ARTICLE

Convolutional Neural Network Model for Fire Detection in Real-Time Environment

Abdul Rehman, Dongsun Kim[*] and Anand Paul

School of Computer Science and Engineering, Kyungpook National University, Daegu, Korea

*Corresponding Author: Dongsun Kim. Email: darkrsw@knu.ac.kr

ABSTRACT

Disasters such as conflagration, toxic smoke, harmful gas or chemical leakage, and many other catastrophes in the industrial environment caused by hazardous distance from the peril are frequent. The calamities are causing massive fiscal and human life casualties. However, Wireless Sensors Network-based adroit monitoring and early warning of these dangerous incidents will hamper fiscal and social fiasco. The authors have proposed an early fire detection system uses machine and/or deep learning algorithms. The article presents an Intelligent Industrial Monitoring System (IIMS) and introduces an Industrial Smart Social Agent (ISSA) in the Industrial SIoT (ISIoT) paradigm. The proffered ISSA empowers smart surveillance objects to communicate autonomously with other devices. Every Industrial IoT (IIoT) entity gets authorization from the ISSA to interact and work together to improve surveillance in any industrial context. The ISSA uses machine and deep learning algorithms for fire-related incident detection in the industrial environment. The authors have modeled a Convolutional Neural Network (CNN) and compared it with the four existing models named, FireNet, Deep FireNet, Deep FireNet V2, and Efficient Net for identifying the fire. To train our model, we used fire images and smoke sensor datasets. The image dataset contains fire, smoke, and no fire images. For evaluation, the proposed and existing models have been tested on the same. According to the comparative analysis, our CNN model outperforms other state-of-the-art models significantly.

KEYWORDS

Fire detection; industrial surveillance system; smart devices; smart social agent (SSA); machine learning algorithms; CNN

1 Introduction

IoT reformed as a mature field that promises extensive connections to the Internet. The researcher's passion is transforming every real-world object into a smart one. So, the unification of social networks, mobile communication, and the Internet has brought a revolution in information technology, and it is widely accepted as Social IoT (SIoT). Ultimately, IoT has entrusted an inspiring collocation to build a vehement industrial system and recently installed massive industrial IoT (IIoT) and its applications. For example, automobile's locations, oversee their odyssey, and predict their approaching neighborhood and traffic conditions [1–3] are easily trackable by the administration using intelligent transportation system (ITS) with the assistance of IoT. The primary goal of IoT

is to establish secure autonomous connections between intelligent devices and applications for data exchange. However, searching for information within such a complex network can be challenging due to issues such as time complexity, redundant information, and unwanted data. To mitigate these issues, researchers have proposed a model named an individual's small-world, which reduces network complexity and enables efficient and precise information retrieval [4–6]. Our previous work proposes a system for monitoring and detecting events early in an industrial setting that combines the social IoT paradigm with a small-world network. This system leverages the collaboration of all IoT devices with a Smart Social Agent (SSA) [7–10].

The early identification and prevention of fires, which can have catastrophic effects on both human life and the environment, is a critical component of industrial safety. Traditional fire detection systems have showed potential, especially those that use computer vision techniques like Convolutional Neural Networks (CNNs) [11], but they still have challenges including manual feature selection, a lot of computation, and slow detection speed [12]. A trustworthy algorithm that can accurately identify fires, automate feature selection, and ultimately save lives and save the environment is therefore urgently needed [13].

Natural disasters like fires are incredibly harmful because of the havoc they can wreak on human life and the natural world. The detection of fires in open areas has recently emerged as a critical issue regarding human life safety and a formidable challenge. Australia's bushfires, which started in 2019 and continued through March 2020, were just one of many wide-ranging wildfires also called forest fires that broke out worldwide that year. Approximately 500 million animals perished in the fire. "Wildfires in 2020" [14] refers to an article about a similar, deadly fire in California, a state in the United States. There has been a growing focus on the importance of fire detection systems in recent years, and these systems have proven invaluable in preventing fires and saving lives and property. Sensor detection systems can detect fire signs like light, heat, and smoke [15].

To protect people, detecting fires in the open air has become a challenging and essential task. According to data gathered worldwide, fires significantly threaten manufactured structures, large gatherings, and densely populated areas. Property loss, environmental damage, and the threat to human and animal life are all possible results of such events. The environment, financial systems, and lives are all put at risk due to these occurrences. The damage caused by such events can be reduced significantly if measures are taken quickly. Automated systems based on vision are beneficial in spotting these kinds of occurrences. To address the requirement for early detection and prevention of industrial accidents, with a specific focus on fire detection, this study proposes a thorough framework for an Intelligent Industrial Monitoring System. To provide safety and security in industrial settings, the suggested system incorporates IIoT devices, such as sensors, actuators, cameras, unmanned aerial vehicles (UAVs), and industrial robots. These gadgets communicate socially with the Industrial SSA (ISSA), facilitating smart cooperation and communication in urgent circumstances. CNN models based on deep learning are used by the IIMS' intelligent layer to identify fires with a high degree of recall and accuracy. The ISSA reduces false alarms and verifies fire events by activating all surveillance equipment in urgent situations. After verification, the system creates alerts, notifies the appropriate authorities, and launches UAVs and industrial robots to monitor and evacuate the impacted area. The cloud infrastructure facilitates a communication route between the intelligent layer and the application layer, which provides real-time data. The SSA maintains the emergency report on the cloud and uses machine learning techniques to identify distinct situations.

In order to fill in research gaps and advance the field of early detection and prevention of industrial accidents, particularly in fire detection, the proposed Intelligent Industrial Monitoring System (IIMS) makes several significant contributions. The following are our work's main contributions:

- Development of an Intelligent Industrial Monitoring System (IIMS) that integrates Industrial Internet of Things (IIoT) devices, including sensors, actuators, cameras, Unmanned Aerial Vehicles (UAVs), and industrial robots, to ensure safety and security in industrial settings.
- Introduction of a Smart Social Agent (SSA) that facilitates intelligent communication and collaboration among IoT devices during critical situations, enhancing the efficiency and effectiveness of the monitoring system.
- Utilization of deep learning-based Convolutional Neural Networks (CNNs) in the intelligent layer of the IIMS to achieve high accuracy and recall in fire detection, addressing the limitations of manual feature selection, computation requirements, and detection speed.
- Establishment of a robust communication channel between the intelligent layer and the application layer through cloud infrastructure, enabling real-time information sharing and facilitating timely decision-making.
- Mitigation of false alarms through the collaborative behavior of IoT devices, where nearby sensors, devices, cameras, and UAVs are activated to sense the environment instead of generating unnecessary alarms.
- Generation of warnings, notifications, and emergency reports on the cloud, enabling seamless communication with concerned authorities such as fire brigades, police, and ambulance services.
- Activation of industrial robots and UAVs for evacuation and monitoring of affected areas, leveraging the capabilities of automation to enhance emergency response and safety measures.

The proposed architecture has several advantages, including:

- The social behavior of IoT devices provides in-depth surveillance. For example, if smoke, heat, or light sensors sense a value higher than the threshold, nearby sensors, devices, cameras, and UAVs will be activated to sense the environment instead of generating an alarm. Devices send perceived and captured data to the intelligent layer for validation.
- The intelligent layer receives data and uses CNN models to detect and validate fire incidents. If two or more surveillance devices detect the fire, ISSA activates actuators and robots for first aid.
- ISSA updates the emergency report on the cloud and alerts concerned authorities such as the fire brigade, police, and ambulance.

The manuscript is organized into several sections to present a clear and systematic account of the research. Specifically, Section 2 offers an overview of the related literature. Section 3 presents the proposed methodology for the Industrial Internet of Things (IIoT), while Section 4 details the experimental setup. The findings of the study are then discussed in Section 5. Finally, Section 6 provides a conclusion that summarizes the key findings and their implications in a concise and professional manner.

2 Related Work

The essential technology for IoT is Radio-Frequency Identification (RFID) innovation. It enables microchips to transfer the identification info of a visitor through wireless communication. It uses interconnected smart sensors to sense and monitor. Wireless Sensors Network (WSN) and RFID have contributed substantially to the advancement of IoT [16]. As a result, IoT has gained popularity in numerous industries, including logistics, manufacturing, retailing, and medicine [17–21].

Additionally, it also influences new Information & Communication Technologies (ICT) and venture systems innovations. Compatibility, effectivity and interoperability achieved in IoT by following standardization procedures at global scale [22–24]. Several countries and organizations can bring incredible financial benefits due to the amelioration of IoT standards. Organizations such as the International Telecommunication Union (ITU), International Electro-technical Commission (IEC), China Electronics Standardization Institute (CESI), the American National Standards Institute (ANSI) and several other are working for fulfilling IoT requirements [25].

Many companies are thriving in the enrichment of IoT criteria. Solid coordination among standardization companies is fundamental need to collaborate between international, national, and regional organizations. Mutually accepted programs and requirements, customers can apply for IoT applications and solutions. It will be deployed and utilized while conserving development and upkeep costs effectively in future perspective. IIoT technology will flourish with innovation by following the standard procedures set by ISO and it will be used in all walks of life. Schlumberger monitors subsea conditions by taking the trip of oceans for gathering relevant information up to many years without using the human force with the help of UAVs. Moreover, mining markets may also get benefit with the advancement of IIoT in terms of remote tracking and sensing as it will reduce the risk of accidents. A leading mining company of Australia, Rio Tinto, which plans to remove human resources for mining purpose by following autonomous mining procedures [26].

Despite the pledge, numerous obstacles in realizing the chances supplied by IIoT define future research. The essential difficulties stem from the need for energy-efficient operation and real-time performance in vibrant settings. As per the statistics shared by The International Labor Organization (ILO), "151 employees face work-related injuries in every 15 s". The IoT has addressed safety and security problems and conserved $220 billion yearly in injuries and health problem prices. RFID cards issued to all works of different industries including gas, oil and coal mining, and transportation sector which collect the live time location data as well as it monitor the heart beat rate, galvanic skin action, skin temperature, and other specifications. The collected data will be evaluated in the cloud compared to the contextual information. Any irregular behavior detection in the body generates an alert and avoids mishaps [26,27].

The automatic detection of fire using deep learning models and computer vision techniques has opened multiple research avenues for several academic communities due to the similarities between fire and other natural phenomena, such as sunlight and artificial lighting [28]. Although methods such as [29] show promise, there is a better solution for image-based problems. Deep learning techniques played pivotal role for problem solving for computer vision [30]. The use of deep learning has become pervasive across a range of real-time applications, including image and video object recognition and classification, speech recognition, natural language processing, and more. This technique has proven highly effective in enabling these applications to recognize and classify data in real-time with remarkable accuracy [31]. Therefore, this research article provides the comprehensive overview and authenticate the visual analysis-based early fire detection systems.

Computer vision and deep neural networks based models like CNNs employed for fire detection, which yielded promising results in this proposed research article. Therefore, CNNs models have gain the interest of few researchers, as they believe that CNNs could improve fire detection performance. Current literature mentions a variety of shapes, colors, textures, and motion attributes as potential solutions for fire detection systems. By analyzing the kinetic properties of smoke, reference [32] created an algorithm for smoke detection. CNN was used to generate suspect features using a machine learning-based strategy; the approach taken was background dynamic update, and the methodology

used was a dark channel prior algorithm; the result could be implemented with relative ease and was widely applicable.

3 Intelligent Industrial Surveillance System

Certainly! Fig. 1 illustrates the sub-architecture of the fire detection system in a format of hardware-based block diagram. This diagram consists of three major modules, which are:

- Surveillance Area Module: This module monitors the environment for potential fire hazards. It includes various sensors and cameras that detect smoke, heat, and other fire signs. The data from these sensors is then fed into the next module.
- IoT Fire System Module: This module receives data from the Surveillance Area Module and processes it using IoT technologies. The IoT Fire System Module could include hardware such as microcontrollers or IoT gateways communicating with the cloud or other remote servers. This module could also include software algorithms that analyze the data and determine whether a fire has started or is likely to begin soon. This module can alert the third module if a fire is detected.
- Responders Module: This module is responsible for dispatching responders such as firefighters, ambulances, or police to the location of the fire. The Responders Module receives alerts from the IoT Fire System Module and can use various communication channels such as mobile phones, two-way radios, or other wireless devices to alert and coordinate the responders. The Responders Module could also include GPS technology to help responders navigate to the location of the fire.
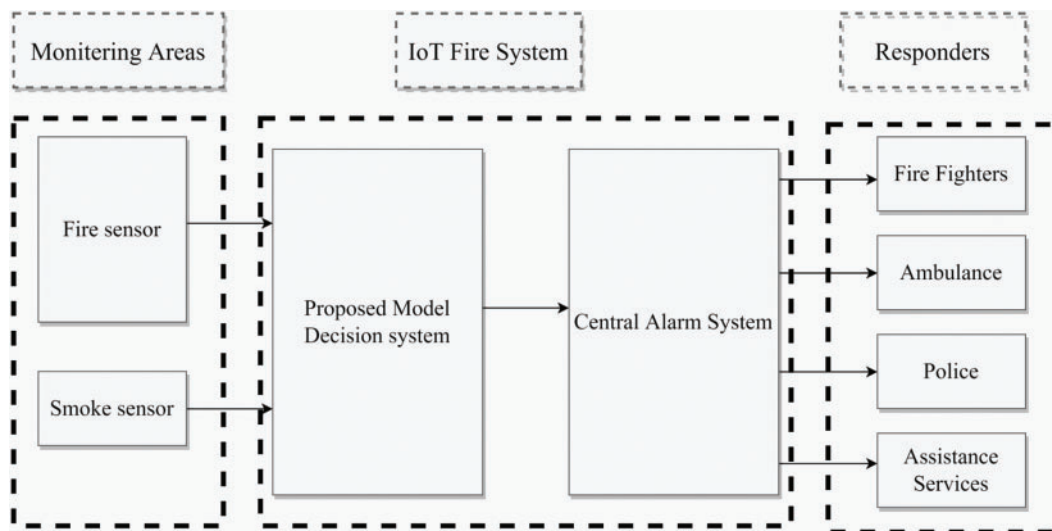

Figure 1: Hardware-based block diagram

Overall, this flow diagram represents a robust fire detection system that utilizes a hardware-based approach to quickly and effectively respond to fire emergencies. The three modules work together seamlessly to detect fires, analyze data, and dispatch responders to the scene. With this system in place, it is possible to reduce the damage caused by fires and protect human life and property.

The IIoT paradigm is crucial in various industries by monitoring industrial environments and preventing monetary and social damage. The architectural design aims to minimize damage by

providing attentive surveillance planning and monitoring of the surrounding environment. In case of an emergency, all devices in the ISIoT paradigm communicate with each other, and the ISSA validates the event using machine learning-based algorithms to prevent false alarms. The IIMS architecture comprises various components such as architectural style, intelligent objects, communication systems, cloud services, Intelligent layers, and application interfaces. The architecture emphasizes the importance of expandability, scalability, modularity, and interoperability to support intelligent objects critical to industrial environments. The industrial settings require a flexible architecture to support continuously moving or interacting intelligent objects. The distributed and diverse nature of the SIoT necessitates an event-driven architecture that is capable of achieving interoperability among various devices.

3.1 Hardware Layer

The SIoT is a network of socially allied smart devices in terms of co-workers, co-location, co-ownership, etc., in which objects may interact to conduct environmental monitoring as a team. We introduced a 5-layered surveillance architecture for IIoT, shown in Fig. 2.
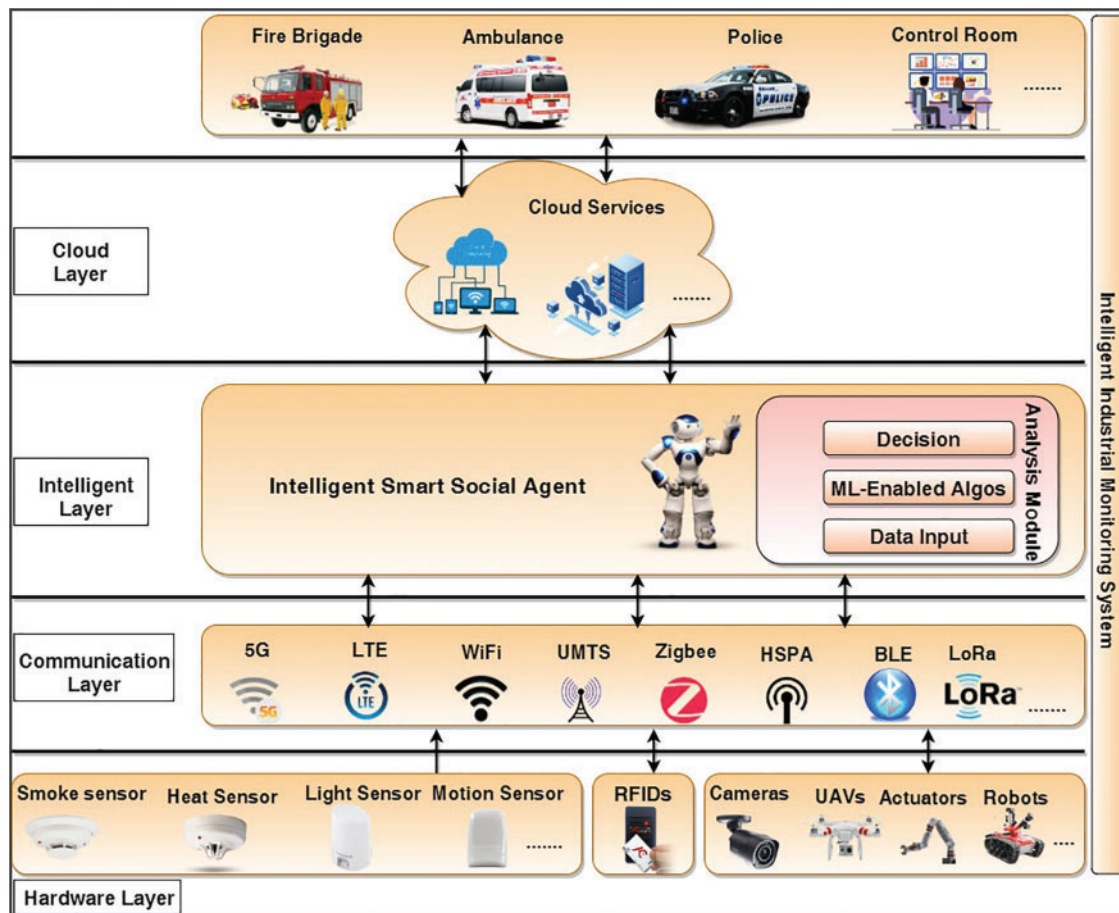

Figure 2: Intelligent industrial surveillance system

In the first layer, called the "hardware layer," electronic devices such as cameras, drones, and other sensors are activated for perceiving the environment and transmitting data. Nowadays, almost all

firms use these clever devices for several purposes. RFIDs are widely used to track and count items, while sensors are crucial in sensing the surroundings. In addition to detecting fires, poisonous gases, and suspicious movements, these surveillance systems monitor manufacturing quality, count goods, preserve energy, manage household and irrigation water, assess agricultural land, etc. Closed-circuit television (CCTV) and a surveillance drone have significantly improved or bolstered the monitoring system. CCTV offers visual monitoring, while surveillance drones and unmanned aerial vehicles (UAVs) monitor locations where it is impossible to place static cameras or sensors. In addition, industrial robots and actuators have been suggested for this design. By functioning autonomously in crisis scenarios, these self-governing devices safeguard the ecosystem from massive harm.

3.2 Communication Layer

Communication solutions between edge devices or between edge devices and clouds allow intelligent device connectivity and data exchange. The communication layer helps by linking all smart things and enabling them to exchange data with other connected devices. Additionally, the communication layer may collect data from existing IT infrastructures (e.g., agriculture, healthcare, and so on). The Internet of Things encompasses various electrical equipment, mobile devices, and industrial machines. Each device has its own set of capabilities for data processing, communication, networking, data storage, and transmission. Smartwatches and smartphones, for example, serve different purposes. Effective communication and networking technologies are essential for enabling intelligent devices to interact with each other seamlessly. Smart objects can utilize either wired or wireless connections for this purpose. In the IoT realm, wireless communication technologies and protocols have recently seen significant advancements. Communication protocols such as 5G, Wi-Fi, LTE, HSPA, UMTS, ZigBee, BLE, Lo-Ra, RFID, NFC, and LoWPAN have contributed immensely to data transmission between connected devices. As technology evolves, the IoT is expected to play an increasingly important role in developing wireless communication technologies and protocols.

3.3 Intelligent Layer

The Intelligent Layer, which establishes the Industrial Smart Internet of Things (ISIoT) paradigm and integrates the Social Smart Agent (SSA) into an industrial context, is the central element of the proposed architecture. The management of the surveillance system and facilitation of seamless communication between the connected devices are the primary goals of the Intelligent Layer. In order to ensure vigilant event detection, avoid damage, and reduce false alarms, the SSA is crucial. The layer is made up of a number of parts, such as sensors, closed-circuit television (CCTV) cameras, and unmanned aerial vehicles (UAVs), which constantly monitor the environment and send the data they collect to the cloud. The SSA starts a notification process among nearby devices when an incident occurs to see if other devices have also noticed the event. The SSA verifies the occurrence of an incident by comparing sensor values against a predetermined threshold. As a result, it produces an emergency alarm, turns on actuators to secure the incident site, and notifies the appropriate authorities right away. In parallel, the surveillance drone is used to record in-depth footage of the incident scene, minimizing interference with daily life. Convolutional neural networks (CNNs) are then used to detect fires using the drone images that were collected. To manage industrial surveillance and avert potential risks, the Intelligent Layer's architecture should put a strong emphasis on effective event detection, dependable communication, sound decision-making, and coordinated actions.

Convolutional neural networks

One type of deep neural network that draws inspiration from biology is the convolutional neural network [33]. Applications of deep convolutional neural networks (CNNs) in computer vision, such as image restoration, classification, localization [34–37], segmentation [38,39], and detection [40,41], are highly effective and efficient. The core idea behind CNN is to continuously break down the problem into smaller chunks until a solution is found. By training the model from a raw pixel value to a classifier, we can avoid the complex preprocessing steps common in ML. An elementary model of CNN is a multi-layered feedforward network with stacked convolutional and subsampling layers. The deepest layers of CNNs are used for classification based on extensive reasoning. Here is a breakdown of what each layer entails.

Convolution layers

In convolutional layers, the image (input) undergoes a convolutional operation, and then the resulting data is passed to the following layer. Each node in the convolutional layer comprises receptive fields built from the units in the layers below it. The neurons in these fields derive fundamental visual features, such as corners, endpoints, and oriented edges. Multiple features can be extracted from the many feature maps in this layer. All units on a given feature map share the same biases and weights, ensuring that the features detected apply equally to all input locations. Researchers commonly use the expression [42] to describe the shape of a convolution layer, see Eq. (1).

$$X_j^l = f\left[\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_i^l\right] \tag{1}$$

In convolutional layers, the input maps collection is denoted by M_j, where k is the kernel size determining the extent of convolution applied to the image. Additionally, b represents a bias, and X_j^l denotes the output as the j_{th} feature map of that convolutional layer.

Subsampling layers

The pooling and subsampling layer plays a crucial role in reducing the complexity of the feature map's resolution by performing sub-sampling and local averaging. This layer is responsible for downsampling the convolutional layer's output, which reduces the computation required for subsequent layers.

In addition to this, it takes away the sensitivity of the output. The representation of a sub-sampling layer looks like this: [42], see Eq. (2).

$$X_j^l = f\left[\beta_{jdown}^l \ (x_i^{l-1}) + b_i^l\right] \tag{2}$$

where (down) refers to a process that is known as sub-sampling. In most cases, the sub-sampling (down) function will offer an n-by-n block to calculate the final output in the input picture. This will result in a normalized output and n times smaller than the original. Where b represents the additive bias and represents the multiplicative bias. We recommend using the following CNN model:

Proposed CNN model

The CNN model utilized in this study was based on the design of AlexNet [43], with a few minor modifications tailored to our specific problem. To reduce complexity, we limited the number of output neurons to just two. Our model consists of ten layers, including five convolutional layers, five max-pooling layers, and two fully connected layers, each comprising a total of 4096 vertices. When presented with an input image x of dimensions H × W × C, where H, W, and C represent the image's height,

width, and depth, respectively, a filter (also known as a kernel) w of dimensions h × w × c is applied to extract local features from the image.

A feature map y of dimension $(H - h + 1) \times (W - w + 1) \times F$ is produced by the following mathematical operation presented in Eq. (3):

$$y_{i,j,f} = \sum_{k=0}^{h-1} \sum_{l=0}^{w-1} \sum_{m=0}^{c-1} x_{i+k,j+l,m} \cdot w_{k,l,m,f} \tag{3}$$

Here, i indicates the height index of a feature map, the width index is indicated by j whereas, f is used for depth. Similarly, the index of height, width, and depth for the filter is represented by k, l, m respectively.

In addition, we switch the order of the max-pooling and normalizing layers, which were previously located between the first and second convolutional layers, and move them to the fifth position. Max pooling is a downsampling operation that reduces the spatial dimensions of the feature map while retaining the most important features. Pooling filter of dimension h x w is applied to accomplish the Max pooling with a stride of s over a feature map y of dimension $(H \times W \times F)$ to retain the max value of every region. The Eq. (4) is the mathematical formulation of Max pooling.

$$y_{i,j,f} = max_{k=0}^{h-1} \, max_{l=0}^{w-1} y_{i \cdot s+k, \, j \cdot s+l, f} \tag{4}$$

Here, i indicates the height index of a feature map after pooling, the width index is indicated by j whereas, f is used for depth. Similarly, the index of height, width, and depth for the pooling filter is represented by k, l, m respectively. In the last layer of our classification system, we used the SoftMax classifier to make high-reason determinations. The addition of non-linearity is essential for neural networks, and this is achieved through the use of activation functions. One such function is the rectified linear unit (ReLu), which has been widely used due to its effectiveness. Another variation of ReLu is the leaky ReLu activation function, which introduces a small positive slope for negative inputs, thereby addressing the problem of "dead neurons" that can occur with ReLu. Mathematically, the leaky ReLu activation function can be expressed as following operation in Eq. (5) that defines the function.

$$f(x) = \begin{cases} x & if \ x > 0 \\ ax & if \ x \leq 0 \end{cases} \tag{5}$$

where x is the input to the activation function, y is the output, and a is a small positive slope for negative inputs.

The final part of the CNN is the fully connected layers, which perform the final classification of the input image. Given an input x of size n, a fully connected layer is a matrix multiplication of the input and a weight matrix w of size $x \times m$ followed by an activation function. The operation that defines a fully connected layer can be expressed mathematically as Eq. (6):

$$y = f(w \cdot x + b) \tag{6}$$

where f is the activation function, b is the bias term, and y is the output of the fully connected layer. Common activation functions in fully connected layers include the sigmoid, ReLu, and softmax functions. The activation functions in a convolutional neural network (CNN) play a crucial role in determining the output of the network. The sigmoid function is commonly used to map input values to a range between 0 and 1, while the ReLu function is used to transform negative input values to 0 and positive input values to their original value. The softmax function is used in the final layer of a CNN

to classify the input image by computing the exponential of each input value and then normalizing the result to produce a probability distribution over the classes. The class with the highest probability is considered as the final output of the CNN.

In Fig. 3, the architecture of our model is presented, which involves resizing the input image to 256 pixels in width, 256 pixels in height, and 3 pixels in depth using our CNN model. The 1st layer applies a filter with 96 kernels of size $11 \times 11 \times 3$ and a stride of 4 pixels. The outcome of this layer undergoes pooling, which reduces the data's complexity and dimensionality. Next, the 2nd layer applies 256 kernels of size $5 \times 5 \times 64$ with a stride of 2, followed by pooling. The 3rd layer employs 384 kernels of size $3 \times 3 \times 256$, without a pooling procedure. The rest of the layers use filters with kernels of 384 and 256, respectively, with a stride of 1. After the 5th layer, the pooling layer uses 3×3 filters. Finally, the last classification step is performed on two fully connected layers, each with 4096 neurons, and the output layer has two neurons that classify the final output as either a picture of fire or no fire.
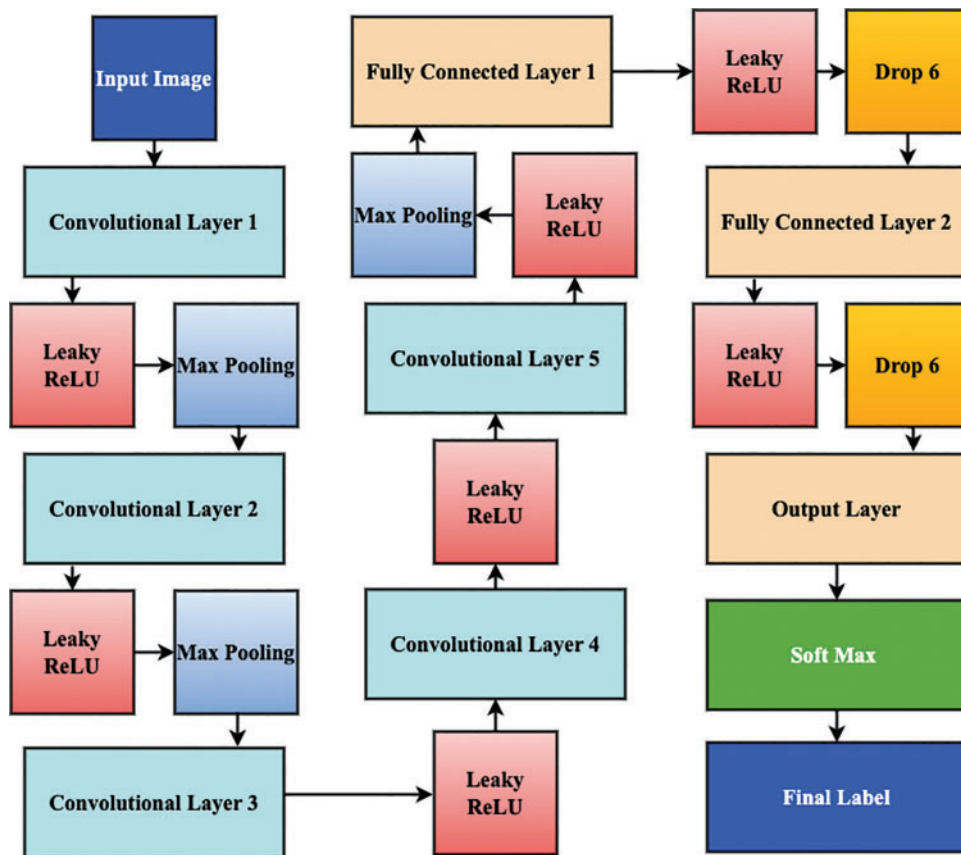
Figure 3: Convolutional neural network framework

3.4 Cloud Layer

Every second, billion of IoT devices create vast amounts of data. Researchers from all over the world are utilizing data for various objectives. The volume of data is expanding exponentially over time, making ordinary computer systems incapable of handling it. Cloud services like data storage, processing, and sharing are critical to coping with this vast volume of data. Since the previous decade, the IoT business has grown fast, and all industries are using IoT infrastructure. An increasing number

of IoT devices are being installed by companies, which are continuously generating data. The cloud layer maintains the generated data for further analysis and processing. Monitoring architecture is paramount for high-speed computing systems that can analyze data in microseconds, aid in emergency detection, and safeguard the environment.

3.5 Service Layer

For real-time surveillance, SSA constantly updates the data in the cloud, and all service centers get updates from the cloud. When an emergency occurs, the service provider receives the alert and sends a service provider to the affected area. A brief event-based service selection scenario is shown in Fig. 4.
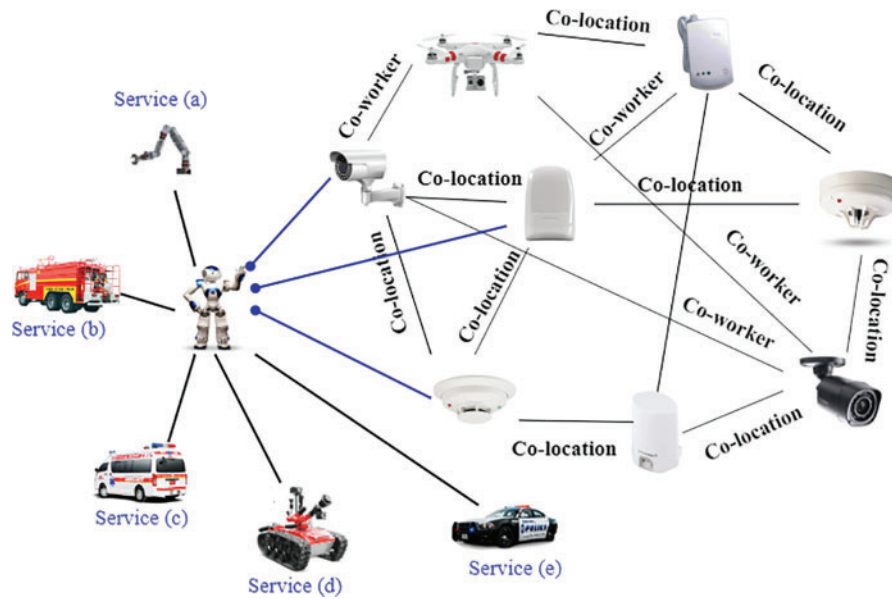

Figure 4: Event-based service search scenario by exploiting social IoT in industrial environment

In the suggested architecture; we recommend both human and artificial intelligence service providers (e.g., robots). Control room and service providers contacted by ISSA as well as it activates robots and industrial actuators simultaneously and generates an alert in case of an emergency. The suggested paradigm intends to prevent environmental degradation without extreme measures. Leaving the building during an emergency alert is usually recommended to protect ourselves. Therefore, determining the afflicted area and circumstances is usually a challenging attempt. The SIoT paradigm allows IoT devices to connect and interact with one another to detect an event and ascertain the precise position and the affected place. UAVs, actuators, and industrial robots play a pivotal role in dealing with this problem, as well as UAVs generates constant visual reports, which minimizes the load of service providers and save the lives of service providers from hazards.

4 Experimental Work

To evaluate the proposed model, we performed an experiment using the Foggia video fire data set [44], the Chino smoke data set [45], and additional gas and heat datasets [46]. In the field of fire detection, the Foggia video fire dataset [44] is a frequently used benchmark dataset. It is made up of video clips that were taken in a variety of settings with various fire scenarios. A realistic representation

of fire incidents, including various fire sizes, types, and intensities, is provided by the dataset. The training and assessment of fire detection models are made possible by the annotation of each video sequence with ground truth labels indicating the presence of fire. The Chino smoke dataset [45] is dedicated to the detection of smoke. It includes pictures that were taken under various smoke-presence conditions. The dataset offers a wide variety of smoke patterns, densities, and lighting situations that mimic real-world smoke scenarios. The Chino dataset is annotated with ground truth labels for smoke presence, much like the Foggia dataset, making it easier to train and test smoke detection models. To improve the model's capacity to identify gas leaks and unusual heat patterns, gas and heat datasets [46] were added. These datasets most likely include temperature readings, sensor readings, or other pertinent information gathered from industrial settings.

We implemented the proposed CNN model in python programming language using TensorFlow and Keras libraries. Following are the specifications of the machine we used for training the model: Intel® Core i5-3570 CPU @ 3.40 GHz or 3.80 GHz, with a Windows operating system and GTX 1080 graphics card. Table 1 presents the system specifications in tabular form.

Table 1: Detailed environment parameters used for implementation

Sr. No.	Name	Experiment environment parameter
1	Operating system	Windows 10
2	CPU	Intel Core i5-3570 CPU @ 3.40 GHZ 3.80 GHz
3	GPU	NVIDIA GeForce GTX1080
4	Memory	24 GB
5	Development tool	Python 3.6
6	Library	TensorFlow, Keras

We have trained our CNN model with 70% of the data, while the remaining 30% is used for validating and testing the model. We used 43,376 photos for training, 19,251 for validation, and the remaining 1,543 for testing, from a total of 64,170 images. The training consisted of 80,000 iterations with 128 batch sizes. Initially, the learning rate was 0.01, but because of the step-decay learning process, it decreases by a factor of 0.5 every 1000 iterations. After 40,000 iterations, our model's learning rate is locked at 0.001 percent. In addition, we set the momentum to 0.9. Accuracy, Precision, and Recall are just a few of the parameters that CNN-based models utilize to gauge their performance. Recall indicates how accurate predictions were made to the actual data, whereas Precision reflects the percentage of accurate predictions. Precision and Recall are calculated using the following equations, see Eqs. (7) and (8):

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \tag{7}$$

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \tag{8}$$

- True Positive = True proposals predicted as true labeled class.
- True Negative = Background proposal predicted as background.
- False Positive = Background proposal predicted as true labeled class.
- False Negative = True Proposals predicted as background.

5 Results and Discussion

Fire images and smoke sensor datasets were used for training and testing the model. The dataset contains fire, no-fire, and smoke images. We used 70% of the data for training the model, 20% for validation, and 10% for evaluation. Two libraries, TensorFlow and Keras, were used to implement the model in Python. While implementing the model, we used the Leaky ReLu activation function and the step decay algorithm for training with the Adam optimizer. Initially, we applied data preprocessing and resized our images to fit the model input. Similarly, we enforced data preprocessing for sensor data, which went through various filters like normalization, redundancy filtering, irrelevance filtering, and data cleaning. The proposed CNN and other existing models used the Leaky ReLu activation function in their hidden layers.

The detailed analysis of the models efficiency was carried out by visualization of the learning curves of all models and the combined learning curve after adding the last fully connected layer. The learning curve helps to analyze the model's performance over numerous epochs of training data. Thorough analysis of the learning curve enabled us to deduce whether the model is learning new knowledge from the input or merely memorizing it. It is evident that, high learning-rate, bias, and the learning-curve may be skewed in training and testing that indicates model's incompetency to learn from its errors. Likewise, a big difference in errors (training and testing errors) reveals higher variation. The model needs to improve in both directions, resulting in erroneous generalizations. Overfitting is a phenomenon that occurs when the training error is very less but presents more testing error. It shows that the model remembers rather than learning. As a result, it is challenging to extrapolate from the model in these cases. In addition, overfitting is avoidable by using the dropout approach and terminating learning early.

By training and testing, we calculated the accuracy of the proposed and existing models. Fig. 5a shows the training loss, while Fig. 5b accuracy curves for the CNN models. Fig. 6 shows the CNN model's Receiver Operating Characteristics (ROC) curve. The model uses 10% of the data that contains fire, smoke, no fire, and blurry images for testing. In our test dataset, some scenic views contain multiple substances that look like fire and smoke but are not actual fire and smoke. Therefore other-based techniques create a false alarm on the images that give the impression of fire. Our model performed very efficiently during testing. Fig. 7 presents the performance of the proposed architecture on testing images.
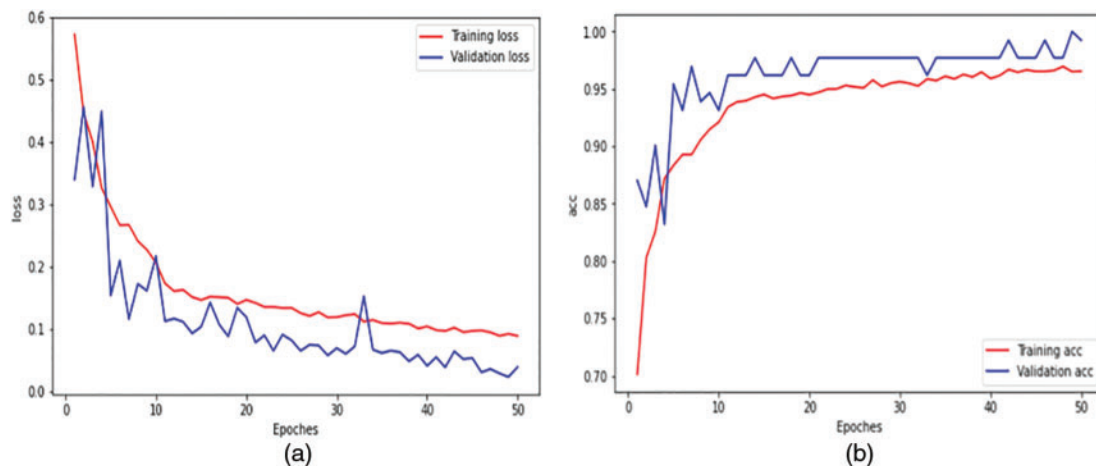

Figure 5: Training and validation loss *vs.* accuracy of proposed CNN model

Figure 6: ROC curve of proposed CNN model

Figure 7: Model performance on testing data

We compared the novel CNN architecture with existing advanced fire detection methods. Our CNN architecture is a shallow network as well as contains less number of trainable parameters, which highly contribute to its strength: only 7.45 MB (646,818) space utilized on disk. It is important to note that other, higher-performing fire detection solutions can be found in published works. The degree to which something performs better depends on the tools, and the data set used to train it. To give

just a few examples, already advanced CNN models have a capability of detection around 4–5 frames per second with utilization of more space on the disk with low-cost embedded hardware. However, the proposed model's superior fire detection capabilities stem from the fact that it was trained on a much more varied dataset and was created expressly for this purpose. At up to 24 frames per second, the Proposed CNN model's real-time fire detection feature is nearly as fast as human visual cognition thanks to this powerful combination. Fig. 8 shows the comparative analysis of the proposed model with state-of-the-art methods using the loss curve. Meanwhile, Fig. 9 compares the Proposed CNN model with existing models on Accuracy and Precision. The figure clearly shows that the proposed model's performance is better than others. Table 2 compares the proposed model with existing models using different matrices, i.e., Accuracy, Precision, Recall, False Positive, and False Negative.
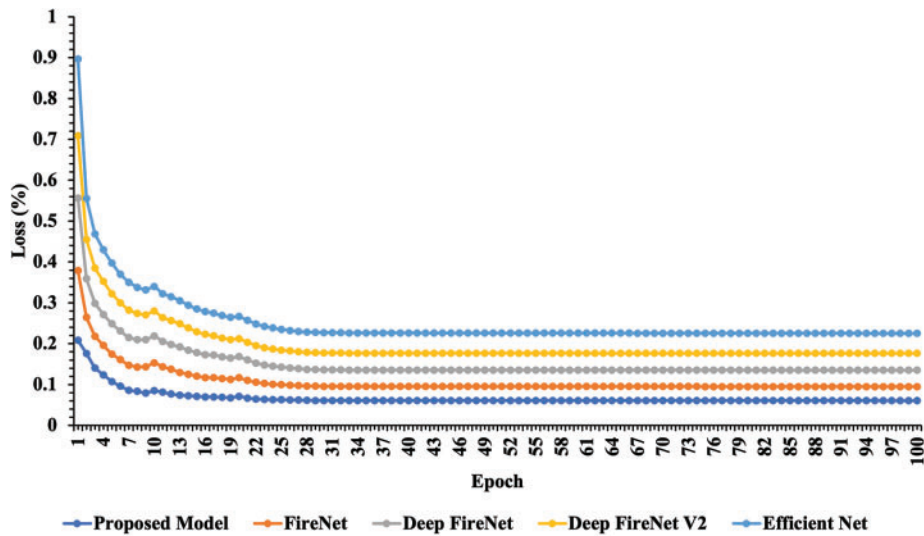

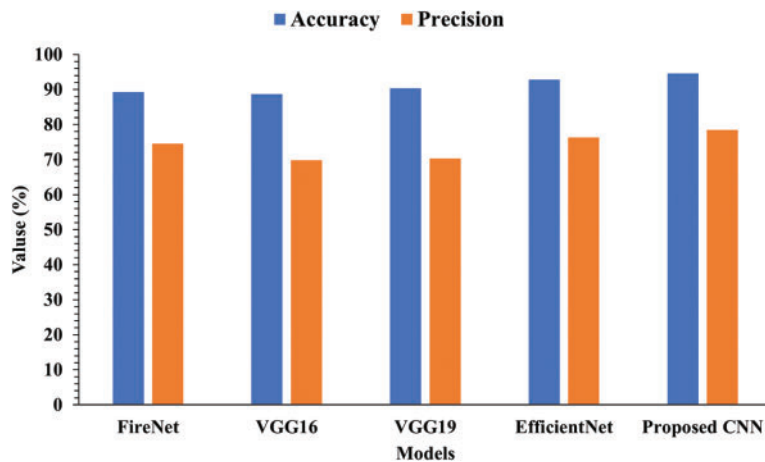**Figure 8:** Training loss comparison of proposed and existing models

Figure 9: Accuracy and precision comparison of proposed and existing models

Table 2: Comparison of different state-of-the-art methods using accuracy, precision, recall, false positive and false negative

Model	Accuracy	Precision	Recall	False positive	False negative
FireNet	91.56%	0.80	0.84	6.93%	4.03%
Deep FireNet	92.86%	0.81	0.86	7.04%	3.12%
Deep FireNet V2	93.46%	0.82	0.89	7.54%	3.29%
EfficientNet	93.67%	0.84	0.92	7.69%	3.42%
Proposed model	**94.50%**	**0.86**	**0.97**	**8.87%**	**2.12%**

As we can see from Table 2, the proposed model achieved the highest accuracy (94.5%) and recall (0.97) among all the models evaluated. It also had a relatively low false positive rate (8.87%) and a false negative rate (2.12%). The results suggest that the proposed model is more effective at detecting fires than the other models evaluated in this study.

6 Conclusion and Future Scope

The use of industrial IoT technology is crucial for disaster prevention in industrial settings, but its effectiveness is limited. Industrial accidents such as fires, toxic gas leaks, chemical spills, and unsafe working conditions can result in significant financial losses and loss of human life. Early detection and swift action are critical to mitigating the impact of such disasters. This study presents an innovative approach using an "industrial smart social agent" (ISSA) that utilizes both IIoT and modern AI techniques to enhance surveillance and detect fire hazards. The proposed CNN-based model, implemented in Python, outperforms four existing fire detection models in detecting fires. Upon detection of the fire, ISSA triggers an alarm and sends alerts to relevant authorities for swift action. The proposed system effectively detects events early, minimizes financial and human losses, and outperforms existing state-of-the-art methods.

Author Contributions: Dr. Abdul Rehman proposed the idea, performed all the experimental work, and wrote the manuscript. Dr. Dongsun Kim has supervised this research work, refined the ideas in several meetings, and managed the funding. Prof. Anand Paul reviewed the paper and extensively helped revise the manuscript.

Availability of Data and Materials: The data used in this paper can be requested from the corresponding author upon request.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] A. Rehman, M. M. Rathore, A. Paul, F. Saeed and R. W. Ahmad, "Vehicular traffic optimisation and even distribution using ant colony in smart city environment," *IET Intelligent Transport Systems*, vol. 12, no. 7, pp. 594–601, 2018.

[2] E. Jang, S. Karshiev, A. Rehman and J. Kim, "Quickest path selection towards the destination in urban environment," in *Proc. Eleventh Int. Conf. on Ubiquitous and Future Networks*, Zagreb, Croatia, pp. 623–628, 2019.

[3] S. Din, A. Paul and A. Rehman, "5G-enabled Hierarchical architecture for software-defined intelligent transportation system," *Computer Networks*, vol. 150, no. 1, pp. 81–89, 2019.

[4] A. Rehman, A. Paul, M. J. Gul, W. H. Hong and H. Seo, "Exploiting small world problems in a siot environment," *Energies*, vol. 11, no. 8, pp. 1–18, 2018.

[5] A. Rehman, A. Paul, A. U. Rehman, F. Amin, R. M. Asif *et al.,* "An efficient friendship selection mechanism for an individual's small world in social internet of things," in *Proc. 2020 Int. Conf. on Engineering and Emerging Technologies*, Lahore, Pakistan, pp. 1–6, 2020.

[6] F. Amin, R. Abbasi, A. Rehman and G. S. Choi, "An advanced algorithm for higher network navigation in social internet of things using small-world networks," *Sensors*, vol. 19, no. 9, pp. 20, 2019.

[7] A. Rehman, A. Paul, M. A. Yaqub and M. M. U. Rathore, "Trustworthy intelligent industrial monitoring architecture for early event detection by exploiting social IoT," in *Proc. 35th Annual ACM Symp. on Applied Computing*, Brno, Czech Republic, pp. 2163–2169, 2020.

[8] T. Sher, A. Rehman and D. Kim, "COVID-19 outbreak prediction by using machine learning algorithms," *Computers, Materials & Continua*, vol. 74, no. 1, pp. 1561–1574, 2023.

[9] A. Rehman, A. Paul and A. Ahmad, "A query based information search in an individual's small world of social internet of things," *Computer Communications*, vol. 163, no. 1, pp. 176–185, 2020.

[10] A. Rehman, A. Paul, A. Ahmad and G. Jeon, "A novel class based searching algorithm in small world internet of drone network," *Computer Communications*, vol. 157, no. 1, pp. 329–335, 2020.

[11] X. Li, Z. Chen, Q. M. J. Wu and C. Liu, "3D parallel fully convolutional networks for real-time video wildfire smoke detection," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 1, pp. 89–103, 2020.

[12] P. Li and W. Zhao, "Image fire detection algorithms based on convolutional neural networks," *Case Studies in Thermal Engineering*, vol. 19, no. 1, pp. 1–11, 2020.

[13] Y. Cao, F. Yang, Q. Tang and X. Lu, "An attention enhanced bidirectional LSTM for early forest fire smoke recognition," *IEEE Access*, vol. 7, no. 1, pp. 154732–154742, 2019.

[14] B. Bir, "Wildfires, forest fires around world in 2020," https://www.aa.com.tr/en/environment/wildfires-forest-fires-around-world-in-2020/2088198

[15] F. Saeed, A. Paul, P. Karthigaikumar and A. Nayyar, "Convolutional neural network based early fire detection," *Multimedia Tools and Applications*, vol. 79, no. 13–14, pp. 9083–9099, 2020.

[16] X. Jia, Q. Feng, T. Fan and Q. Lei, "RFID technology and its applications in Internet of Things (IoT)," in *Proc. 2012 2nd Int. Conf. on Consumer Electronics, Communications and Networks*, Yichang, China, pp. 1882–1885, 2012.

[17] C. Sun, "Application of RFID technology for logistics on Internet of Things," *AASRIProcedia*, vol. 1, no. 1, pp. 106–111, 2012.

[18] A. Rehman, S. Din, A. Paul and W. Ahmad, "An algorithm for alleviating the effect of hotspot on throughput in wireless sensor networks," in *Proc. 2017 IEEE 42nd Conf. on Local Computer Networks Workshops*, Singapore, pp. 170–174, 2017.

[19] P. Lade, R. Ghosh and S. Srinivasan, "Manufacturing analytics and industrial Internet of Things," *IEEE Intelligent Systems*, vol. 32, no. 3, pp. 74–79, 2017.

[20] L. Maglaras, L. Shu, A. Maglaras, J. Jiang, H. Janicke *et al.,* "Editorial: Industrial Internet of Things (I2oT)," *Mobile Networks and Applications*, vol. 23, no. 4, pp. 806–808, 2018.

[21] W. He and L. Da Xu, "Integration of distributed enterprise applications: A survey," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 1–8, 2014.

[22] D. Dujovne, T. Watteyne, X. Vilajosana and P. Thubert, "6TiSCH: Deterministic IP-enabled industrial Internet (of Things)," *IEEE Communications Magazine*, vol. 52, no. 12, pp. 36–41, 2014.

[23] A. U. Rehman, A. Jiang, A. Rehman and A. Paul, "Weighted based trustworthiness ranking in social Internet of Things by using soft set theory," in *Proc. 2019 IEEE 5th Int. Conf. on Computer and Communications*, Chengdu, China, pp. 1644–1648, 2019.

[24] A. U. Rehman, A. Jiang, A. Rehman, A. Paul, S. Din *et al.,* "Identification and role of opinion leaders in information diffusion for online discussion network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 1, no. 1, pp. 1–13, 2020.

[25] D. Bandyopadhyay and J. Sen, "Internet of Things: Applications and challenges in technology and standardization," *Wireless Personal Communications*, vol. 58, no. 1, pp. 49–69, 2011.

[26] Y. C. Zhang and J. Yu, "A study on the fire IoT development strategy," *Procedia Engineering*, vol. 52, no. 1, pp. 314–319, 2013.

[27] A. U. Rehman, R. A. Naqvi, A. Rehman, A. Paul, M. T. Sadiq *et al.,* "A trustworthy SIoT aware mechanism as an enabler for citizen services in smart cities," *Electronics*, vol. 9, no. 6, pp. 1–19, 2020.

[28] 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.

[29] S. Mohdiwale, M. Sahu, G. R. Sinha and V. Bhateja, "Statistical wavelets with harmony search-based optimal feature selection of EEG signals for motor imagery classification," *IEEE Sensors Journal*, vol. 21, no. 13, pp. 14263–14271, 2021.

[30] S. K. Khare, V. Bajaj and G. R. Sinha, "Adaptive tunable Q wavelet transform-based emotion identification," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 12, pp. 9609–9617, 2020.

[31] B. Kim and J. Lee, "A video-based fire detection using deep learning models," *Applied Sciences*, vol. 9, no. 14, pp. 1–19, 2019.

[32] 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.

[33] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[34] S. Anwar, K. Hwang and W. Sung, "Fixed point optimization of deep convolutional neural networks for object recognition," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, South Brisbane, QLD, Australia, pp. 1131–1135, 2015.

[35] T. H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng *et al.,* "PCANet: A simple deep learning baseline for image classification?" *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5017–5032, 2015.

[36] B. Jiang, J. Yang, Z. Lv, K. Tian, Q. Meng *et al.,* "Internet cross-media retrieval based on deep learning," *Journal of Visual Communication and Image Representation*, vol. 48, no. 1, pp. 356–366, 2017.

[37] J. Yang, B. Jiang, B. Li, K. Tian and Z. Lv, "A fast image retrieval method designed for network big data," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 5, pp. 2350–2359, 2017.

[38] W. Li, R. Zhao, T. Xiao and X. Wang, "DeepReID: Deep filter pairing neural network for person re-identification," in *Proc. 2014 IEEE Conf. on Computer Vision and Pattern Recognition*, Columbus, OH, USA, pp. 152–159, 2014.

[39] P. Luo, Y. Tian, X. Wang and X. Tang, "Switchable deep network for pedestrian detection," in *Proc. 2014 IEEE Conf. on Computer Vision and Pattern Recognition*, Columbus, OH, USA, pp. 899–905, 2014.

[40] K. Vadim, O. Maxime, C. Minsu and I. Laptev, "Contextlocnet: Context-aware deep network models for weakly supervised localization," *European Conference on Computer Vision*, vol. 1, no. 2, pp. 350–365, 2016.

[41] W. L. Zhang, R. J. Li, H. T. Deng, L. Wang, W. L. Lin *et al.,* "Deep convolutional neural networks for multi-modality isointense infant brain image segmentation," *NeuroImage*, vol. 108, no. 1, pp. 214–224, 2015.

[42] T. Çelik, H. Özkaramanlı and H. Demirel, "Fire and smoke detection without sensors: Image procecessing based approach," in *Proc. 15th European Signal Processing Conf.*, Poznan, Poland, pp. 1794–1798, 2007.

[43] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

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

[45] D. Y. T. Chino, L. P. S. Avalhais, J. F. Rodrigues and A. J. M. Traina, "BoWFire: Detection of fire in still images by integrating pixel color and texture analysis," in *Proc. 28th SIBGRAPI Conf. on Graphics, Patterns and Images*, Salvador, Brazil, pp. 95–102, 2015.

[46] F. Saeed, A. Paul, A. Rehman, W. H. Hong and H. Seo, "IoT-based intelligent modeling of smart home environment for fire prevention and safety," *Journal of Sensor and Actuator Networks*, vol. 7, no. 1, pp. 1–11, 2018.