from this domain fail to detect fire at a larger distance or small amount of fire.

To cope with these issues, recently convolutional neural networks (CNN) are explored for fire detection using edge intelligence. For instance, Frizzi et al., [21] presented a CNN based method for fire and smoke detection. Their work is based on a limited number of images and having no comparison with existing methods that could prove its performance. Sharma et al., [22] explored VGG16 and Resnet50 for fire detection. Their dataset is very small (651 images only) and the reported testing accuracy is less than 93%. Their work is compared with [21] with testing accuracy of 50%. Muhammad et al., [23] presented a CNN based early fire detection method for surveillance networks using two benchmark datasets. They also nominated a prioritization mechanism for cameras in a specific surveillance setup and explored cognitive radio networks-assisted channel selection approach for reliable data transmission. The main issue with this work is the huge size of model (238 MB), making its deployment restricted for resource constrained devices. In another work [24], a reasonable trade-off was maintained between the fire detection accuracy and false alarms rate, keeping the model size reasonable. A more efficient CNN based approach for both fire detection and localization was devised in [25] with model size of 3 MB, reasonable accuracy, and false alarm rate.

The aforementioned CNN based approaches are applicable to only certain environment with limited performance in uncertain surveillance environment. In addition, deploying huge-sized models on resource constrained devices is expensive and not recommended for surveillance networks. Furthermore, the fire detection accuracy and false alarm rate still need improvement, considering the critical nature of fire detection systems for disaster management. These issues are resolved in the current work with the following major contributions:

Manuscript received July 11, 2018; Accepted: XXX, Published: XXXX. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No.2016R1A2B4011712). This paper was recommended by Associate Editor XYZ. (*Corresponding author: Sung Wook Baik*)

Khan Muhammad is with the Department of Software, Sejong University, Seoul 143-747, South Korea. (e-mail: khan.muhammad@ieee.org)

Salman Khan and Sung Wook Baik are with Intelligent Media Laboratory, Digital Contents Research Institute, Sejong University, Seoul 143-747, Republic of Korea (Email: salmank@ieee.org, sbaik@sejong.ac.kr)

Mohamed Elhoseny is with Faculty of Computers and Information Mansoura University (Email: mohamed_elhoseny@mans.edu.eg)

Syed Hassan Ahmed is with Department of Computer Science Georgia Southern University Statesboro, GA, 90458, USA (Email: sh.ahmed@ieee.org)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2897594, IEEE Transactions on Industrial Informatics

1. We propose an efficient CNN based method for fire detection in videos captured in uncertain surveillance environment. Our method dominates the state-of-the-art in terms of accuracy and rate of false alarms.
2. Our method uses light-weight deep neural networks with no dense fully connected layers, making it computationally inexpensive. This favors our method for adaptation in surveillance networks with constrained resources in general and 5G TI-enabled surveillance in particular.
3. Our method results in an inference model of reasonable size (approximately 13 MB), which is easily deployable on mobile devices with embedded vision. We believe that our method is superior compared to state-of-the-art as verified from experiments and a suitable candidate for integration with disaster management systems.
4. The literature contains benchmark fire detection datasets for certain environment, however, there is no benchmark dataset specially created for uncertain surveillance environment. We created a dataset, consisting of synthetic fire images with fog and real fire foggy images. The dataset will be publically available for research community to mature fire detection algorithms for uncertain environment.

The rest of this paper is structured as follows. We disclose our method in Section 2 with its experimental validation in Section 3. Finally, we conclude this paper in Section 4 with key findings and several future directions for research community.

II. THE PROPOSED FRAMEWORK

The time-consuming efforts of features engineering makes fire detection a tedious job especially when the surveillance environment is uncertain with snow, fog, and smoke etc., or the fire is very small in size or at a long distance. In such situations, generally, the traditional fire detection systems produce a significant number of false alarms with limited fire detection accuracy. Recently, CNN based approaches are also explored for fire detection but their running time, size, and limited performance in several challenging situations (shadows, fire-like objects, uncertain scenes with smoke, snow, and fog etc.), make them infeasible for resource-constrained surveillance networks. Considering these challenges, we propose an efficient CNN based method for fire detection in videos captured in uncertain environment. To keep our method computationally inexpensive and effective for small-sized fire at a larger distance, we use light-weight deep neural networks with no dense fully connected layers. Our system is detailed in Fig. **1**.

A. CNN based Fire Detection

Literature shows that CNNs have achieved state-of-the-art performance for many real-world and challenging problems such as image classification, object detection and recognition [26], action and activity recognition [27, 28], segmentation, localization, image reconstruction, authentication [29], prioritization, indexing [30], and retrieval [31, 32]. The underlying factor behind this success is their hierarchical architecture consisting of convolution, pooling, and fully connected layers via which they automatically learn rich features from raw data. A convolution layer results in large number of feature maps from which high activations are selected by a pooling layer for dimensionality reduction and translation invariance. A fully connected layer learns high-level information needed for the target classification problem. In case of fire detection, a CNN architecture is usually changed such that the final fully connected layer has two classes i.e., fire and non-fire. The input fire data is provided to the intended CNN for training during which the weights of a large number of neurons are adjusted and learnt for classification into fire and non-fire.

B. Details of the Proposed Architecture for Fire Detection

The research community agrees that CNNs can automatically learn rich and discriminative features from raw data. However, much effort is needed to obtain the optimal setting, considering results through evaluation metrics, the amount of available data and its quality, and the problem under consideration. We explored different CNNs with different parameter settings for fire detection considering both certain and uncertain scenarios. After extensive experimentations, we found MobileNet version (V2) better than other models such as AlexNet [33], GoogleNet [34], and SqueezeNet [35]. Thus, we use a model with similar architecture to MobileNet [36] and modify it according to fire detection problem in uncertain surveillance environment. Similar to AlexNet, SqueezeNet, and GoogleNet, the baseline MobileNet is trained on ImageNet dataset for classification of objects into 1000 classes. Since MobileNet learns much rich features than other CNN models, thus we focused on re-using its learned features for accurate fire detection. To this end, we kept the number of neurons to two instead of 1000 in the final layer of our architecture, enabling classification into fire and non-fire. The architecture of MobileNet (V2) is modified by adding an expansion layer to the main building block. The modified block is given in Fig. **2**. To this end, we kept the number of neurons to two instead of 1000 in the final layer of our architecture, enabling classification into fire and non-fire. The architecture of MobileNet (V2) is modified by adding an expansion layer to the main building block. The modified block is given in Fig. **2**.

The expansion layer expands the number of channels in the input data before it is passed to the next layer of depthwise convolution. The amount of expansion can be controlled by the expansion factor, which is 6 by default. The second layer depthwise convolution filters the input while the projection layer makes the number of channels smaller. Each layer is also followed by a batch normalization with activation function "ReLU6". ReLU 6 is employed due to its robustness when used with low-precision computation. The projection layer is not followed by any activation function because its output is low-dimensional data and such activation function can affects the useful information. Overall, the employed architecture has 17 blocks similar to Fig. **2**, followed by a 1x1 convolution and the classification pipeline given in Fig. **1**. For getting inference on an input image, it is passed through the proposed architecture given in Fig. **1**, which outputs two probabilities. The highest probability indicates the final label of the input image as given in Fig. **3** for several sample images.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2897594, IEEE Transactions on Industrial Informatics

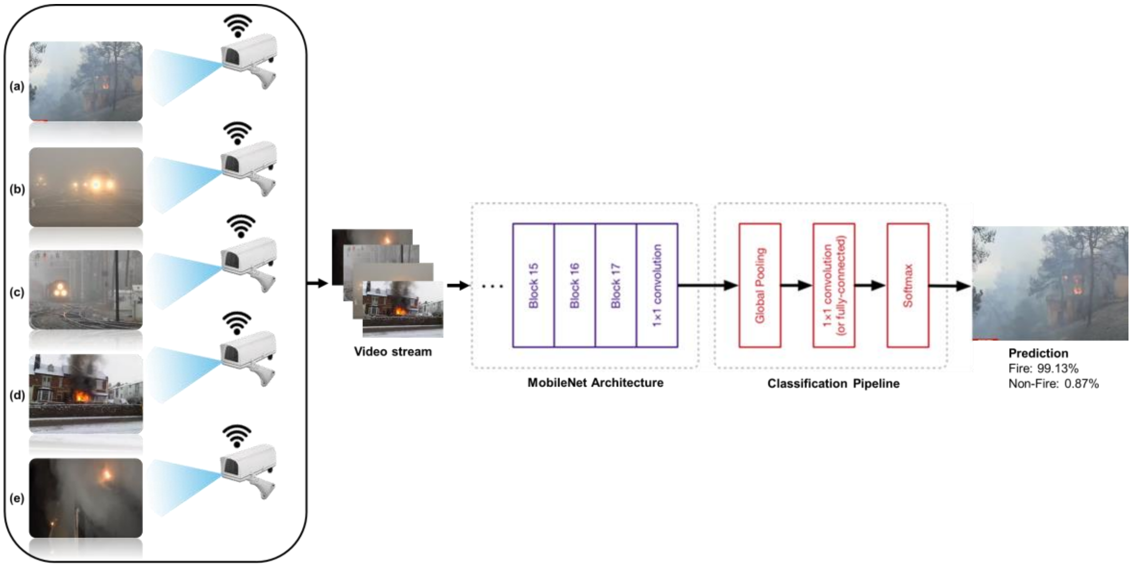


Fig. 1: Efficient deep CNN for fire detection in videos captured in uncertain environment. (a), (b), and (c): video stream from foggy surveillance. (d) video frames of surveillance from snowy scenes. (e) video stream with smoke and fire.

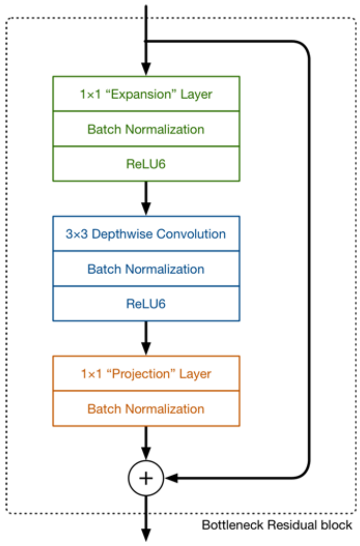


Fig. 2: Main building block of MobileNet V2 architecture

The expansion layer expands the number of channels in the input data before it is passed to the next layer of depthwise convolution. The amount of expansion can be controlled by the expansion factor, which is 6 by default. The second layer depthwise convolution filters the input while the projection layer makes the number of channels smaller. Each layer is also followed by a batch normalization with activation function "ReLU6". ReLU 6 is employed due to its robustness when used with low-precision computation. The projection layer is not followed by any activation function because its output is low-dimensional data and such activation function can affects the useful information. Overall, the employed architecture has 17 blocks similar to Fig. **2**, followed by a 1x1 convolution and the classification pipeline given in Fig. **1**. For getting inference on an input image, it is passed through the proposed architecture given in Fig. **1**, which outputs two probabilities. The highest probability indicates the final label of the input image as given in Fig. **3** for several sample images.

C. Motivations of using MobileNet (V2) for Fire Detection

Model selection is a critical step especially in resource constrained environment and for applications of critical nature such as disaster management where minor delay can result in huge loss in terms of humanity and economy. Compared to other CNN models, we use MobileNet due to its higher feasibility for memory and bandwidth-restricted hardware architectures such as FPGAs, smart sensors, and raspberry Pi and its suitability to 5G TI-enabled surveillance. The motivation of using MobileNet (V2) [36] compared to MobileNet version 1 (V1) [37] is its reduced size both in terms of number of computations and learned parameters with comparable accuracy. The statistics of both versions are given in Table **I**.

TABLE I
COMPARATIVE STATISTICS OF BOTH VERSIONS OF MOBILENET

Parameters	MobileNet V1_1.0_224	MobileNet V2_1.0_224
MACs (millions)	569	300
Parameters (millions)	4.24	3.47
Top-1 Accuracy (%)	70.9	71.8
Top-5 Accuracy (%)	89.9	91.0

Here "1.0" shows the version number of MobileNet V2 while "MAC" refers to multiply-accumulate operations, measuring the number of calculations required for getting inference on an image of size 224×224×3 pixels. Based on this metric, V2 is two times faster than V1. Besides this, memory access is also important and quite slower than computation on mobile devices. To this end, V2 has quite less number of parameters compared to V1. Finally, V2 performs better than V1 in terms of image classification accuracy on ImageNet dataset. These characteristics verify the choice of MobileNet V2 in our architecture.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2897594, IEEE Transactions on Industrial Informatics

Fig. 3: Classification predictions by our system for images captured from uncertain scenes with smoke, snow, and fog. (a) and (b): fire with smoke, (c) and (d): fire in night time, (e): fire with snow, (f) and (g): normal with fog and fire-colored regions, (h), (i), and (j): normal with fire-colored lighting at night.

III. Experimental Results and Discussion

This section provides details about the datasets, followed by experimental evaluation and comparison of our method with CNN and hand-engineered features based fire detection methods. Next, the robustness of our system is evaluated compared to existing methods with discussion on system parameters and its feasibility to uncertain surveillance environment. Finally, the discussion is concluded by highlighting the importance of the proposed framework for 5G TI-enabled fire detection system for surveillance in uncertain industrial environments. Similar to CNNFire [25], we call our method "EMN_Fire", [23] as "ANetFire", and [24] as "GNetFire" in the remaining of the paper for ease of interpretation.

A. Details of the Datasets

For experimentations, we have created a new dataset from two benchmark datasets: Dataset1 (DS1) [9] and Dataset2 (DS2) [38] with two classes "Fire" and "Non-fire". For the creation of new dataset, we selected random images from both classes in which fog is added synthetically. Smoke and snow images from Internet are included to the newly created dataset to cover the uncertain environment. The integrated dataset comprises of a total of 30,776 images. To train and test the system, we used our recent strategy [25] by using 20% data of the dataset for training and rest of 80% for testing. With this approach, our model is trained with 1844 fire images and 6188 non-fire images. The statistics of training and testing data is given in Table **II**. A few representative images from DS1 and DS2 with their remarks are given in Fig. **4**. For further details about the two benchmark datasets and the reasons for their usage in experiments, refer to [25].

TABLE II
Statistics of training and testing data for our system

	Dataset source	Total images	Fire images		Non-fire images	
			Certain	Uncertain	Certain	Uncertain
Training Data	DS1	8032	1604	240	4807	1381
Testing Data	DS1	22518	6000	1003	10515	5000
	DS2	226	100	19	90	17

B. Comparison with CNN based Fire Detection Methods

In this section, the performance of our system is compared with CNN based fire detection methods using the results, collected on both datasets of DS1 and DS2. Two different sets of evaluation metrics are employed to evaluate the performance of each method from all perspectives. The first set of metrics contain accuracy, false-negatives, and false-positives (also referred as false alarm rate) [25]. Using this set up, the proposed system is compared with the most recent work [25] and two other CNN based fire detection systems [23, 24]. The experimental results using both datasets are given in Table **III**.

TABLE III
Comparative results using DS1 and DS2 on evaluation set1

Dataset	Method Name	False-Positives (%)	False-Negatives (%)	Accuracy (%)
DS1	EMN_Fire	**0**	*0.14*	**95.86**
	CNNFire	*0.06*	1.24	94.61
	GNetFire	**0**	1.09	*93.66*
	ANetFire	6.78	**0.08**	94.27
DS2	EMN_Fire	**9.34**	6.72	**92.04**
	CNNFire	*18.69*	*2.52*	*89.82*
	GNetFire	24.29	6.72	84.96
	ANetFire	23.36	**1.68**	88.05

It can be seen that ANetFire achieved the best false-negatives (0.08), however, its false positives rate is high as well as its accuracy is 94.27%. GNetFire achieved similar false alarm rate to our proposed method, however, its accuracy is the worst using DS1. Our proposed system achieved the best combination of accuracy, false alarm rate, and false negatives using DS1, thus dominating other CNN models. The results using DS2 are also reported in Table **III**. DS2 is a small but challenging dataset. From results of DS2, we can see that GNetFire performed worst in terms of all metrics. CNNFire achieved better performance compared to ANetFire and GNetFire. The best combination is still achieved by our proposed system with minimum false alarm rate of 9.34% and highest accuracy of 92.04%. Thus, our method is better than existing approaches using the first set of metrics on both datasets.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2897594, IEEE Transactions on Industrial Informatics

Fig. 4: Representative images of fire and non-fire from both datasets with their descriptions.

TABLE IV
COMPARATIVE RESULTS USING DS1 AND DS2 ON EVALUATION SET2

Dataset	Method Name	Precision	Recall	F-Measure
DS1	EMN_Fire	**1**	**0.99**	**0.99**
	CNNFire	*0.99*	*0.98*	*0.98*
	GNetFire	**1**	0.97	*0.98*
	ANetFire	0.93	**0.99**	0.96
DS2	EMN_Fire	**0.90**	0.93	**0.92**
	CNNFire	*0.83*	*0.97*	*0.90*
	GNetFire	0.79	0.93	0.85
	ANetFire	0.80	**0.98**	0.88

To further investigate the performance of fire detection methods under consideration, we use another set of evaluation metrics including precision, recall, and F-measure. The complete details of these metrics can be found in [38, 39]. The incurred results using both DSI and DS2 through the second set of evaluation metrics are given in Table **IV**. Overall, the performance of ANetFire is worst on DS1, considering the precision and F-measure score. The performance of GNetFire and CNNFire [25] is almost same. The proposed system dominated other competing methods in terms of precision, recall, and F-measure score, showing its strength on DS1. Referring to DS2, GNetFire performed worst both in terms of precision and F-measure. ANetFire is better than GNetFire, however, it failed to beat CNNFire [25]. As shown, the proposed system successfully outperformed the competing fire detection systems, both in terms of precision and F-measure. The improvement is due to the deep but light-weighted neural networks used in the employed architecture for effectively learning discriminative features for fire detection.

C. Comparison with Hand-Crafted Features based Fire Detection Methods

This section investigates and analyzes the performance of the proposed system with respect to traditional fire detection methods and presents a comparison using both DS1 and DS2. The same two sets of evaluation metrics are used as mentioned

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2897594, IEEE Transactions on Industrial Informatics

in **Section III (B)**. Using the first set of evaluation metrics and DS1, the proposed system is compared with six representative methods that are based on color, motion, and shape characteristics of the fire. The comparative results are given in Table **V**. From the results, the worst method is [15] using RGB color model with highest false alarm rate of 41.18% and smallest accuracy of 74.20%. The method [14] has 0% false negatives, however, its accuracy is only 83%. The best method in the given existing methods in terms of false positives, is [16] with 5.88%. Similarly, in terms of false negatives, [14], [9], and [17] performed well. The reasonable combination in all three scores is achieved by [9] with accuracy 93.55% and false negatives 0%. However, the false alarm rate is really high and better accuracy is preferable, considering the critical nature of disaster management systems. Our proposed system has resolved these issues and has boosted the accuracy to 95.86% with 0% false alarms and negligible false negatives of 0.14%.

TABLE V
COMPARISON WITH DIFFERENT HAND-CRAFTED FEATURES BASED FIRE DETECTION METHODS ON DS1

Method Name	False Positives (%)	False Negatives (%)	Accuracy (%)
EMN_Fire	**0**	*0.14*	**95.86**
[9]	11.67	**0**	*93.55*
[17]	13.33	**0**	92.86
[16]	*5.88*	14.29	90.32
[15] (RGB)	41.18	7.14	74.20
[15] (YUV)	17.65	7.14	87.10
[14]	29.41	**0**	83.87
[12]	11.76	14.29	87.10

For further investigation, we compared our system using DS2 with the second set of evaluation metrics. The results are shown in Table **VI**. For readers' information, it is worth notable that DS2 is not used in training process of the proposed system, CNNFire, GNetFire, and ANetFire. From the results, it can be seen that the worst method is [12] with an F-measure score of 0.25 from which [40] is better. [14] and [41] have similar results. The recent method BoWFire [38] achieved better performance compared to other existing methods. Interestingly, our proposed system outperformed all existing methods in terms of all three metrics using DS2, showing its effectiveness.

TABLE VI
COMPARISON WITH HAND-CRAFTED FEATURES BASED FIRE DETECTION METHODS ON DS2

Method Name	Precision	Recall	F-Measure
EMN_Fire	**0.90**	**0.93**	**0.92**
BoWFire [38]	0.51	*0.65*	*0.57*
[41]	0.63	0.45	0.52
[40]	0.39	0.22	0.28
[14]	0.55	0.54	0.54
[12]	*0.75*	0.15	0.25

D. Robustness Analysis

For uncertain environment, it is important that the fire detection system is robust against well-known attacks. In this section, we have evaluated the robustness of our system against noise and fire blockage attack and have compared its results with state-of-the-art as shown in Fig. **5**. It can be noted that the proposed method provides best result in majority of the cases while second best result in some cases, reflecting its superiority for fire detection in uncertain environments with different weather conditions.

E. System Feasibility Analysis for Uncertain Environment

Besides simulation, it is important to investigate the feasibility of a system for deployment in real-world. This section is aimed at providing similar details about our system for deployment in uncertain 5G TI-enabled IoT surveillance environment. To this end, we tested our system on two settings with: 1) NVidia TITAN X (Pascal) having 12 GB onboard memory with a deep learning framework [42] running with Intel Core i5 CPU with Ubuntu OS and 64 GB RAM and 2) a Raspberry Pi 3 having 1024 MiB SDRAM and 1.2 GHz 64-bit ARMv8 Cortex-A53. Based on these two settings, our proposed system can process 34 fps and 5 fps, respectively. Since processing few frames in real-time are enough for detection of fire and the conventional cameras can capture 25~30 fps, thus our system is significant enough for real-time fire detection. The comparison of our system in terms of fps, accuracy, and false alarm rate with state-of-the-art using DS1 is given in Table **VII**.

From the incurred results, [17] seems to be the best method in terms of fps, however, its accuracy is low and this method is tested on a very small dataset, which is not benchmark. Also, it has a false alarm rate of 6.67% and the deployment details are not known. The method [9] achieved better processing speed of 60 fps with reasonable accuracy, however, the false alarm rate of 11.67% is high and not much recommendable. The CNNFire [25] achieved 20 fps with 94.50% accuracy, however, its false alarm rate of 8.87% is still high and not preferable, considering the critical nature of disaster management systems. Our proposed method achieved the best accuracy of 95.86% with false alarm rate of 0% using DS1. The running time of our method is 34 fps with setting 1 and 5 fps with setting 2, showing its superiority over the state-of-the-art.

TABLE VII
IMPLEMENTATION DETAILS WITH COMPARATIVE PERFORMANCE OF THE PROPOSED SYSTEM AND STATE-OF-THE-ART

Method	Fps	Accuracy (%)	False alarm rate (%)	Remarks
EMN_Fire	34	95.86	0	Setting 1
EMN_Fire	5	95.86	0	Setting 2
CNNFire [25]	20	94.50	8.87	Setting 1
CNNFire [25]	4	94.50	8.87	Setting 2
[9]	60	93.55	11.67	Intel dual core T7300 with 4 GM RAM
[9]	3	93.55	11.67	Raspberry Pi B (ARM processor with 700 MHz and 512 MiB RAM
[17]	70	92.59	6.67	-
[16]	20	90.32	5.88	Dual core 2.2 GHz

Besides the better performance, our employed architecture is light-weighted with fewer mega floating-point operations per second (MFLOPS) and reasonable size as given in Table **VIII.** It can be seen that our method needs fewer MFLOPS/image compared to other models, enabling it to execute several surveillance streams. Similarly, the size of our model (13.23

MB) is also reasonable and easily deployable on resource constrained devices. Another motivational point of our system is that it can be easily run on a raspberry Pi device (such as raspberry Pi 3), whose price is much affordable ($35). Considering the overall performance evaluation metrics, model size, and MFLOPS/image, we can claim that our system is the best candidate for early fire detection in certain surveillance in general and uncertain surveillance environment in particular, compared to existing fire detection systems.

TABLE VIII

MODELS SPECIFICATION IN TERMS OF MEGA FLOATING POINT OPERATIONS (MFLOS)/IMAGE AND SIZE

Method Name	MFLOPS/image	Size (MB)
EMN_Fire	**300**	*13.23*
CNNFire	833	**3.06**
GNetFire	1500	43.30
ANetFire	*720*	233

F. *5G Tactile Internet-Enabled Fire Detection System for Surveillance in Uncertain Industrial Environments*

According to the International Telecommunication Union, the Tactile Internet is an internet network that combines ultra-low latency with extremely high availability, reliability, and security". Unlike IoT that interconnects smart devices, the TI is going to control the IoT in real-time, needing ultra-reliable infrastructure [1]. The reason is that several tasks of critical nature (e.g., early fire detection in uncertain scenes during industrial surveillance) need to be executed remotely and instantly, requiring cheap edge infrastructure for ease of scalability. Considering these constraints, 5G can be a suitable underlying network infrastructure for such environment.

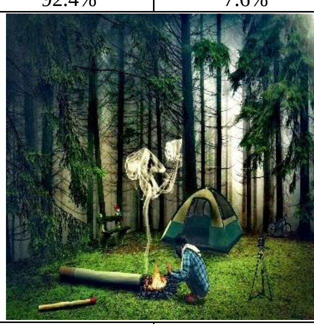
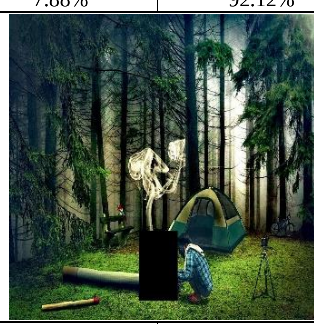
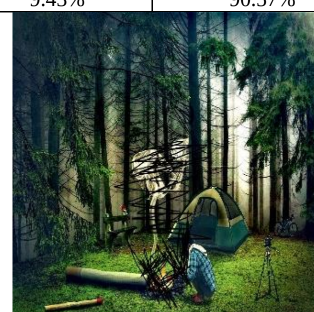
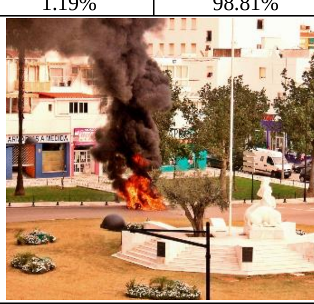
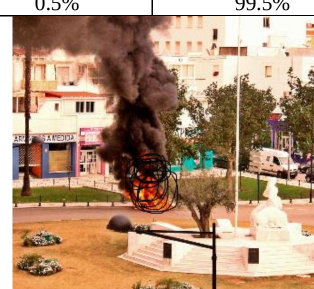
Method	Original input image		Fire blocked		Fire region affected with noise	
						
	Fire	Normal	Fire	Normal	Fire	Normal
EMN_Fire	**100.0%**	0.0%	1.87%	*98.13%*	**85.4%**	14.6%
CNNFire	97.9%	2.1%	1.8%	**98.2%**	9.0%	91.0%
GNetFire	65.08%	34.9%	2.9%	97.1%	46.6%	53.33%
ANetFire	92.4%	7.6%	7.88%	92.12%	9.43%	90.57%
						
	Fire	Normal	Fire	Normal	Fire	Normal
EMN_Fire	**59.01%**	40.98%	38.69%	61.31%	**64.43%**	35.56%
CNNFire	0.2%	99.8%	0.02%	**99.98%**	0.02%	99.98%
GNetFire	*48.5%*	51.49%	0.83%	*99.17%*	*48.9%*	51.11%
ANetFire	1.19%	98.81%	7.32%	92.68%	0.5%	99.5%
						
	Fire	Normal	Fire	Normal	Fire	Normal
EMN_Fire	**99.88%**	0.12%	3.51%	*96.49%*	*51.42%*	48.55%
CNNFire	97.47%	2.53%	0.98%	**99.02%**	**53.8%**	46.2%
GNetFire	9.36%	90.64%	4.15%	95.85%	4.4%	95.6%
ANetFire	88.5%	11.5%	11.1%	88.85%	51.34%	48.6%

Fig. 5: Robustness analysis using fire blocking and noise attack for the proposed method and other state-of-the-art. Best result is shown in bold font while the 2nd comparable result is shown in italic font

TI can intelligently combine multiple technologies at network and application level, enabling intelligence via mobile edge computing and data transmission over a 5G network. As described in previous sections, recently several CNN based fire detection approaches using edge intelligence are presented. These methods achieved reasonable accuracy for surveillance in certain IoT environment. However, their performance is limited in terms of fire detection in uncertain environment such as smoke, fog, and snow that can happen frequently in surveillance. Furthermore, the fire detection alert and representative video frames need reliable and instant reporting, considering the critical nature of disaster management. This goal can be achieved using a 5G TI-enabled fire detection system for which our proposed framework fits well, considering its promising accuracy, minimum false alarm rate, and response time. Furthermore, the size of the proposed model is reasonable due to usage of light-weight deep neural networks that favors its running time, making it suitable for fire detection during surveillance in uncertain industrial environments for mobile and embedded vision applications.

IV. Conclusion and Future Work

With the recent achievements of CNNs for solving numerous problems, researchers have applied them for abnormal event detection such as fire. Early detection of fire is very important to disaster management systems for which several CNN based fire detection methods using edge intelligence are presented to date. These methods have reasonable accuracy and execution time and are applicable to only certain environment. In case of uncertain environment having fog, smoke, and snow, their performance is limited. In addition, it is difficult to deploy computationally expensive fire detection models on resource constrained devices. Considering these motivations, an efficient CNN based method is proposed in this work for fire detection in videos of uncertain environment. Our method provides several advantages compared to recent fire detection approaches of complex and huge-sized CNN models such as AlexNet, SqueezeNet, and GoogleNet. First, our method is based on light-weight deep neural networks with no dense fully connected layers, making it computationally inexpensive. Second, the size of the resultant model is approximately 13 MB, which is easily deployable on mobile devices with embedded vision. Lastly, our method dominates state-of-the-art in terms of fire detection accuracy and number of false alarms as verified from experimental results. In addition, the robustness of our method against different attacks and its feasibility analysis also verify its effectiveness. We believe that our method is superior compared to state-of-the-art and a suitable candidate for integration with disaster management systems under the umbrella of 5G TI and industrial surveillance.

Our current method is focused on fire detection with reasonable model size for resource constrained devices in uncertain environment. This work can be extended for extraction of detailed contextual information from fire scenes such as object on fire, burning degree, and fire growth rate etc. Furthermore, a hybrid system can be developed by integrating smoke detection methods with the current work for intelligent management of fire disasters. Finally, our framework can be combined with industrial systems, 5G IoT, traffic, and robotics for more safe automation, traveling, richer, and trustworthy experience [43-47].

References

[1] M. Maier, M. Chowdhury, B. P. Rimal, and D. P. Van, "The tactile internet: vision, recent progress, and open challenges," *IEEE Communications Magazine,* vol. 54, pp. 138-145, 2016.

[2] M. Simsek, A. Aijaz, M. Dohler, J. Sachs, and G. Fettweis, "5G-enabled tactile internet," *IEEE Journal on Selected Areas in Communications,* vol. 34, pp. 460-473, 2016.

[3] K. Muhammad, R. Hamza, J. Ahmad, J. Lloret, H. H. G. Wang, and S. W. Baik, "Secure Surveillance Framework for IoT systems using Probabilistic Image Encryption," *IEEE Transactions on Industrial Informatics,* 2018.

[4] N. Pogkas, G. E. Karastergios, C. P. Antonopoulos, S. Koubias, and G. Papadopoulos, "Architecture design and implementation of an ad-hoc network for disaster relief operations," *IEEE Transactions on Industrial Informatics,* vol. 3, pp. 63-72, 2007.

[5] 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.

[6] U. L. Puvvadi, K. Di Benedetto, A. Patil, K.-D. Kang, and Y. Park, "Cost-effective security support in real-time video surveillance," *IEEE Transactions on Industrial Informatics,* vol. 11, pp. 1457-1465, 2015.

[7] S. Mumtaz, A. Bo, A. Al-Dulaimi, and K.-F. Tsang, "5G and beyond mobile technologies and applications for industrial IoT (IIoT)," *IEEE Transactions on Industrial Informatics,* pp. 1-1, 2018.

[8] D. Guha-Sapir, P. Hoyois, and R. Below, "Annual disaster statistical review 2015: the numbers and trends. https://www.cred.be/sites/default/files/ADSR_2015.pdf."

[9] 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.

[10] B. U. Töreyin, Y. Dedeoğlu, U. Güdükbay, and A. E. Cetin, "Computer vision based method for real-time fire and flame detection," *Pattern recognition letters,* vol. 27, pp. 49-58, 2006.

[11] D. Han and B. Lee, "Development of early tunnel fire detection algorithm using the image processing," in *International Symposium on Visual Computing*, 2006, pp. 39-48.

[12] T.-H. Chen, P.-H. Wu, and Y.-C. Chiou, "An early fire-detection method based on image processing," in *Image Processing, 2004. ICIP'04. 2004 International Conference on*, 2004, pp. 1707-1710.

[13] G. Marbach, M. Loepfe, and T. Brupbacher, "An image processing technique for fire detection in video images," *Fire safety journal,* vol. 41, pp. 285-289, 2006.

[14] T. Celik and H. Demirel, "Fire detection in video sequences using a generic color model," *Fire Safety Journal,* vol. 44, pp. 147-158, 2009.

[15] A. Rafiee, R. Dianat, M. Jamshidi, R. Tavakoli, and S. Abbaspour, "Fire and smoke detection using wavelet analysis and disorder characteristics," in *Computer Research and Development (ICCRD), 2011 3rd International Conference on*, 2011, pp. 262-265.

[16] Y. H. Habiboğlu, O. Günay, and A. E. Çetin, "Covariance matrix-based fire and flame detection method in video," *Machine Vision and Applications,* vol. 23, pp. 1103-1113, 2012.

[17] R. Di Lascio, A. Greco, A. Saggese, and M. Vento, "Improving fire detection reliability by a combination of videoanalytics," in *International Conference Image Analysis and Recognition*, 2014, pp. 477-484.

[18] P. V. K. Borges and E. Izquierdo, "A probabilistic approach for vision-based fire detection in videos," *IEEE transactions on circuits and systems for video technology,* vol. 20, pp. 721-731, 2010.

[19] M. Mueller, P. Karasev, I. Kolesov, and A. Tannenbaum, "Optical flow estimation for flame detection in videos," *IEEE Transactions on Image Processing,* vol. 22, pp. 2786-2797, 2013.

[20] 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.

[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] J. Sharma, O.-C. Granmo, M. Goodwin, and J. T. Fidje, "Deep Convolutional Neural Networks for Fire Detection in Images," in *International Conference on Engineering Applications of Neural Networks*, 2017, pp. 183-193.

[23] 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.

[24] 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.

[25] 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.

[26] M. Sajjad, M. Nasir, K. Muhammad, S. Khan, Z. Jan, A. K. Sangaiah, *et al.*, "Raspberry Pi assisted face recognition framework for enhanced law-enforcement services in smart cities," *Future Generation Computer Systems,* 2017/11/16/ 2017.

[27] A. Ullah, J. Ahmad, K. Muhammad, M. Sajjad, and S. W. Baik, "Action Recognition in Video Sequences using Deep Bi-Directional LSTM With CNN Features," *IEEE Access,* vol. 6, pp. 1155-1166, 2018.

[28] Y. Liu, L. Nie, L. Liu, and D. S. Rosenblum, "From action to activity: Sensor-based activity recognition," *Neurocomputing,* vol. 181, pp. 108-115, 2016.

[29] M. Sajjad, S. Khan, T. Hussain, K. Muhammad, A. K. Sangaiah, A. Castiglione, *et al.*, "CNN-based anti-spoofing two-tier multi-factor authentication system," *Pattern Recognition Letters,* 2018.

[30] J. Ahmad, K. Muhammad, J. Lloret, and S. W. Baik, "Efficient Conversion of Deep Features to Compact Binary Codes Using Fourier Decomposition for Multimedia Big Data," *IEEE Transactions on Industrial Informatics,* vol. 14, pp. 3205-3215, 2018.

[31] J. Ahmad, K. Muhammad, S. Bakshi, and S. W. Baik, "Object-oriented convolutional features for fine-grained image retrieval in large surveillance datasets," *Future Generation Computer Systems,* vol. 81, pp. 314-330, 2018.

[32] N. Rahim, J. Ahmad, K. Muhammad, A. K. Sangaiah, and S. W. Baik, "Privacy-preserving image retrieval for mobile devices with deep features on the cloud," *Computer Communications,* vol. 127, pp. 75-85, 2018/09/01/ 2018.

[33] 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.

[34] 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.

[35] F. N. Iandola, M. W. Moskewicz, K. Ashraf, S. Han, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 1MB model size," *arXiv preprint arXiv:1602.07360,* 2016.

[36] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510-4520.

[37] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, *et al.*, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861,* 2017.

[38] D. Y. Chino, L. P. Avalhais, J. F. Rodrigues, and A. J. Traina, "BoWFire: detection of fire in still images by integrating pixel color and texture analysis," in *2015 28th SIBGRAPI Conference on Graphics, Patterns and Images*, 2015, pp. 95-102.

[39] K. Muhammad, M. Sajjad, M. Y. Lee, and S. W. Baik, "Efficient visual attention driven framework for key frames extraction from hysteroscopy videos," *Biomedical Signal Processing and Control,* vol. 33, pp. 161-168, 2017.

[40] L. Rossi, M. Akhloufi, and Y. Tison, "On the use of stereovision to develop a novel instrumentation system to extract geometric fire fronts characteristics," *Fire Safety Journal,* vol. 46, pp. 9-20, 2011.

[41] S. Rudz, K. Chetehouna, A. Hafiane, H. Laurent, and O. Séro-Guillaume, "Investigation of a novel image segmentation method dedicated to forest fire applicationsThis paper is dedicated to the memory of Dr Olivier Séro-Guillaume (1950–2013), CNRS Research Director," *Measurement Science and Technology,* vol. 24, p. 075403, 2013.

[42] 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.

[43] G. P. Fettweis, "The tactile internet: Applications and challenges," *IEEE Vehicular Technology Magazine,* vol. 9, pp. 64-70, 2014.

[44] M. K. Afzal, Y. B. Zikria, S. Mumtaz, A. Rayes, A. Al-Dulaimi, and M. Guizani, "Unlocking 5G Spectrum Potential for Intelligent IoT: Opportunities, Challenges, and Solutions," *IEEE Communications Magazine,* vol. 56, pp. 92-93, 2018.

[45] S. Mumtaz, A. Alsohaily, Z. Pang, A. Rayes, K. F. Tsang, and J. Rodriguez, "Massive Internet of Things for industrial applications: Addressing wireless IIoT connectivity challenges and ecosystem fragmentation," *IEEE Industrial Electronics Magazine,* vol. 11, pp. 28-33, 2017.

[46] Z. Yan, P. Zhang, and A. V. Vasilakos, "A survey on trust management for Internet of Things," *Journal of network and computer applications,* vol. 42, pp. 120-134, 2014.

[47] S. Mumtaz, K. M. S. Huq, M. I. Ashraf, J. Rodriguez, V. Monteiro, and C. Politis, "Cognitive vehicular communication for 5G," *IEEE Communications Magazine,* vol. 53, pp. 109-117, 2015.

Khan Muhammad (S'16–M'18) received the Ph.D degree in Digital Contents from Sejong University, South Korea. He is currently an Assistant Professor in the Department of Software, Sejong University, South Korea. His research interests include medical image analysis (brain MRI, diagnostic hysteroscopy and wireless capsule endoscopy), information security (steganography, encryption, watermarking and image hashing), video summarization, computer vision, fire/smoke scene analysis, and video surveillance. He has published over 50 papers in peer-reviewed international journals and conferences in these research areas with target venues as IEEE COMMAG, TII, TIE, TSMC-Systems, IoTJ, Access, TSC, Elsevier INS, Neurocomputing, PRL, FGCS, COMCOM, COMIND, JPDC, PMC, BSPC, CAEE, Springer NCAA, MTAP, JOMS, and RTIP, etc. He is also serving as a professional reviewer for over 40 well-reputed journals and conferences.

Salman Khan (S'19) received his Bachelor's degree in Computer Science from Islamia College Peshawar, Pakistan in 2017 with research in medical image analysis. Currently, he is pursuing MS degree in software convergence from Sejong University, Seoul, Republic of Korea. He is working as a research assistant at Intelligent Media Laboratory (IM Lab). His research interests include Medical Image Analysis, Computer Vision, Deep Learning, Fire/Smoke Scene Analysis, and Video Surveillance.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2019.2897594, IEEE Transactions on Industrial Informatics

Mohamed Elhoseny completed his B.S in Computer and Information Sciences from Mansoura Univeristy, Faculty of Computers and Information, Egypt. Then, Dr. Elhoseny received his Ph.D. in Computer and Information Sciences from Mansoura University Egypt (in a scientific research channel with Department of Computer Science and Engineering, University of North Texas, USA). His Ph.D. thesis was awarded the best PhD thesis prize (2016) at Mansoura University. Dr. Elhoseny is currently an Assistant Professor at the Faculty of Computers and Information, Mansoura University, Egypt where he also teaches several courses such as "Security in Computing", "Mobile Computing", and "Information Theory". Collectively, Dr. Elhoseny authored/co-authored over 50 International Journal articles, Conference Proceedings, Book Chapters, and 1 Springer brief book. His research interests include Network Security, Cryptography, Machine Learning Techniques, Internet of Things, and Quantum Computing. He has several publications in reputed and high impact journals published by IEEE, Elsevier, Springer, and others.

Syed Hassan Ahmed (S'13–M'17–SM'18) received the B.S. degree in computer science from the Kohat University of Science and Technology, Pakistan, and the master's combined Ph.D. degree from the School of Computer Science and Engineering (SCSE), Kyungpook National University (KNU), South Korea. In 2015, he was a Visiting Ph.D. Student at Georgia Tech, Atlanta, GA, USA. He is currently an Assistant Professor with the Computer Science Department, Georgia Southern University (GSU), Statesboro, GA, USA. Before starting at GSU, he was a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Central Florida, Orlando, FL, USA. He has authored/co-authored over 100 international publications, including journal articles, conference proceedings, book chapters, and three books. His research interests include sensor and ad hoc networks, cyber-physical systems, vehicular communications, and future Internet. He is an ACM Professional Member. From 2014 to 2016, he consequently received the Research Contribution Awards by SCSE and KNU. In 2016, his work on robust content retrieval in future vehicular networks leads him to receive the Qualcomm Innovation Award at KNU.

Sung Wook Baik (M'16) received the B.S degree in computer science from Seoul National University, Seoul, Korea, in 1987, the M.S. degree in computer science from Northern Illinois University, Dekalb, in 1992, and the Ph.D. degree in information technology engineering from George Mason University, Fairfax, VA, in 1999. He worked at Datamat Systems Research Inc. as a senior scientist of the Intelligent Systems Group from 1997 to 2002. In 2002, he joined the faculty of the College of Electronics and Information Engineering, Sejong University, Seoul, Korea, where he is currently a Full Professor and the Chief of Sejong Industry-Academy Cooperation Foundation. He is also the head of Intelligent Media Laboratory (IM Lab) at Sejong University. His research interests include computer vision, multimedia, pattern recognition, machine learning, data mining, virtual reality, and computer games. He is a member of the IEEE.