Neurocomputing for Internet of Things: Object Recognition and Detection Strategy

Kashif Naseer Qureshi[1], Omprakash Kaiwartya[2], Gwanggil Jeon[3*], Francesco Piccialli[4*]

[1]Department of Computer Science, Bahria University, Islamabad
[2]School of Science and Technology, Nottingham Trent University, Nottingham NG11 8NS, UK
[3]Department of Embedded Systems Engineering, Incheon National University, Incheon, 22012, Korea
[4]Department of Mathematics and Applications "R. Caccioppoli", University of Naples Federico II, Italy
*gjeon@inu.ac.kr, francesco.piccialli@unina.it

Abstract

Modern and new integrated technologies have changed the traditional systems by using more advanced machine learning, artificial intelligence methods, new generation standards, and smart and intelligent devices. The new integrated networks like the Internet of Things (IoT) and 5G standards offer various benefits and services. However, these networks have suffered from multiple object detection, localization, and classification issues. Conventional Neural Networks (CNN) and their variants have been adopted for object detection, classification, and localization in IoT networks to create autonomous devices to make decisions and perform tasks without human intervention and helpful to learn in-depth features. Motivated by these facts, this paper investigates existing object detection and recognition techniques by using CNN models used in IoT networks. This paper presents a Conventional Neural Networks for 5G-Enabled Internet of Things Network (CNN-5GIoT) model for moving and static objects in IoT networks after a detailed comparison. The proposed model is evaluated with existing models to check the accuracy of real-time tracking. The proposed model is more efficient for real-time object detection and recognition than conventional methods.

Keywords: *Neural Network, Deep Learning, Object Detection, Image Processing, Localization, Classification, Convolutional Neural Network*

1. Introduction

In the last few decades, the number of smart devices and objects has been increased rapidly and offered cost-effective and intelligent services. Internet of Things (IoT) is one of the emerging technology based on new data communication standards like 5G to improve people's daily life and services [1]. Most IoT devices have video and image capabilities to obtain information about the surrounding environment for other security and monitoring purposes. Understanding collected information in the shape of video or images is an essential part and component for object detection applications [2-4]. The IoT networks are based on interconnected objects or devices for home automation, intelligent transportation systems, the internet of industries, smart grids systems, healthcare, and smart education systems. The connected objects are sensor nodes, actuators, controllers, and other wireless and wired devices. The devices collect the required information with the help of new and advanced Wi-Fi, WIMAX, Bluetooth, and 5G technologies [5]. Object detection is a technique in computer technologies based on digital image processing and computer vision. In these techniques, the objects are detected instances of semantic objects in a specific class like a chair, table, person, desk, car in digital video, and image data.

Object detection is categorized into many sub-fields such as pedestrian application, face detection, body movement detection, fix and mobile object detection, human behaviour detection, and autonomous driving. Mostly, the existing object detection literature has focused on the big object covering in images. Small object detection from the images still exists and has unsatisfactory results [6, 7]. Learning systems, especially Conventional Neural Networks (CNN), have achieved excellent object detection techniques [8]. The object detection modules are divided into three major components: region selection, feature extraction, and classification [9]. Deep learning models have been adopted to provide an efficient solution by using different algorithms. Deep learning is a useful machine learning method based on various layers of features.

The fundamental objective of these methods is to distinguish subordinate-level categories. Informative region selection refers to a complete image with multi-scale slicing to identify the distinct objects with different sizes or ratios. This strategy can find all possible positions of objects. If an unfixed number of slicing windows is used, then produces redundant windows and leads to computational burden [10]. If a fixed slicing window is used, then an unwanted region may be produced. In feature extraction, visual features can provide robust and semantic representation that is significant for recognising the different objects [11]; however, because of the diversity of backgrounds, appearances, and illumination conditions. In order to describe all kinds of objects, a manual design feature descriptor does not cover all aspects. In classification, a classifier distinguishes between different objects of classes. Representation becomes more revealing, semantic, and hierarchical for visual recognition. Inequitable learning of graphical models allows for building a high precision part-based model to diversify object classes in images.

This paper investigates deep learning methods used in IoT in different domains such as automotive systems, home automation, and smart safety. After a detailed review, we proposed a Conventional Neural Networks for 5G-Enabled Internet of Things Network (CNN-5GIoT) model for object recognition and detection. The proposed model is working with 5G and cloud-based services where users receive a message on their smart device. The proposed model is evaluated in terms of object detection and recognition. The other objectives of this paper are as follows:

- Discussed deep learning methods for object detection and recognition.
- Proposed a 5GIoT data communication model for fast and reliable data services to support the IoT networks.
- Suggested edge and cloud-based systems for better services.
- Proposed CNN based object detection and recognition model for IoT networks

The rest of the paper is organized as follows: Section 2 presents the deep learning usage for object detection and recognition in IoT networks. Section 3 presents the existing work related to deep learning for object detection in IoT networks. Section 4 presents the proposed CNN-5GIoT model for object detection. Section 5 presents the results of the proposed model with state of the art models. The paper concludes with the future direction in the last section.

2. Deep Learning for Object Detection and Recognition in IoT

Usually, deep learning models are those neural network models that have deep structures. Gradually, deep learning has been gained popularity after the arrival of graphical processing units and large-scale data such as Image Net. A deep learning model can learn a high level of representation of features to detect pedestrian detection. Switchable layers can quickly learn low-level and high-level features. Deep learning performance is better than classical machine learning in terms of accuracy and performs better when a large amount of data and powerful computing resources are available. Deep Learning usually has large layers for processing [12]. Deep learning is one of the branches of machine learning based on deep neural networks, whereas CNN is a popular neural network architecture. CNN can learn high-level representation extensively. In real-world scenarios, objects are frequently subject to occlusions and sensor noise. CNN is based on different layers and filters; filter size depends upon image and requirement [13, 14]. CNN uses hierarchical features that are based on multi-level feature representation. It can learn automatically from raw data; one of the advantages of CNN or deep learning is that it extracts feature hierarchically. Mainly CNN has three layers' convolutional layer, the pooling layer, and the fully connected layer. In the first two layers, convolutional and pooling layers' have different types of activation functions for use. Convolutional layer extract features automatically from images, also extract a high level of abstraction. The first level is responsible for low-level features, and the second layer pooling layer is responsible for reducing parameters and decreasing computation cost. A fully connected layer is responsible for classifying between distinct classes [15]. Figure 1 shows the basic CNN architecture and its layers.

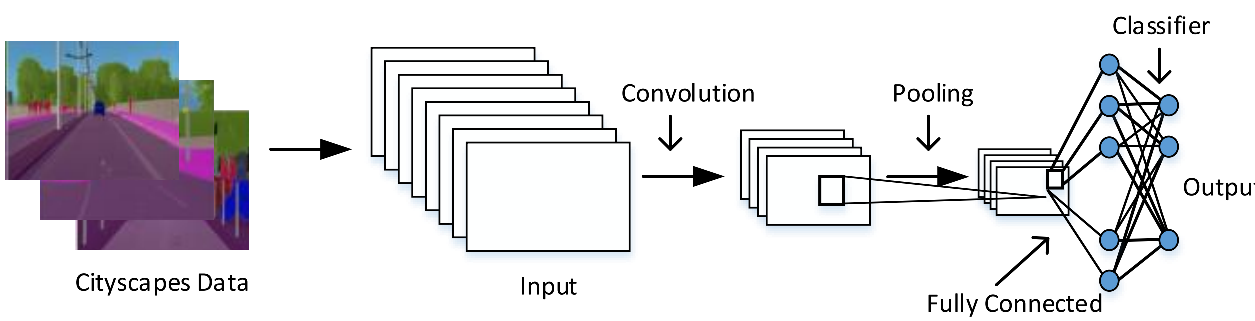

Figure 1: Convolutional Neural Network Framework

Figure 1 clearly shows that how the data is processed with the CNN framework by using different input, convolution, pooling, and classifier methods or layers and generate the high detection output.

2.1 Object Detection in Automobile Systems, Home Automation, and Smart Safety

Automation or transportation system is one of the core areas where automobiles need continuous improvement for various applications such as traffic management, object recognition, safety, and monitoring purposes. Traffic sign recognition is one of the popular areas for regulating traffic [16]. There are various areas in the automation field where machine and deep learning methods have been adopted for object detection such as traffic sign recognition, vehicle classification, driver behaviour recognition, road damage detection, vehicle attribute recognition, and license plate detection. The three layers of CNN extracted the image features and classify the image or video data. The repetition of a stack of several convolution layers and pooling layers has been adopted for feature extraction. Then, the different feature maps are applied for feature extraction from the image. The main objective of pooling is to decrease the parameters and prevent over-fitting. For example, in a car's image, any part detects the car features like four tires, four doors, and car shape. After this process, the next step is the classification where the flattened vector is used from the previous convolution layer [17, 18]. In order to train the CNN model, the feed-forward neural networks and backpropagation models have been adopted in each iteration. Different parameters are applied in CNN architecture, such as optimization and estimation and hyperparameters, which are set manually to help estimate the optimized parameters.

In face-related applications, especially face recognition, face detection is vital for several applications in IoT networks. Also, use in face expression analysis, face synthesis, and face recognition. In this task, many pixels are available for processing nearly 300 pts vs. 1000 pts.

Large faces have different structural settings such as skin colour, eye colour, distribution of different face parts. Nevertheless, a large visual difference of faces makes a challenge for diverse applications in the perspective of pose variations, illumination adjustments, and occlusions. For conventional face detection procedures, many annotations and high computational expenses are required to accomplish reasonable results [19]. Deep learning models can learn deep representation in images and later classify according to obligation. Face recognition is a broad problem for recognizing objects in videos. Steps to recognize objects divide into several parts such as detection, alignment, feature extraction, and classifying the object. In this decade, deep learning can detect faces more efficiently than humans.

Pedestrian detection is another area of object detection, such as video surveillance systems and robotics. This area has gained popularity due to security and other variabilities. However, with many features, this area is suffered from dense occlusion due to obstacles. It is observed that 70% of street images' pedestrian data occluded in at least one video frame. The existing solutions have addressed the occlusion issue by using a grouping strategy into two types, including distinct occlusion underlying variables. The first type is constructing a specific detector and needs prior information of occlusion types such as train a series of unfair classifiers for bottom-up and right-left occlusions. The second type is separated pedestrian into several parts and inferred their familiarity with latent variables [20]. Figure 2 shows a few examples of object detection and face recognition.

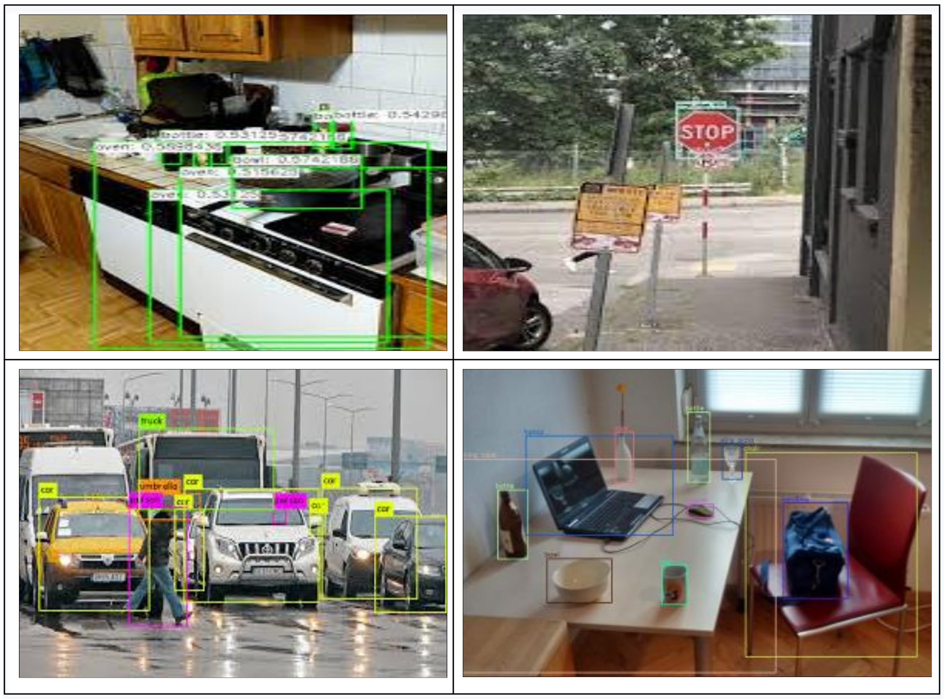

Figure 2: Object Detection using Neural Networks

Table 1 shows the well-known deep learning models for object detection and recognition with technical aspects.

Table 1: Deep Learning Models used for Object Detection

S#	Model Name	Category	Learning Model	Input Data Type	Features	Usage in IoT Applications
1	AE (Auto-Encoder)	Generative	Unsupervised	Various	Best for dimensional reduction works with unlabeled data	Emotion Diagnostic and machine faults recognition
2	RBM (Restricted Boltzmann Machine)	Generative	Unsupervised	Various	Best for dimensionally reduction and classification	Used for indoor localization and energy consumption prediction
3	DBN (Deep Belief Network)	Generative	Unsupervised	Various	Best for hierarchal feature discovery and applied greedy training of the network layer	Used for fault detection and security threats identification
4	RNN (Recurrent Neural Networks)	Discriminative	Supervised	Serial-time series	Through internal memory processes sequence of data	Used for object movement patterns and behavior detection
5	LSTM (Long Short-Term Memory)	Discriminative	Supervised	Serial-time series, long time-dependent data	Better performance in terms of data of long time lag	Used for people movement activities and mobility prediction

6	CNN (Convolutional Neural Network)	Discriminative	Supervised	2D, Image, Sound	Convolution layers provide the biggest part of computations and take fewer connections compared to DNN	Used for traffic sign detection and plant disease detection.

3. Related Work

In [21], the authors designed a data-driven object detection framework by using deep features and hierarchical object appearances. This framework uses a latent topic model algorithm and SVM to constitute the hierarchical classification ensemble. The proposed framework is explored as a generalized ability of an object in terms of its categorization and localization. The object detector is used to evaluate the detection tasks by using datasets based on thousands of images of real-world scenes. This framework's main objective is to predict the bounding boxes of objects in images and overlap predicted bounding boxes considered a true positive. The proposed framework achieved deep features and classifiers assembled for each node.

In [22], the authors presented a performance analysis of deep learning by using a single shot detector in IoT networks. This system is designed for home automation services and detected the objects based on the Model View Controller (MVC) architecture. This architecture is deployed in clouds where the users remotely monitor their homes. For data communication, message queuing telemetry transport protocol is used where IoT devices and are connected. The authors also proposed a distributed concept for load balancing. Different experiments are conducted to evaluate the proposed system performance where some devices are used, such as raspberry pi and camera for object detection based on deep learning algorithm and OpenCV library. The light intensity level and objects' distance from cameras are the parameters for system evaluation. The experiment results indicated that the system delay is very low as compared to the Raspberry Pi. Also, the accuracy of the proposed system depends on the object's distance from the camera. However, object detection accuracy depends on the object distance from the camera because of the object size.

In [23], the authors presented a method based on an unsupervised co-segmentation algorithm for image object detection. This method is using CNN to effectively distinguished the background and foreground appearance, which is distributed into regions and pixels. The CNN method classified the images and achieved semantic segmentation. This model decoded the channel and trained the standard to estimate the coding bits and estimated the decoders for accuracy and channel noise estimation. The experimental results are indicated that the limited performance of traditional image semantic segmentation methods compared to CNN methods. In this proposed method, high accuracy is achieved for image recognition and semantic segmentation.

Authors in [24], presented a fusion-based object detection model for object detections. This model is using arbitrarily oriented region CNN, and nine types of pansharpening methods are used to fuse multi-source images. The authors also used a faster region-based CNN structure for large scale satellite images. In order to generate the axially aligned bounding boxes, this framework also adopted a region proposal network and extracted the features by using pooling layers with different sizes. The selected features are used for classification to adjust the bounding boxes and predicted the inclined boxes and the abjectness and non-abjectness score. For detection results, the authors used the non-maximum supervision method. The proposed framework has achieved a better detection rate with proper validation of its effectiveness and efficiency.

Authors in [25], presented an improved adaptive genetic algorithm by using the CNN method. This model is designed for object recognition and moving objects based on acquiring the target information in a specific time series. This method also adopted the gradient descent algorithm for training and learning and determined the initial threshold and weights to avoid optimal local state. CNN's weights and thresholds are optimized by using an adaptive genetic algorithm to overcome the slow convergence issues. The experimental results indicated that the proposed model had achieved significant performance accuracy and efficiency, especially for object detection recognition. A better classification and recognition results are achieved for mobile object recognition..

Authors in [26], proposed a 5G intelligent IoT architecture for big data analysis, especially to optimize the communication channels. This architecture has adopted data mining, deep learning, and reinforcement learning. The proposed architecture is using cellular networks, cloud computing to incorporate device-to-device data communication. These networks especially integrated with 5G have produced a large volume of data where this data processing has been a challenge. In order to address this issue, this study uses three building blocks of architecture, including processing, object processor, and sensing region. The authors also applied some techniques like data mining and deep learning for data optimization. These methods have improved these network performances by analyzing the changes and the key evaluation indicators.

In [27], the authors presented the experimental study using the ANN methods to overcome the analytic challenge in 5G-based IoT networks. The sensitive data transmission in these networks always needs attention, especially data of ultra-reliable low-latency applications. Audio steganography is one of the low latency method used in wireless links. This study has used a brute force data analysis to check the presence of eventually hidden messages in a file. The experiments are based on MP3 files to detect secret data accuracy after determining data integrity and validity. This study's main findings are finding the stego-objects in MP3 files by extracting the features like the energy of sound, mean, standard deviation, and correlation coefficient.

Table 1: Comparison of Discussed Studies

S#	**Proposed Technique**	**Technologies Used**	**Finings and Achievements**	**Limitations**
1	Data-driven object detection framework [21]	SVM, Deep learning	Predict the bounding boxes of objects	Unsatisfactory prediction accuracy
2	Single-shot detector method [22]	Deep learning, Cloud Computing	Better results in terms of delay as compared to Raspberry Pi	Depends on the object distance from the camera because of the object size

3	Unsupervised co-segmentation algorithm [23]	CNN, Unsupervised co-segmentation Algorithm	High accuracy is achieved	Overhead in processing
4	Fusion-based object detection model [24]	CNN, Satellite Images.	Achieved a better detection rate	Limited in scope
5	Improved adaptive genetic algorithm [25]	CNN, Threshold and Weights Methods	Achieved significant performance in terms of accuracy and efficiency	Suffered from high mobility objects detection
6	5G I-IoT [26]	5G, Deep Learning, Data Mining	Achieved QoS parameters	Other data preprocessing methods are neglected
7	Evolutionary Detection Accuracy of Secret Data Method [27]	5G, VoIP, Audio MP3	Achieved accuracy	Limited scope because only focused on sound data

3.1 Discussion

After a detailed discussion about object recognition and CNN usage for feature extraction and provide accuracy and efficiency, it is observed that CCN has better performance features and choices. However, the discussed models have some limitations where these methods have ignored the image background, definition, and noise factors that have impacts, especially in IoT object recognition. Some of the methods do not consider accuracy and training time constraints and because of these not showing progressive results. Object recognition needs more hidden layers means more weight and consequently, more iteration. In order to minimize time complexity, decomposition of tensor will be needed. Besides rapid development in this trending field, some issues are still left. There are still some problems left, such as detecting small objects from the perspective of localization accuracy. Network modification is required to handle problems like occlusion. Multitask joint optimization is studied by several researchers. Other than this, many subtasks are important for object detection such as multi-object detection, instance segmentation, and militia person posture estimation. Scale Adaption is another crucial phase in the object detection process because like pedestrian detection objects exist in different scales. To train scale-invariant harmful sample mining, reverse connection, subcategory all is fruitful. Scale adaptive detectors combine knowledge graph, cascade network, and attentional mechanism.

4. Proposed CNN-5GIoT Model

The proposed CNN-5GIoT model consists of three modules including the user and IoT applications module, object detection and finding module, and cloud and edge module for data processing. These three layers have their responsibilities and systems. Figure 3 shows the three layers' model.

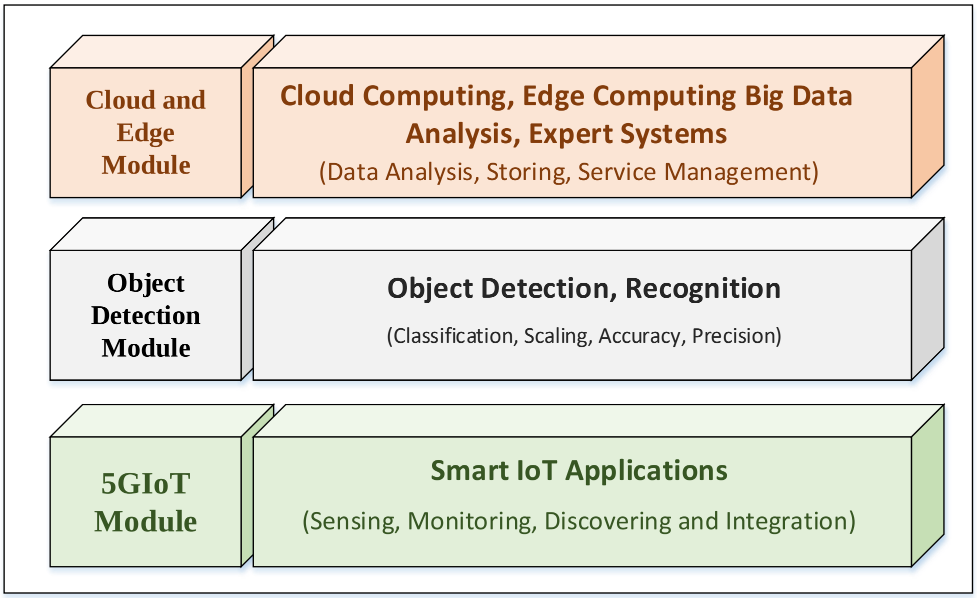

Figure 3: Proposed CNN-5GIoT Three Layers Model

4.1 5GIoT Services and User Module

The first module provides all types of applications and services for IoT smart devices. This layer also provides all services related to automobile systems, home automation, smart safety, and smart home management. Users have smart mobile phones and devices connected with 5G cellular services with more enhanced connectivity [28, 29]. The 5G services are flexible and scalable and connected with resilient cloud-native core networks with end-to-end support. The 5G provides features to support time-sensitive communication with enhanced location services [30]. IoT networks are based on three basic components, including hardware components, middleware, and presentation services. On the hardware side, smart sensor nodes, smart mobile devices, and embedded communication devices are connected to make fast and efficient data communication. The middleware and presentation services are based on visualization tools, data storage, and analysis handling services. The IEEE 802.11 and 802.15.4 standards and protocols are used in IoT networks. The 5G technologies address the most of previous challenges. They offer higher data rates, large bandwidth, low end-to-end latency, cost-effective services, affordable computational capabilities, and device intelligence services.

The proposed 5G architecture for IoT provides a wide variety of scalable networks, virtualization and network densification, radio access control, resource allocation facilities. The proposed architecture uses HetNet, Multiplexing, Massive MIMO, MmWAVE Spectrum, and Multi-RAT features. These technologies are providing massive connectivity and high date-rates. MmWAVE provides high band frequency by using big chunks of bandwidth. Flexible time-frequency multiplexing is another feature of agile 5G networks. The MIMO techniques are also adopted to increase high directional mmWAVE communication. The HetNet standard of LTE/LTE Advanced

(LTE-A) technology provides spectral efficiency by reusing the spectrum tightly with low uplink and down-link power transmission. HetNet also provides inter-cell interference coordination and able to handle massive traffic and large node density [31]. Figure 4 shows the proposed 5G-IoT network architecture. The proposed architecture detail technical specifications are presented in Table 2 with a description.

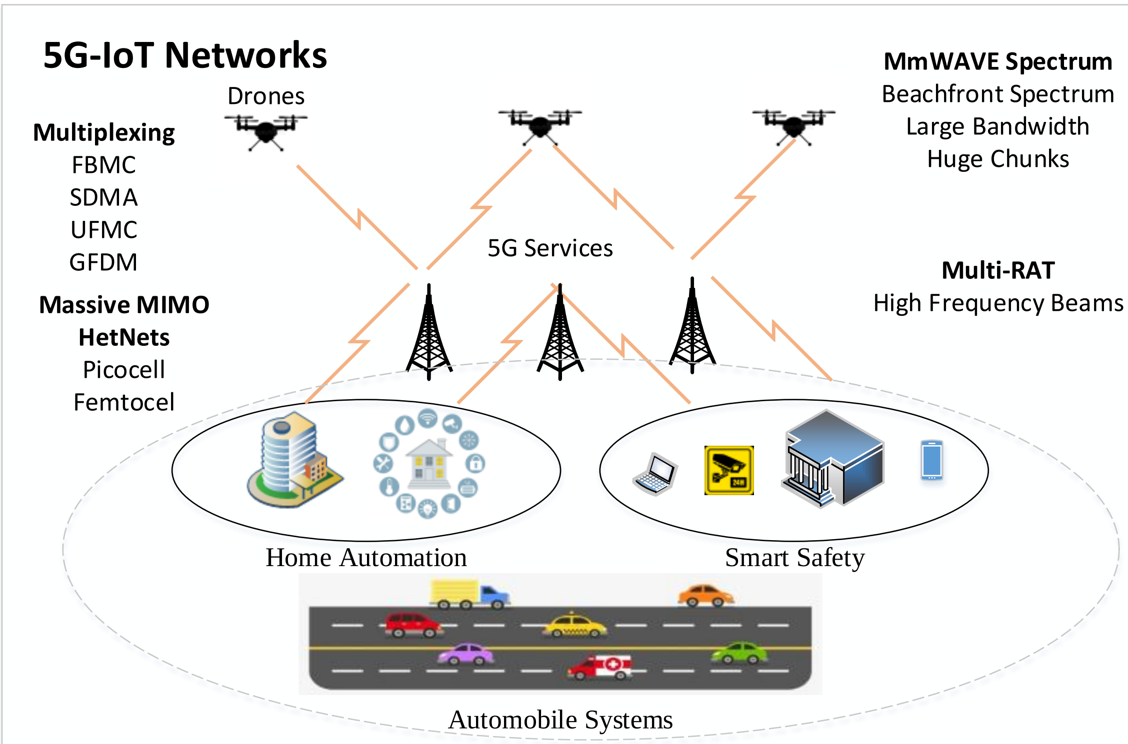

Figure 4: Proposed 5G-IoT Network Architecture

Table 2: Technical Specifications and Description of Proposed Architecture

S#	Technical Specifications	Description
1	mmWAVE	This offers a vast spectrum in-network by suing directional interface for spatial capabilities
2	Congestion Avoidance	This feature is useful to reduce congestion by using advanced methods
3	Energy Management	C-RAN offers energy management capabilities
4	Time-Frequency	New formats provide time-frequency packaging
5	Synchronism	This feature offers synchronism and nonorthogonality
6	OFDM	Offers potential alternatives to OFDM
7	MAC Functionalities	New and advanced MAC functionalities in 5G networks
8	Multiplexing	Provide multiplexing for variable delay spreads.
9	Low Latency	Fulfill the low latency requirements

10	Multi-tenancy	Offers multi-tenancy methods
11	Information-Centric	Provides information-centric framework
12	Connectivity	Unlicensed types of Wi-Fi and D2D communication

We analyze the proposed 5GIoT model to check its performance during data communication in the IoT network. We check the scalability we consider that T_m is the minimum throughput of every device and M_m denotes the mobile device. If x is the percentage of IoT and D is the maximum data rate, then we evaluate the number of IoT devices by using Equation 1.

$$5GIoT = \frac{(\frac{x}{100}) \times (D)}{M_{min}} \qquad (1)$$

Uplink and downlink delay is used to evaluate the round trip latency of the proposed system by using the hybrid automatic repeat request re-transmission delay. The total round-trip latency is calculated as shown in Equation 2.

$$Latency = Uplink_D + Downlink_D \qquad (2)$$

$$Uplink_D = \frac{TotalUplink + Loss \times AR_{uplink}}{N+1} \quad (3)$$

$$Downlink_D = \frac{TotalDownlink + Loss \times AR_{Downlink}}{N+1} \quad (4)$$

$$AR_{Uplink} = Loss \times TotalUplink + Loss \times TotalUplink^2 \ldots \ldots + Loss \times TotalUplink^N \qquad (5)$$

$$AR_{Downlink} = Loss \times TotalDownlink + Loss \times TotalDownlink^2 \ldots \ldots + Loss \times TotalDownlink^N \qquad (5)$$

Where D denotes the delay of uplink and downlink and AR is using for automatic repeat request for calculating the average of uplink and downlink without any loss where N is the total number of re-transmission.

4.2 Object Detection and Finding Module

In this module, we proposed an object detection architecture by using CNN methods. This architecture provides a feasible mechanism to detect the objects and categorized and localized the IoT network objects. Object detection is one of the complex processes due to image ambiguities in inter-class appearance. We proposed a CNN based object detection and module. Since IoT networks have multiple objects such as different devices, vehicles on roads, buildings, towers, drones, and computers, most object detection methods are based on single object detection.

Also, most of the object detection methods are using low-resolution images, and the size of the testing and training samples are fixed, which is not practical for real scenes. The proposed model contains three layers, including input, pooling, and output layer. Figure 5 shows the complete flow chart of the proposed model..

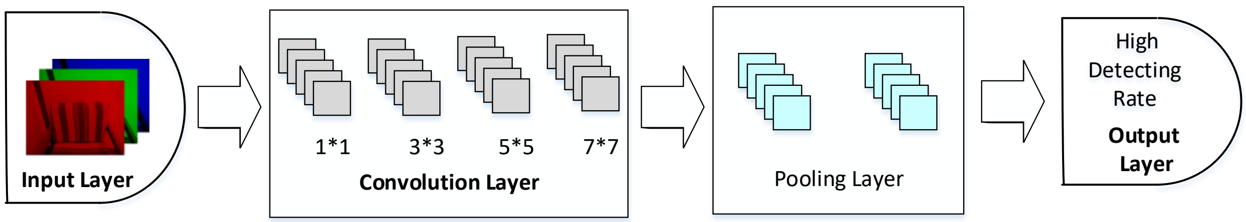

Figure 5: Proposed CNN based Object Detection Model

As shown in Figure 5, the input layer is processing with the convolution layer and then forward to the pooling layer for a high detection rate. The IoT network objects the original image is separated into three colours (red, green, and blue) channels. All channels are preprocessed, and colour values of pixels are reflected 0 and 1. This method is useful to reduce the oversize issue. The three colour channels are further connected with four sub-channels. The channels are regarded as a linear transformation of the original image to prevent information loss from parameter sharing. Equation 1 shows the feature map from the conventional layer.

$$a(x) = af(x * y + d) = af(\sum x * y + d) \qquad (1)$$

Where the af denotes the activation function with different types and y is the input matrix and x and d are weight vectors and bias. The number of channels except the first kernel is modified for recognition accuracy results. The initial weight value is set from Xavier's initialization due to its better performance [32]. The weight distribution range is set as in Equation 2.

$$\left[-\sqrt{\frac{6}{l+m}, \frac{6}{l+m}}\right] \qquad (2)$$

Where the l is used for input data and m is for output data dimension. The initial weights are generated randomly from the 1024 range.

In the pooling layer, the max-pooling is used with a 2*2 scanning window of stride 2 as no maximum suppression. All other positions are set to 0 where max value is excluded and they remove positions to reduce the map dimensions. For backpropagation processing, all the positions are recorded. The pooling layer is a direct relation with the output layer. For the loss function, the basic Euclidean metric is used as shown in Equation 3.

$$Loss(Y_{ij}) = \frac{1}{2}(x - y_{ij})^2 \qquad (3)$$

The backpropagation algorithm is used for training and we adopted the mini-batch strategy for train the network [33].

For the dataset, we used two datasets PASCAL VOC 2007 and 2012 [34] for the object detection task. This dataset is based on large scale data and aerial images. This dataset has around 2806 images collected from different networks. Another reason to adopt this dataset is that because it has multiple object images such as vehicles for smart automation systems, and building and homes related to IoT network infrastructures. The main objective is to find and predict the bounding box overlap with a true positive rate.

4.3 Cloud and Edge Module

Cloud is one of the prominent technology with significant influence to support the IoT networks. Cloud means a group of centralized devices connected to serve a different number of clients according to their needs with the assistance of virtualization. Cloud computing can also be defined as open, on-need, distributed accessibility to the resource pool, and network infrastructure. These storage cloud services can be implemented rapidly with minimal managerial activity. Cloud computing offers an infrastructure technology, shared storage, service computing, and decentralized computing [35]. Edge computing is another well-suited support to overcome the cloud computing load by offering the devices on edge. The 5G enables networks to support IoT networks' cloud, and edge systems were augmented reality, localization of objects, fast data communication, and big data analysis have great workload. We proposed a cloud and edge computing model to support the IoT networks, especially for object detection and recognition systems.

5. Results and Discussion

This section is divided into two sections including results related to 5GIoT networks and results of the CNN model for object detection and recognition. Different experiments detail discusses to check the performance of the proposed CNN-5GIoT model.

5.1 5GIoT Model Results

To evaluate the proposed 5GIoT network's performance, we used the ns-3 (mmWave ns-3 module) simulation to test the network performance in terms of scalability. We set the data rate for 5G as 1Gbps and for Long-Term Evolution (LTE), a standard for wireless broadband communication with a speed of 300 Mbps. The mobile devices in the IoT network data rate are 10 Kbps. Five-time retransmission is applied with 46 dBm (macro-cell) transmission power and 20 dBm set for the small cell. Figure 6 shows the scalability comparison of the proposed 5GIoT model with LTE standard.

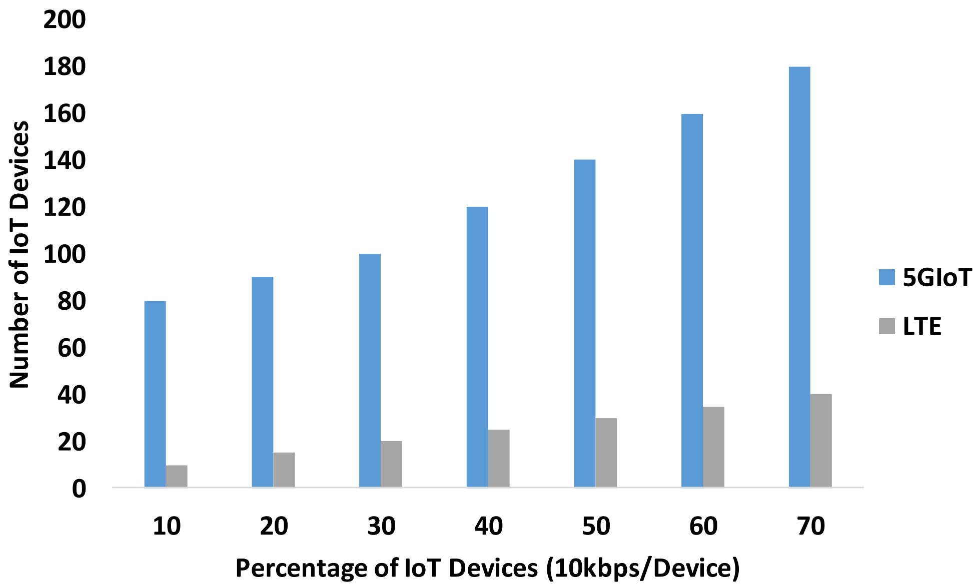

Figure 6: Scalability Comparison of LTE and proposed 5GIoT model

Figure 6 shows the varying percentage of IoT devices where it is observed that the 5GIoT model has better results of scalability. This tern shows the unprecedented device proliferation in IoT networks. This model also addresses the coexistence and heterogeneity issues in mobile to mobile communication processes. Figure 7 shows the latency results in the presence of different channel conditions.

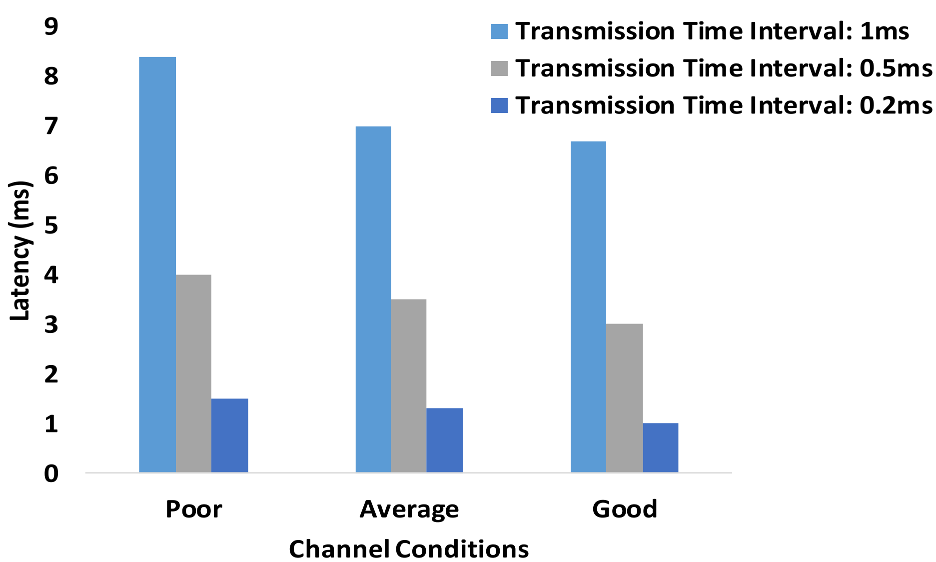

Figure 7: Latency Comparison of 5GIoT Model

Figure 7 shows the latency comparison of the 5GIoT model where we proposed a short Transmission Time Interval. For LTE, the transmission time interval is defined 1ms and almost all the communication system has adopted this standard. Recent advancements developed a shorter time from 0.5 ms to 1 ms such as for OFDM standard. This result is based on downlink data processing transmission to check the poor, average, and good channel conditions. This result indicated that the smaller transmission time interval supports the 5G networks and better options for delay-sensitive networks like IoT.

5.2 CNN Model Results

Cityscapes dataset [36] is used for experiments that focus on urban areas in real-world situations. In the first stage, we conducted two experiments we trained the detector, and in the second experiment, the knowledge transfers learning. Finally, we compared the proposed model with state of the art methods. The experiment setup requirements are shown in Table 3.

Table 3: Hardware and Software Requirements and Specifications

S#	Requirements	Description
1	Computer System	Core(TM) i7-6500U
2	Operating System	Windows 10 (64 bits)
3	Tool	Visual Studio (OpenCV 2.4.0)
4	Image Samples	50 for each class and 100 for the testing group
5	Data Sets	PASCAL VOC 2007 and 2010

The main objective of the proposed model is to detect multiple objects from one image. In the first experiment, we test a pure color background image that has ten objects.

5.1 Accuracy Analysis

In this analysis, the proposed object detection model is used to detect the corresponding accuracy labels. The transformation process leads to information loss, especially on edge regions and also causes of reducing detection accuracy. Every sample contains one object for counting the correct labels. The average no of indications is 8.5 for every image and can be indicated with the right location for its category assignment. The correct label's average number is 6.5, and in most of the images, the objects are correctly indicated in terms of their category and position. All the accuracy results show in Table 4.

Table 4: Accuracy Results of Multiple Object Detection

	Image-1	Image-2	Image-3	Image-4	Image-5	Image-6	Image-7	Image-8	Image-9
Number of Indication	8.4	8.3	8	7.6	7.2	7	8	8	9
Number of correct Labels	6.5	6.3	6.2	7	4	6	5	6	6

As shown in the above Table, that average number of the correct label is 6, and 6 objects are correctly indicated only image 5 and 7 have fewer objects due to black background or unexpected feature. The proposed model can detect multiple object detection. In the second experiment, accuracy is measured in terms of scaling where width, height, and width and height scaling perform simultaneously. The scaling levels are separated onto sixteen and set the pooling layer condition into even matrix size. For this experiment, we set around 2975 samples test. For this test, all images are resized in terms of image height and width. The acceptable value for the accuracy rate is set at 0.9. The scaling accuracy results show in Figure 8.

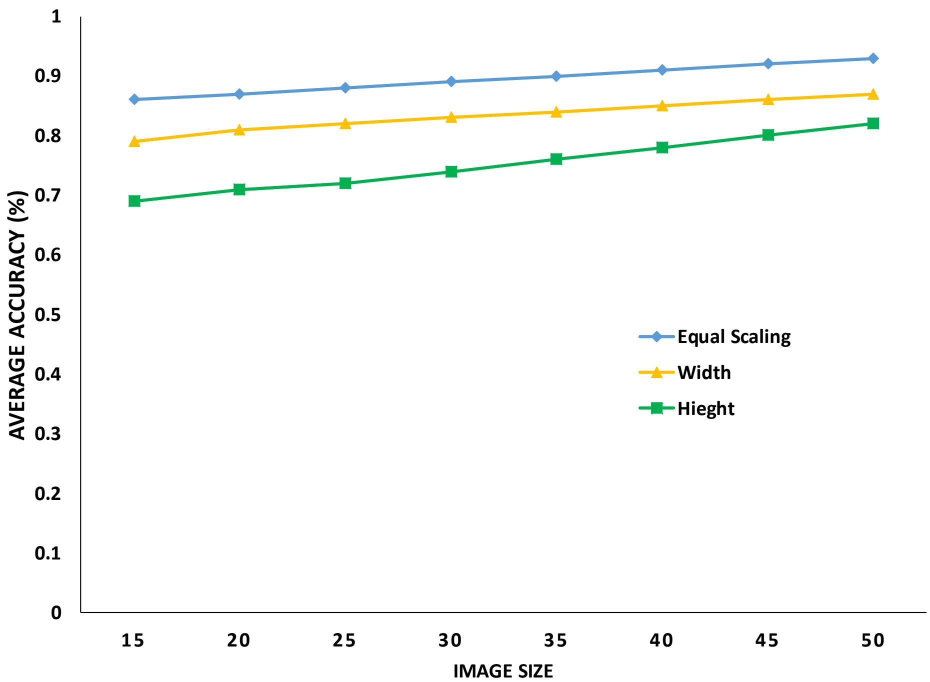

Figure 8: Accuracy Results in Terms of Different Scales

The accuracy results indicate that the proposed model has anti-scaling features. The results also showed that the larger images are stable. However, the large images are taking more processing time and also cause ignorance of small size objects. After detecting and scaling accuracy, we tested the proposed model accuracy in terms of object recognition rate with different objects and compared it with the existing fusion mode [23]. From Table 5, it can see that the proposed model has better segmentation and recognition ability as compared to the fusion model.

Table 5: Recognition Rate of Proposed Model

S#	Image Class	Accuracy Rate (%) of Model Fusion	Accuracy Rate (%) of Proposed Model
1	Road	94.57	96.5
2	Person	87.65	89.3

3	Vehicle	83.59	85.4
4	Building	87.17	89.2
5	Pole	69.43	72.1
6	Vegetation	75.19	78.3
7	Sky	63.48	67.3
8	Background	79.81	84.2

The proposed CNN-5GIoT model has high recognition accuracy. The accuracy rate of the existing fusion model degrades to recognize the object as compared to the proposed model where the multiple object detection and recognition accuracy is high. The proposed CCN-5GIoT model has better results for adopting the 5G standard and also in object recognition in IoT networks.

6. Conclusion

New advanced technologies provide fast and reliable data communication services, especially 5G networks. It is observed that most modern networks like IoT have multiple services where these networks are involved in data processing, object recognition, localization, and utilizing cloud and edge systems for better and fast services. For object recognition and detection, the more advanced machine learning, artificial intelligence methods have been adopted. This paper proposed a CNN-5GIoT model with more advanced CNN-based object detection and recognition systems and 5 G-based data communication services for IoT networks. This model offers various benefits and services. The object detection model provides accuracy and better scaling percentage for better localization and recognition. The huge and complex IoT data is processed with the help of 5G networks and edge and cloud computing models and provides fast data delivery, fast processing with more bandwidth capabilities. The proposed model has achieved better results in scalability, latency during data communication, and better accuracy and scaling percentage during object detection. In the future, we will expand these models with more services and integrated other backbone networks and datasets for learning and training.

References

[1] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future generation computer systems,* vol. 29, no. 7, pp. 1645-1660, 2013.

[2] F. Mehmood, I. Ullah, S. Ahmad, and D. Kim, "Object detection mechanism based on deep learning algorithm using embedded IoT devices for smart home appliances control in CoT," *Journal of Ambient Intelligence and Humanized Computing,* pp. 1-17, 2019.

[3] K. N. Qureshi, A. Iftikhar, S. N. Bhatti, F. Piccialli, F. Giampaolo, and G. Jeon, "Trust management and evaluation for edge intelligence in the Internet of Things," *Engineering Applications of Artificial Intelligence,* vol. 94, p. 103756, 2020.

[4] K. N. Qureshi, S. Din, G. Jeon, and F. Piccialli, "Link quality and energy utilization based preferable next hop selection routing for wireless body area networks," *Computer Communications,* vol. 149, pp. 382-392, 2020.

[5] K. N. Qureshi, S. Din, G. Jeon, and F. Piccialli, "Internet of Vehicles: Key Technologies, Network Model, Solutions and Challenges With Future Aspects," *IEEE Transactions on Intelligent Transportation Systems,* 2020.

[6] C. Chen, M.-Y. Liu, O. Tuzel, and J. Xiao, "R-CNN for small object detection," in *Asian conference on computer vision*, 2016, pp. 214-230: Springer.
[7] P. Li, X. Chen, and S. Shen, "Stereo r-cnn based 3d object detection for autonomous driving," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 7644-7652.
[8] W. Ouyang, X. Wang, X. Zeng, S. Qiu, P. Luo, Y. Tian, H. Li, S. Yang, Z. Wang, and C.-C. Loy, "Deepid-net: Deformable deep convolutional neural networks for object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 2403-2412.
[9] M. J. J. Ghrabat, G. Ma, I. Y. Maolood, S. S. Alresheedi, and Z. A. Abduljabbar, "An effective image retrieval based on optimized genetic algorithm utilized a novel SVM-based convolutional neural network classifier," *Human-centric Computing and Information Sciences,* vol. 9, no. 1, p. 31, 2019.
[10] A. Dhillon and G. K. Verma, "Convolutional neural network: a review of models, methodologies and applications to object detection," *Progress in Artificial Intelligence,* vol. 9, no. 2, pp. 85-112, 2020.
[11] D. Cao, Z. Chen, and L. Gao, "An improved object detection algorithm based on multi-scaled and deformable convolutional neural networks," *Human-centric Computing and Information Sciences,* vol. 10, pp. 1-22, 2020.
[12] D. S. W. Ting, L. R. Pasquale, L. Peng, J. P. Campbell, A. Y. Lee, R. Raman, G. S. W. Tan, L. Schmetterer, P. A. Keane, and T. Y. Wong, "Artificial intelligence and deep learning in ophthalmology," *British Journal of Ophthalmology,* vol. 103, no. 2, pp. 167-175, 2019.
[13] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2117-2125.
[14] L. Zhang, G. Zhou, Y. Han, H. Lin, and Y. Wu, "Application of Internet of Things technology and convolutional neural network model in bridge crack detection," *Ieee Access,* vol. 6, pp. 39442-39451, 2018.
[15] I. Cong, S. Choi, and M. D. Lukin, "Quantum convolutional neural networks," *Nature Physics,* vol. 15, no. 12, pp. 1273-1278, 2019.
[16] S. Hussain, M. Abualkibash, and S. Tout, "A survey of traffic sign recognition systems based on convolutional neural networks," in *2018 IEEE International Conference on Electro/Information Technology (EIT)*, 2018, pp. 0570-0573: IEEE.
[17] H. Alanazi, A. Abdullah, K. Qureshi, and A. Ismail, "Accurate and dynamic predictive model for better prediction in medicine and healthcare," *Irish Journal of Medical Science (1971),* vol. 187, no. 2, pp. 501-513, 2018.
[18] H. O. Alanazi, A. H. Abdullah, and K. N. Qureshi, "A critical review for developing accurate and dynamic predictive models using machine learning methods in medicine and health care," *Journal of medical systems,* vol. 41, no. 4, p. 69, 2017.
[19] X. Sun, P. Wu, and S. C. Hoi, "Face detection using deep learning: An improved faster RCNN approach," *Neurocomputing,* vol. 299, pp. 42-50, 2018.
[20] D. Tomè, F. Monti, L. Baroffio, L. Bondi, M. Tagliasacchi, and S. Tubaro, "Deep convolutional neural networks for pedestrian detection," *Signal processing: image communication,* vol. 47, pp. 482-489, 2016.

[21] B. Lee, E. Erdenee, S. Jin, and P. K. Rhee, "Efficient object detection using convolutional neural network-based hierarchical feature modeling," *Signal, Image and Video Processing,* vol. 10, no. 8, pp. 1503-1510, 2016.

[22] F. Mehmood, I. Ullah, S. Ahmad, and D. Kim, "Object detection mechanism based on deep learning algorithm using embedded IoT devices for smart home appliances control in CoT," *Journal of Ambient Intelligence and Humanized Computing,* pp. 1-17, 2019.

[23] L. Zhang, Z. Sheng, Y. Li, Q. Sun, Y. Zhao, and D. Feng, "Image object detection and semantic segmentation based on convolutional neural network," *Neural Computing and Applications,* pp. 1-10, 2019.

[24] Y. Ya, H. Pan, Z. Jing, X. Ren, and L. Qiao, "Fusion object detection of satellite imagery with arbitrary-oriented region convolutional neural network," *Aerospace Systems,* vol. 2, no. 2, pp. 163-174, 2019.

[25] S. Xiao, T. Li, and J. Wang, "Optimization methods of video images processing for mobile object recognition," *Multimedia Tools and Applications,* vol. 79, no. 25, pp. 17245-17255, 2020.

[26] D. Wang, D. Chen, B. Song, N. Guizani, X. Yu, and X. Du, "From IoT to 5G I-IoT: The next generation IoT-based intelligent algorithms and 5G technologies," *IEEE Communications Magazine,* vol. 56, no. 10, pp. 114-120, 2018.

[27] M. J. Alhaddad, M. H. Alkinani, M. S. Atoum, and A. A. Alarood, "Evolutionary detection accuracy of secret data in audio steganography for securing 5G-enabled internet of things," *Symmetry,* vol. 12, no. 12, p. 2071, 2020.

[28] A. Ghosh, A. Maeder, M. Baker, and D. Chandramouli, "5G evolution: A view on 5G cellular technology beyond 3GPP release 15," *IEEE Access,* vol. 7, pp. 127639-127651, 2019.

[29] N. Akkari and N. Dimitriou, "Mobility management solutions for 5G networks: architecture and services," *Computer Networks,* vol. 169, p. 107082, 2020.

[30] K. Shafique, B. A. Khawaja, F. Sabir, S. Qazi, and M. Mustaqim, "Internet of things (IoT) for next-generation smart systems: A review of current challenges, future trends and prospects for emerging 5G-IoT scenarios," *IEEE Access,* vol. 8, pp. 23022-23040, 2020.

[31] S. Li, L. Da Xu, and S. Zhao, "5G Internet of Things: A survey," *Journal of Industrial Information Integration,* vol. 10, pp. 1-9, 2018.

[32] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics,* 2010, pp. 249-256.

[33] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint arXiv:1609.04747,* 2016.

[34] M. Everingham, S. A. Eslami, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes challenge: A retrospective," *International journal of computer vision,* vol. 111, no. 1, pp. 98-136, 2015.

[35] W. Yu, F. Liang, X. He, W. G. Hatcher, C. Lu, J. Lin, and X. J. I. a. Yang, "A survey on the edge computing for the Internet of Things," vol. 6, pp. 6900-6919, 2017.

[36] (2020). *Cityscapes Dataset*. Available: https://www.cityscapes-dataset.com/

Dr. Kashif Naseer Qureshi received the M.S. degree in Information Technology from the Institute of Management Sciences, Peshawar, Pakistan, in 2012, and the Ph.D. degree from the University of Technology Malaysia (UTM) in 2016. He is currently an Associate Professor with Bahria University, Islamabad. He is a Cisco and Microsoft Certified Network Professional. His research interest focuses on Internet of connected Vehicles (IoV), Electronic Vehicles charging management planning & recommendation (EV), and IoT use cases implementation in Wireless Sensor Networks. He has been a reviewer for various reputable academic journals. He has many years' research experience in a number of areas, e.g. VANET, WSN, WBAN, Security, privacy-preserving data aggregation, computer forensics, and cloud security. He has published around 110 papers in international journals and conference proceedings, and served in a number of conference IPCs and journal editorial boards. He has also various projects related to routing and cyber security domain.

Omprakash Kaiwartya is currently working as a Lecturer at the School of Science & Technology, Nottingham Trent University (NTU), UK. Previously, He was a Research Associate at the Northumbria University, Newcastle, UK, and a Postdoctoral Research Fellow at the Universiti Teknologi Malaysia (UTM). He received his Ph.D. degree in Computer Science from Jawaharlal Nehru University, New Delhi, India. His research interest focuses on IoT centric future technologies for diverse domain areas including Transport, Healthcare, and Industrial Production. His recent scientific contributions are in Internet of connected Vehicles (IoV), Electronic Vehicles Charging Management (EV), Internet of Healthcare Things (IoHT), and Smart use case implementations of Sensor Networks.

Gwanggil Jeon received the B.S., M.S., and Ph.D. (summa cum laude) degrees from the Department of Electronics and Computer Engineering, Hanyang University, Seoul, Korea, in 2003, 2005, and 2008, respectively.
From 2009.09 to 2011.08, he was with the School of Information Technology and Engineering, University of Ottawa, Ottawa, ON, Canada, as a Post-Doctoral Fellow. From 2011.09 to 2012.02, he was with the Graduate School of Science and Technology, Niigata University, Niigata, Japan, as an Assistant Professor. From 2014.12 to 2015.02 and 2015.06 to 2015.07, he was a Visiting Scholar at Centre de Mathématiques et Leurs Applications (CMLA), École Normale Supérieure Paris-Saclay (ENS-Cachan), France. From 2019 to 2020, he was a Prestigious Visiting Professor at Dipartimento di Informatica, Università degli Studi di Milano Statale, Italy. He is currently a Full Professor at Xidian University, Xi'an, China and at Incheon National University, Incheon, Korea. He was a Visiting Professor at Sichuan University, China, Universitat Pompeu Fabra, Barcelona, Spain, Xinjiang University, China, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand, and University of Burgundy, Dijon, France.
Dr. Jeon is an Associate Editor of Sustainable Cities and Society, IEEE Access, Real-Time Image Processing, Journal of System Architecture, and MDPI Remote Sensing.
Dr. Jeon was a recipient of the IEEE Chester Sall Award in 2007 and the ETRI Journal Paper Award in 2008.

Francesco Piccialli (Member, IEEE) is currently Assistant Professor (tenure track) of Computer Science at the Department of Mathematics and Applications "R. Caccioppoli" (DMA) of the University of Naples Federico II (UNINA). He received a Laurea Degree (BSc+MSc) in Computer Science and a PhD in Computational and Computer sciences from the Unviersity of Naples Federico II.
He is the founder and Scientific director of the M.O.D.A.L. research group that is engaged in cutting-edge on novel methodologies, applications and services in Data Science and Machine Learning fields and their emerging application domains.
He has been involved in research and development projects in the research areas of Internet of Things, Smart Environments, Data Science, Mobile Applications. He is author of many papers (90+) in international conferences and top-level journals (IEEE, Springer, ACM and Elsevier).