

Received August 19, 2020, accepted September 1, 2020, date of publication September 4, 2020,  
date of current version September 18, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3021983

# A Three Layered Decentralized IoT Biometric Architecture for City Lockdown During COVID-19 Outbreak

MANJUR KOLHAR<sup>1</sup>, FADI AL-TURJMAN<sup>2</sup>, (Member, IEEE), ABDALLA ALAMEEN<sup>1</sup>,  
AND MOSLEH M ABUALHAJ<sup>3</sup>

<sup>1</sup>Department of Computer Science, Prince Sattam Bin Abdulaziz University, Wadi Ad-Dawasir 11990, Saudi Arabia

<sup>2</sup>Research Center for AI and IoT, Artificial Intelligence Department, Near East University, 99138 Mersin, Turkey

<sup>3</sup>Department of Networks and Information Security, Al-Ahliyya Amman University, Amman 19328, Jordan

Corresponding author: Manjur Kolhar (m.kolhar@psau.edu.sa)

**ABSTRACT** In this article, we have built a prototype of a decentralized IoT based biometric face detection framework for cities that are under lockdown during COVID-19 outbreaks. To impose restrictions on public movements, we have utilized face detection using three-layered edge computing architecture. We have built a deep learning framework of multi-task cascading to recognize the face. For the face detection proposal we have compared with the state of the art methods on various benchmarking dataset such as FDDB and WIDER FACE. Furthermore, we have also conducted various experiments on latency and face detection load on three-layer and cloud computing architectures. It shows that our proposal has an edge over cloud computing architecture.

**INDEX TERMS** Face detection, cascaded convolution neural network, edge computing, cloud computing, IoT.

## I. INTRODUCTION

It is reported that by 2020, more than 50 billion devices will be connected through radio communications. IoT network and its physical devices such as sensors, mobile devices, radio identification tag, and actuators are programmed to collect data from the user environment. IoT network, are employed in many fields including security, agriculture, utility management, smart homes and health care systems [1]–[6]. These capabilities are further extended to IoT enabled services with smart devices. Such combinations of services are more suitable to users. Because they have certain features which makes more competitive among other wireless devices, and these features are long range communications, low data rate, consumption of less energy and less cost. Smartphone enabled IoT services have become future networking paradigm and service framework in which distributed services and physical object will be arrayed together for useful information that enables intelligent services. Biometric services have been

studied and implemented for a particular services, however, such settings require privacy and security embedded in them to have controlled environment. The surveillance such as face detection in residential area greatly required specially during pandemic with high risk of community spread of contagious disease within a community. Such enriched technologies in small sized embedded physical devices enabled the growth of more advanced computing devices that are equipped with enough computational power to analyze and recognize face on their own without the intervention of Cloud computing architecture. IoT and overlaying technologies that have reached flawless level of management of devices and communication protocol into the level of perfection. The combination of these two technologies made it possible to produce three layered edge computing. Fog computing is expansion of cloud computing which allows building an infrastructure to allocate the workload to balance between two or more Cloud entities.

During the pandemic, the population will be under lockdown and restrictions on the movement of people. To impose such lockdown to curb the pandemic with the help of

The associate editor coordinating the review of this manuscript and approving it for publication was Yudong Zhang<sup>1</sup>.

technology especially IoT and physical smart devices are necessary because the virus can spread easily if we use human power to control residential areas.

The objective of this work is to develop an IoT the biometric framework which can detect and recognition of face using three-layered edge architecture. To this end, we have utilized the capacity of Multi-task Cascaded Convolutional Network architecture (MCCNN), efficiently using Raspberry Pi, for robust detection and recognition of a face. The proposed network compared with the state-of-the-art method on various datasets of challenging benchmarks, which further strengthens the proposed method to avoid the training framework from scratch. In the proposed MCCNN network, only a few features across multiple channels are used, while advanced layers for pre-training remained intact. In the proposed system, we have ported Computer Vision library to the Raspberry Pi computer. This setting allowed us to surveil people's movement necessary tools under lockdown condition. Secondly, we have implemented face detection methods on edge computing architecture using three layered architecture namely edge layer, application layer, and physical layer, and the required configuration items are installed and executed on Raspberry Pi computers. The remainder of this article is structured as follows: Division II reviews the existing state of the art methods, image processing frameworks for handling biometric data. Division III introduces the main concepts of our proposed framework. It also describes the developed three-layered architecture for processing the face data in the residential areas and at any time. Division IV describes the various experimental scenario used to validate the proposed architecture. The main outcomes of the proposed architecture are discussed. Section V concludes our research and discusses further research.

## II. RELATED WORK

During the COVID-19 pandemic outbreak, it becomes necessary for the general public to follow steps as advocated by the city councilors to contain COVID-19. There are many rules and guidelines suggested across the world, such as partially lifting of lockdown, complete lockdown, and allowing few family members to access general and essential goods or city council will take its own course of action to supply essential goods. To follow such rules and keeping an eye on each and every person becomes difficult and cumbersome. So, taking advantage of IoT based solutions becomes the norm. Hence, in this article, we have built an IoT based biometric solution to help people to follow guidelines under lockdown conditions.

IoT based technology has been timely surveyed [27]–[29]. Authors have presented a survey of IoT based technology, their security features, embedded hardware design, and competencies of their networking stack have also studied. Furthermore, IoT based required essential technologies for the IoT service infrastructure, existing trends, and their unique features. IoT based new technologies are also introduced with the help of 3GPP radio technology which incurred low power consumption and cost.

There are two regions specifically used in the face recognition and detection framework [7]–[13], which are named as complete face region and occlusion verification. Most of these proposed algorithms are rare to verify occlusion and hence occlusion is still at large. However, there are a number of researchers have tried to solve the face occlusion by using filter functions [23]. However, various temporal information along with the skin have become important features for occlusions problems [24]. Furthermore, some researchers have discovered more features in parts to recognize [25], [26]. Most of these methods have proved to handle occlusions from different regions of the human body, however, they failed to handle occlusion under severe conditions. Although, a face-based application using IoT based on various methods such as color, contour, matching based methods have contributed to face detection and recognition research area. These methods have a computational requirement when face is covered.

Face detection is one of the widely studied problems in image processing. The authors' Viola and Jones have proposed the two basic image processing algorithms on face detection using boosted cascade structure and simple features. Till now, these algorithms are being used in many applications ranging from industry and academia. However, these algorithms are suitable for normal conditions but not a good candidate for harsh conditions such as low light and occlusion of expression. These two algorithms are built using the Harr technique, and it is proven to be insufficient when faces have variations.

In [13], boosting was introduced with the AdaBoost algorithm, the AdaBoost algorithm combines classifiers to bring more accuracy to the classifiers. It is an ensemble method and works iteratively. AdaBoost combines poorly executing classifiers to form strong classifiers so as to produce high accuracy. Lately, it proved that AdaBoost won't work in multiclass cases. However, there were efforts rendered to enhance AdaBoost for multi-class [14]. Most recent research is also based on AdaBoost, however, but they are utilized facial expressions in the behavioral analysis of emotions, cognitive science, and social interactions [15], [16].

Authors [15], have proposed AdaBoost with dynamic time wrapping for facial expression recognition on boosted features. The features of the face are collected from the collection of facial expression images based on tracking facial landmarks.

In [16], they have used Viola Jones, histograms of oriented gradients, and multi-class SVM. Authors have used SVM (support vector machine) multiclass methods to train data from 5 different angle image captured from CCTV positioned from the walk through the metal detector. Authors [17], have introduces a technique using deep learning to detect a face. They have applied a convolutional neural network model along with AdaBoost.

In [18], the authors presented facial expression recognition model, it detects the emotional states in real time. Further, it allows the host machine to interact with the user and identifies the frontal face along with the region of interest (ROI),

collects features from suitable facial landmarks, and classifies the facial expressions. In [19], authors have proposed a model for breeding classifier ensembles based on clustering. However, a continuous stream of video data makes the complete system burdened large memory. Hence, they have built a framework to collect only the informative data from the memory by applying the clustering technique. Authors [20], have suggested Haar techniques to protect images apart from image detection. They have utilized the cloud resources for collecting Haar feature from privacy-preserved images from the cloud.

In [21], authors have used image processing and execution for handheld devices particularly for sports and entertainment applications. They claimed that complex dependencies of service modules aimed at optimizing the execution time and energy consumption during real-time execution of services.

Authors of [22], have designed a face occlusion technique using an IoT device particularly focused on criminal behavior. Furthermore, it uses deep learning techniques and utilizes gradient and shape cues to yield the best results in face detection using face occlusion.

Nowadays, face detection can be used for a variety of applications apart from just for access control or identification. Facial biometrics has also been studied for demographic information, for example, age, gender, and mood. Such parameters can be fed to the Interactive billboard or sign module for determining reactions. Hence, face detection is not only broadened its scope in terms of application apart from its core issues related to image processing. However, its core issues are still being researched such as time spent in processing is still a challenge, and face occlusions because of various reasons related to the face captured and the surrounding atmosphere. These problems are still at large and the research fraternity to suggest a more efficient solution to avoid the aforementioned issues with respect to image processing and identification.

### III. A THREE LAYERED DECENTRALIZED IoT BIOMETRIC ARCHITECTURE

To have scalability and management of the complete proposed system we have proposed edge-based monitoring and control system for the residents. Hence, the core of the proposed system is edge layer, because it receives important data and it controls the camera other sensing devices. Secondly, we have an application layer that uses smart devices to interact with the edge layer to have data from the sensors and cameras. Device layer implements, core proposal of face detection and recognition to restrict residents. Residents, are required to be seized in their respective homes, however, to stay during the lockdown is difficult and sometimes lockdown is partially relaxed for few hours a week so that people can visit their shopping centers for buying essential goods. For such situations, allowing only one person from each family is accepted to go out for shopping. To enforce such laws, we have designed image processing-based solution. Figure 1 shows the architecture used in this research work.

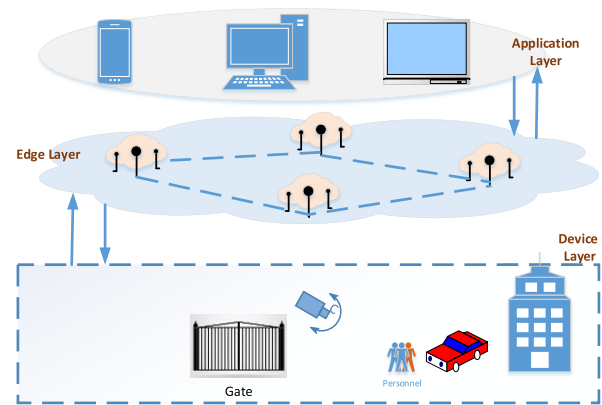


FIGURE 1. Proposed architecture, consists of a device layer, an edge layer and an application layer.

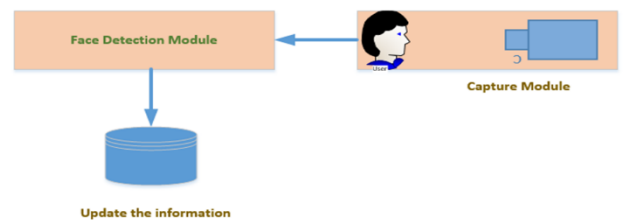


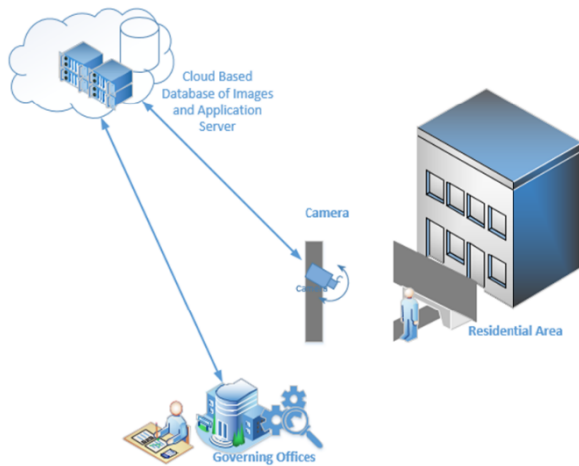
FIGURE 2. Physical Layer communication, its capture module.

The mapping between the application layer, device layer, and edge layer is a straight forward communication. The lowest layer is called the device layer, it is collection of physical devices connected together to achieve certain work assigned to them by application layer through edge layer. The edge layer is the backbone of the complete system which sits between the physical layer and the application layer and has computational units. The application layer, it is completely controlled by the human for their intended jobs, in our case, we have face identification and recognition system.

#### A. PHYSICAL LAYER

It hosts several devices, and their job is to constantly collect face image data from the residential area especially at the entrance. These devices can be fixed cameras and sensor smart devices. Each of these devices has ported on Raspberry Pi device and hence they execute numerous jobs on their own. It contains capture and face detection and database connection modules. The capture module captures the face of the person who intends to leave the apartment. This face image will be sent to the cloud-based application server and, once the face is detected, the person's information is updated.

Fig 2 physical layer, contains a camera to capture the face, and it sends captured images to the cloud-based database and application server. The application server recognizes the face and blocks other family members. The database of images is gathered from each residential area, presently we have used a set of images from (WIDER FACE dataset) This database also hosts house wise or flatwise (house) images of the residents. Presently, our proposal is in the initial stage of



**FIGURE 3.** Proposed Method, edge layer with application server and database interacting with cloud service provider.

development, so it hosts very few images of 100 flats data, but it can be employed to run into a real environment.

### B. THE EDGE LAYER

In this section, we brief the design of the edge layer and its working. Mainly, the edge layer receives and stores information in database originating from physical layer for aiding in decision making (see algorithm 1) to recognize and to transfer information to others. The device layer receives data from the camera and sensors and the data is dispatched inside the device layer for further analyzing purpose. The Raspberry Pi devices nodes allow a quick response time and without compromising the computation of the system. The main job of the edge layer is to combine and compute data and from various kinds of devices. Hence, gathered information from different IoT devices to have a better understanding and interpretation of the ground data and ensure the best decision.

The edge layer, Fig 3, is consists of cloud architecture and hosts application server and cloud database, which is continuously receives data from physical layer devices. Every time a physical layer device captures face image and reported to the edge layer in order to recognize and ultimately take suitable countermeasures.

The application server takes care of the job of bookkeeping of residents who are out & in and block people who are not eligible to leave houses. This application is connected to the local authorities to have a watch on the residents. Through local authorities' police can be alerted to find mischiefs people who break laws of lockdown.

With respect to the application server, which executes major modules called deep learning module built using CNN along with a multi-task cascaded method. It has three modules, namely CNN based series of bounding boxing of the captured face image, secondly, regression of bounding boxes, and image enhancement. Bounding boxes created by the previously yield multiple bounding boxes. Therefore, to avoid overlapping of such bounding boxes, non-maximum suppression (NMS) is adopted to refine the bounding boxes given

by detectors. All these images are fed to another Refiner network, which is the second module of the project. This second module is applied to further enhance the image by using bounding box regression. The third module is further enhancing the image by classifying main features such as the eye (left and right), nose, and mouth (right and left) of the face.

### C. FACE DETECTION USING CNN ARCHITECTURE

Algorithm 1, shows the steps used to detect face using library C++. It is proven that multiple CNN to detect faces have complications because of the fact that lack of multiplicity of weights that have resulted in the production of discriminative descriptions. Hence, face detection needs a smaller number of filters and a greater number of discriminative descriptions. We have used the python library to develop Face Detection with Deep Learning. Our proposed method is available in the form of a programming library called OpenCV library, it is written in the C++ computer vision library which provides a python interface. Hence, pre-trained faces are also employed with this library and it comes with plugins to train any model of your own dataset. Our proposed method has three stages as follows;

The proposed method places the camera at the front gate that captures the face of the person and produces different sized images of the face, which looks like image pyramid. These images are fed to the Multi-task Cascaded Convolutional Network (MTCNN) daemon process, which recognizes the face eventually produces five facial landmarks' positions. In the first stage, the kernel passes through the face image from top left corner browses (technically it is called stride by 2 pixels) till it gets complete landmarks of face coordinates by bounding boxes. Shifting of two cells which is called stride, helps the normal central processing unit to perform a calculation without any bottleneck. The key elements of the first stage, weights and biased are trained so that they can accurately produce bounding boxes for the captured image of size  $12 \times 12$  kernel. Hence, it keeps on generating bounding boxes with certain confidence levels, the daemon process deletes the bounding boxes with fewer confidence levels and considers only the highest confidence levels. However, this process will produce a lot of boxes which overlap each other, finally, the daemon process uses the NMS process, which reduce the number of bounding boxes and produces a list of the highest confidence level of boxes, and then the information will be fed to the second stage.

In the second stage, each bounding box has an array of pixels created to form a new array. For example, if the bounding box is slightly away from the portion of the image then we consider only the pixel values which are inside the bounding box and mark the rest of the pixels as zero, this processing called padding. The processing of padding for a processed image enables a more accurate analysis of the image. When we continue to perform padding, our proposed method removes lower confident boxes, and employ suppression techniques on lower confident boxes to eliminate



them. Then newly formed bounding boxes are built using the P-Net bounding boxes; hence they need conversion to form standard coordinates. The output of the second stage is quite similar to the first stage, however, it contains the coordinates of the new face image, more truthful bounding boxes, for each bounding box.

In the final stage, it starts with padding of bounding boxes who are out of bounds and resize the pixels. This final process yields the coordinates bounding box, facial landmarks, and the confidence level of each bounding box. To get all this, by removing the bounding boxes of lower confidence and standardize the coordinate of bounding boxes and facial landmarks by using the NMS method.

Once the image is received from the camera will be sent to the application server. The application server runs through 17 steps to produce facial landmarks. Apart from running the daemon process, it also handles updating the image-based information in the database and it also helps local governing bodies to look after any miscreants who breaks the law of lockdown.

---

**Algorithm 1** Algorithm for Detection of Face, Using Multi-Task Cascaded Convolutional Network

---

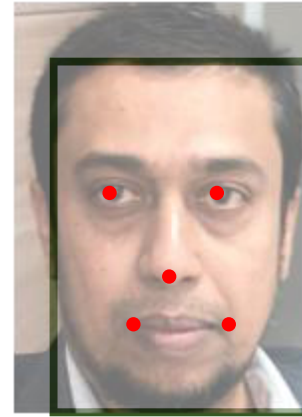
Input: Capture the Image

Output: Confidence of the bounding boxes and facial landmarks.

- 1: Capture the image of the person
  - 2: Pass it to daemon process to process on the image
  - 3: Create pyramid of image; send this to stage 1
  - 4: Develop bounding boxes for each image;  
from top left corner of the image to down right
  - 5: Delete all the bounding boxes who less confidence
  - 6: Induce NMS method, for each scaled image
  - 7: Readjust the bounding boxes
  - 8: Padding
  - 9: Feed output to the second stage
  - 10: Again, delete the bounding boxes
  - 11 Apply NMS methods to all bounding boxes with lower confidence level
  - 12: Re-adjust the bounding boxes
  - 13 Padding
  - 14: Feed output to the third stage
  - 15: Delete unnecessary bounding boxes
  - 16: Convert high confidence bounding into facial landmarks
  - 17: Result: Confidence of the boundin boxes and facial landmarks
- 

#### D. BOUNDING BOXES

In image detection, we have applied bounding boxes to find the facial landmarks. The bounding box is a box that is identified with the help of coordinates in the upper-left corner and lower-right corner of the box. In each stage, we have applied these bounding boxes to find an image with the highest level of confidence. In this research, we have used regression of



**FIGURE 4.** Facial landmark localization using MTCNN method on any given dataset.

predicted bounding boxes and ground truth boxes are computed in each stage with the help of cross-entropy loss.

$$L_i^{det} = -\left(y_i^{det} \log(p_i) + (1 - y_i^{det})(1 - \log(p_i))\right) \quad (1)$$

The aforementioned Eq. 1 is derived from the Kaipeng Zhang et.al. Eq 1 produces face with the ground-truth label. Eq 2, gives the bounding boxes derived from the Euclidean loss for each pixel. It contains regression target denoted as against the ground truth. Each of the stages is run through bounding boxes regressions, this results in multiple edge detection of the face. Hence, to reduce the bounding boxes with lower confidence NMS is applied.

Fig 4. Shows. Facial landmarks, we can also apply facial landmark detection can also be calculated by Euclidean loss [32].

$$L_i^{Landmark} = \left\| \hat{y}_i^{Landmark} - \hat{y}_i^{Landmark} \right\| \quad (2)$$

Eq. 2 consists of facial coordinate and ground truth coordinates. This equation yields us facial landmarks when it is bounding box regressively applied on the face image by using all three phases (P-net, R-Net, and O-Net). Hence, masked persons cannot be identified.

To conveniently describe abbreviations, and corresponding definitions used throughout manuscript are listed in Tab. 1

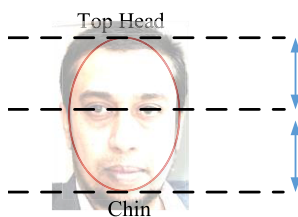
$$L_i^{box} = \left\| y_i^{box} - y_i^{box} \right\| \quad (3)$$

Our proposed method returns similarity index as we input two same person's images, whereas it produces lower similarity index when the face recognition receives a different person images. Hence, it should be trained to achieve the aforementioned capabilities, therefore following are the steps are used to detect face and comparison purpose.

- Detect/identify by using a face detection model
- Predict landmarks for the image identified in step 1
- Using these landmarks calculate face encodings.
- Compare the face encodings of known images with those from saved database images to identify the face.

**TABLE 1. Abbreviations.**

Symbol	Description
$L_i^{det}$	Entropy loss function
$y_i^{det}$	Ground truth
$L_i^{box}$	Bounding box
$W$	Waiting Time
$T_\delta$	Average transmission
$P_Q$	Probability of the Queue
$\mu$	Service rate
$\rho$	Utilization factor
$L_C$	Cloud utilization
$UD$	UE upload and download line
$L_e$	Delay at the edge


**FIGURE 5. Guidelines for drawing ellipses around face regions from FDDDB.**

To have uniformity across all the dataset of multiple face annotators, we used a set of instructions as suggested in our datasets. However, the FDDDB dataset uses ellipse to fit facial landmarks as shown in Fig 5.

### E. EDGE COMPUTING

Edge computing is directly communicating with the physical layer environment node or sensors, camera, and it is collectively called IoT. An IoT is represented as 4 tuples  $\langle C_t, L_g, F_a, N_L \rangle$ , where  $C_t$  is the computational time spent on the task. The performance of edge computing depends on  $C_t$  because IoT consumes more computational time then the performance of the complete framework will be degraded and the throughput of the system will also be further degraded, and it will not be a useful candidate for a real-time decision-making application. The second tuple  $L_g$ , is geolocation of the IoT in our case it is either camera or thermal sensors.  $F_a$ , represents the availability of the IoT device and its computational load will be shared among the available IoT's if it is not available. Finally,  $N_L$  represents the latency of the particular geo-located IoT.

Furthermore, the mean arrival time of the request for face recognition requests can be distributed within a residential area is given by the M/M/c and queuing delay can be had from the following. An M/M/c is a queuing system with impatient residence and synchronous lockdown policy. Residence arriving according to a Poisson process rate. The service of capturing face image by the small computer node who sever residence on a first come first basis and time spent on each residence is exponentially distributed  $\frac{1}{\mu}$ . When a node finds none of the residences in the queue then all nodes who are serving will consume zero bandwidth otherwise, they will

serve residents who are willing to leave the campus. The time spent idle is exponentially distributed with a mean  $\frac{1}{\gamma}$ . Once the node finds faces then all node who are idle will be active and this also exponentially distributed with mean  $\frac{1}{\xi}$ . The delay incurred due to this variation in the service and edge cloud computing which resembles the data center. In edge computing, the arrival of the task is processed by nodes hence, communication in the edge computing increases when the load rises and jobs (residents) are queued till their face is processed.

$$\begin{aligned} \text{Delay} &= d_{IoT} = W + \frac{1}{\mu} + T_\delta \\ &= P_Q * \frac{\rho}{\eta(1-\rho)} + \frac{1}{\mu} + T_\delta \end{aligned} \quad (4)$$

$$\text{Probability of the Queue} = P_Q = \frac{(c\rho)^c}{c!} \frac{1}{1-\rho} \rho_0 \quad (5)$$

$$\text{Initial Probability} = \left[ \sum_{k=0}^c \frac{(c\rho)^k}{k!} + \frac{(c\rho)^c}{c!} \times \frac{1}{1-\rho} \right] \quad (6)$$

$W$  is the waiting time in the queue,  $T_\delta$  it is the average transmission time for our proposed edge computing units and it is further defined in [30].

The Latency of the conventional network consists of few components, but in cloud architecture, it is a combination of the delay incurred due to irreducible propagation, transmission, routing, and the overall processing time. Sometimes few researchers suggested using software-defined networking switching delays as core to the main delay components [31]. Because delay increases as the number of users increase on particular geolocation for the cloud resource. The Following equation is formed for network latency at cloud without a third layer.

$$\begin{aligned} L_C &= (\text{UE upload and download line}) \\ &\quad * (\text{minimum distance of UE to the nearest AP}) \\ &\quad + (\text{AP to Cloud's Upload and download rate}) \\ &\quad * (\text{Distance from AP to Cloud}) \\ &\quad + (\text{Observer delay in IoT node}) \end{aligned}$$

Similarly the total delay observed at edge computing is as follows.

$$\begin{aligned} L_e &= (\text{UE upload and download line}) \\ &\quad * (\text{minimum distance of UE to the nearest AP}) \\ &\quad + (\text{Observer delay in IoT node}) \\ &\quad + \text{Control plane switching latency} \\ &\quad \times \text{from AP to AP with in edge} \end{aligned}$$

Aforementioned equation for latency in cloud and edge represented in Eq 6 and 7.

$$L_C = (UD) * (\min_{i=0}^m (d_{UE-AP})) + (UD_C) * (d_{AP-C}) + (d_{IoT}) \quad (7)$$

$$L_e = (UD) * (\min_{i=0}^m (d_{UE-AP})) + (d_{IoT}) + (d_s) \quad (8)$$

IV. EVALUATION

We have created an environment to evaluate the efficiency of the proposed method. In this section, we have compared our method with another state-of-the-art methods such as the FDDB, WIDER Face and Annotated facial Landmarks (AFLW). We have used FDDB data set which contains annotated 5171. Wider has more than half of them are used for testing categorized into 3 namely difficulty set, training (forty percent) set, and remaining validation subsets (10 %). Our test consists of placing devices and computational units within a building. Camera and thermos sensors placed on the Raspberry Pi computer placed on the device layer using wireless access points. The capture module and heat sensor to detect the temperature of residents are placed on these small computers. The edge layer is consists of server and database entities, all these servers are Linux based POSIX (Portable Operating System Interface) systems and SQLite is used as a database for faster retrieval of data and communication purposes.

A. TRAINING DATA

WIDER FACE dataset is a face detection benchmark dataset, of which images are selected from the publicly available WIDER dataset. Since, MTCNN network uses face detection and alignment, however, we have used the only face detection, hence we have three types of annotations for training purposes. Negatives, we have used less than 0.3 as the ratio of the IoU regions to any ground truth face. The Positive region is more than 0.65 and landmark for facial localization. In the first stage of the MTCNN method is applied to crop WIDER FACES to arrange positive and negative and we have also used CelebA as landmark faces. In the seconds' phase we have detected faces from the dataset of WIDER FACES and landmarked the faces from CelebA. In this article we detect a face, therefore we have utilized various data annotation technique for the training process firstly, Negatives, it is an area which is denoted by Intersection-over-Union (IoU) has less than 0.3 ratios to any ground-truth faces. The second one we have Positives, it is IoU ration is more than 0.65 to a ground truth face. Third, we call it a part of faces and has a ratio between 0.4 and 0.65 to a ground truth face. Finally, we have Landmark faces. The training data for each network as mentioned in the proposal is used as follows.

- P-Net Crop patches of WIDER face (To Derive Positive and negative, Part of face and face landmarks)
- R-Net (Use the out from the first stage to detect faces to mark positive, negative, part of the face and landmark from the CelebA dataset)
- O-Net (We have used this network to receive face data from this network. It is quite similar to the aforementioned network; however, we have utilized for face detection only)

We have performed an evaluation of the proposed method of face detection and compared with the state-of-the-art [1]–[6] methods that have utilized FDDB datasets and WIDER FACE

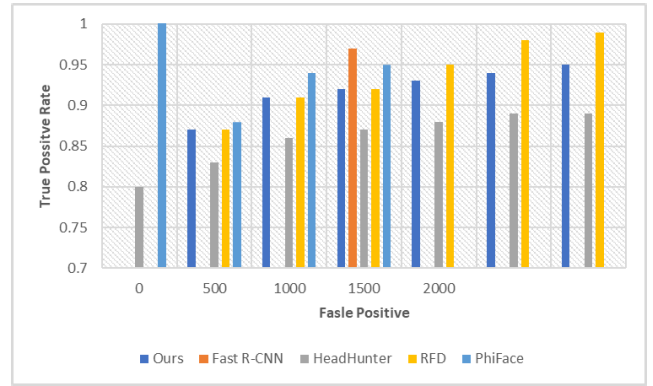


FIGURE 6. Evaluation of the proposed face detection on FDDB and WIDER Face.

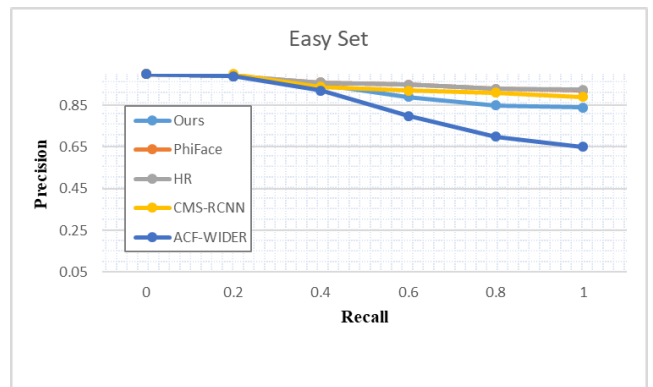


FIGURE 7. Evaluation of the proposed face detection on easy dataset FDDB and WIDER Face.

TABLE 2. Evaluation and comparison of WIDERFACE with state of arts methodology.

Average Precision/Recall/Hardest	Ours	PhiFace	HR	CMS-RCNN	ACF-WIDER
	0.607	0.810	0.806	0.624	0.273

data set. Fig 6 and 7 and Table 2, shows that our proposed research work outclass all the earlier research proposals.

For identifying face, we have collected data for different sized images and the time is taken by the processor to recognize the face. We have captured different sized images such as High definition images 1920 × 1080 to the lowest sized image of 320 × 240 image. The average time taken by the processor to process HD image is 100941 ms, this is result is calculated using 10 images. Similarly, when, the small-sized image processing time is 250ms. However, this processing time is directly proportional to the processing power the Raspberry Pi small computer.

We performed the experiments proposed system with respect to latency and detection. The Following settings were executed to measure the performance of the proposed system.

1. Moved the application server and database to the Edge layer

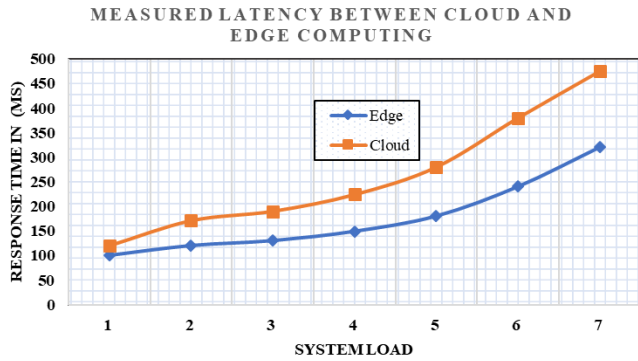


FIGURE 8. Latency of Cloud and Edge computing of Application and database server.

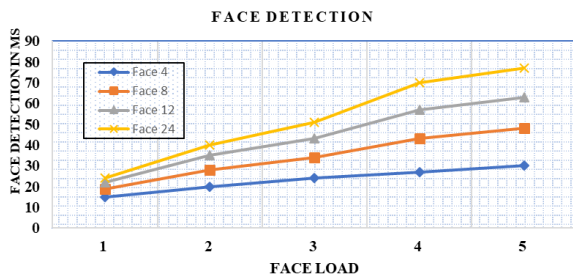


FIGURE 9. Face detection: average time taken by edge.

- Hybrid settings: Equally distributed the load between Cloud and Edge computing
- Moved computation entities to the Cloud

We hosted database server and application server at edge computing as shown in Fig 2. We ran several experiments to prove the efficiency of the proposed system against the settings of cloud-based services under the aforementioned experimental settings. We calculated latency for the aforementioned settings, and additionally, we also increased the device load on the current experiment settings. Fig 8, shows the latency of communication in Cloud and edge computing. Fig 9 shows an increase in face detection. It is the average face detection time in edge and cloud computing configuration. Figure 9 shows response time, measured separately for cloud and edge networks for the different numbers of face load when no bandwidth limitation on the edge and edge cloud architecture.

With these conditions, we have observed 45 ms response time with system load higher than 70% edge computing. For cloud-only architecture we have seen that the network latency making response time higher.

Figure 9 shows the bandwidth limitation of cloud computing on response time with variable load in the form number of faces. The limitation of bandwidth is between edge and cloud computing architecture. As the load increases, the detection time also increases under low bandwidth settings and whereas increased load on large bandwidth settings the average response time is an acceptable range.

Figures 10 and 11 show the average response time of the cloud and edge system on different bandwidth bounds.

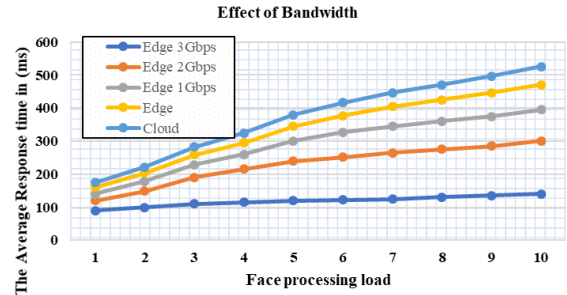


FIGURE 10. Average response time of the cloud and edge with different bandwidth setting.

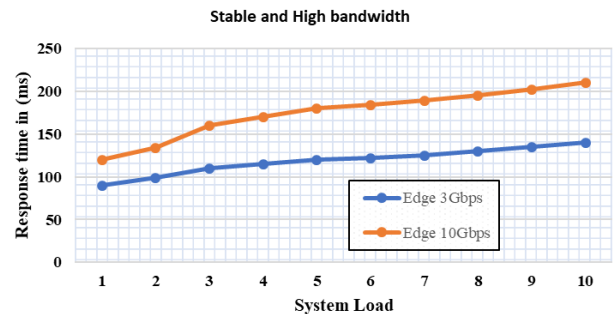


FIGURE 11. Average response time of the cloud and edge with stable and high bandwidth setting.

We have two experiments, one is fixed front haul bandwidth that makes resource utilization. In the second condition set of an experiment in that a cloud-only with inter edge bandwidth. In the first experiments, edge computing only competed with unlimited bandwidth systems and hence it utilizes complete resources of edge only systems. In the second experiment, the cloud-only system performs better than the edge systems in inter edge bandwidth conditions, because when the edge unable to process new requests it passes the new requests to the neighboring edge using inter edge bandwidth. With these settings, response time will be increased because the bandwidth is utilized to satisfy application requirements. Additionally, propagation delay and queueing will also increase response time under extremely high load. As the load increases, the bandwidth will be no more left and hence the more queueing delay. As more response time delay causes application underperformance and it won't handle more load of faces. Acceptable average response time for real-time image processing should not be more than 120ms. But if the bandwidth is increased, then propagation delay and queueing will not be noticed and the application will also have any trouble with the average response time. Figure 10 illustrates the required bandwidth to have constant fronthaul and inter edge bandwidth.

For testing purpose, we have extracted 10 faces per second for any given camera within the vicinity of the residential area. Our proposed method, if it receives 24 frames then the total number of frames is given.

$$\begin{aligned} \text{Total number of frames} &= \text{Faces /seconds} * 60\text{minutes} \\ &10 * 60 * 60 \text{ (minutes)} = 1200 \text{ frames.} \end{aligned}$$



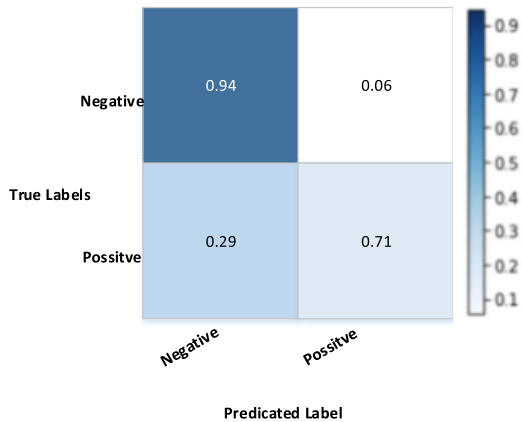


FIGURE 12. Normalized confusion matrix using our own dataset.

Therefore, it takes 10 hours to process one hour for 1200 frames. With the MTCNN, our proposed method has taken  $1200 \text{ frame} * 100 \text{ frames/second} = 1200/100 = 12 \text{ minutes}$ . Fig 12, shows the confusion matrix obtained from the proposal using MCCNN network and obtained satisfactory results which coincides with the proposal.

## V. CONCLUSION

This research presented a decentralized IoT based solution for lockdown issues for the residential population to avoid flooding of people so as to follow guidelines. We have validated our proposal against the state-of-the-art face detection and recognition methods under distributed and centralized architecture in terms of a load of devices and their work is done with respect to response time. For the face detection method, we have utilized a CNN based multi task cascaded framework. The performance shows that our research method performs better than the state-of-the-art proposals on various challenging datasets namely WIDER FACE (WIDER FACE dataset is a face detection benchmark dataset) and Face Detection Data Set and Benchmark. For decentralized conditions, we have compared the results of latency and load of devices on Cloud and Edge computing, on various centralized and distributed conditions of the application and database server. With respect to Edge computing we have achieved a better result on the decentralized load of face detection is achieved better results. In the future, we will study the essential features between face detection and its impact on the communication protocol and payload.

## FUNDING

None.

## ACKNOWLEDGMENT

The authors appreciate the Research Deanship of Prince Sat-tam Bin Abdulaziz University, Saudi Arabia, for providing research resources and equipment and various research programs to encourage research among faculty members.

## CONFLICTS OF INTEREST

Declare conflicts of interest or state "The authors declare no conflict of interest.

## REFERENCES

- [1] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future Gener. Comput. Syst.*, vol. 29, no. 7, pp. 1645–1660, Sep. 2013.
- [2] M. A. Khan and K. Salah, "IoT security: Review, blockchain solutions, and open challenges," *Future Gener. Comput. Syst.*, vol. 82, pp. 395–411, May 2018.
- [3] S. Chen, H. Xu, D. Liu, B. Hu, and H. Wang, "A vision of IoT: Applications, challenges, and opportunities with China perspective," *IEEE Internet Things J.*, vol. 1, no. 4, pp. 349–359, Aug. 2014.
- [4] M. Chiang and T. Zhang, "Fog and IoT: An overview of research opportunities," *IEEE Internet Things J.*, vol. 3, no. 6, pp. 854–864, Dec. 2016.
- [5] B. Xu, L. Da Xu, H. Cai, C. Xie, J. Hu, and F. Bu, "Ubiquitous data accessing method in IoT-based information system for emergency medical services," *IEEE Trans. Ind. Informat.*, vol. 10, no. 2, pp. 1578–1586, May 2014.
- [6] M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for Internet of Things (IoT) security," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1646–1685, 3rd Quart., 2020.
- [7] H. Jiang and E. Learned-Miller, "Face detection with the faster R-CNN," in *Proc. 12th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, May 2017, pp. 650–667.
- [8] A. Ospina and F. Torres, "Counter: Count without bells and whistles," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Cham, Switzerland: Springer, Jun. 2020, pp. 720–735.
- [9] J. Zhang, X. Wu, S. C. H. Hoi, and J. Zhu, "Feature agglomeration networks for single stage face detection," *Neurocomputing*, vol. 380, pp. 180–189, Mar. 2020.
- [10] Z. Zhang, W. Shen, S. Qiao, Y. Wang, B. Wang, and A. Yuille, "Robust face detection via learning small faces on hard images," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2020, pp. 1361–1370.
- [11] S. Wu, M. Kan, S. Shan, and X. Chen, "Hierarchical attention for part-aware face detection," *Int. J. Comput. Vis.*, vol. 127, nos. 6–7, pp. 560–578, Jun. 2019.
- [12] C. Chi, S. Zhang, J. Xing, Z. Lei, S. Z. Li, and X. Zou, "Selective refinement network for high performance face detection," in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, Jul. 2019, pp. 8231–8238.
- [13] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," in *Proc. Eur. Conf. Comput. Learn. Theory*, Berlin, Germany: Springer, Mar. 1995, pp. 23–37.
- [14] T. Hastie, S. Rosset, J. Zhu, and H. Zou, "Multi-class AdaBoost," *Statist. Interface*, vol. 2, no. 3, pp. 349–360, 2009.
- [15] D. Ghimire and J. Lee, "Geometric feature-based facial expression recognition in image sequences using multi-class AdaBoost and support vector machines," *Sensors*, vol. 13, no. 6, pp. 7714–7734, Jun. 2013.
- [16] M. Arafah, A. Achmad, Indrabay, and I. S. Areni, "Face recognition system using Viola Jones, histograms of oriented gradients and multi-class support vector machine," *J. Phys., Conf. Ser.*, vol. 1341, Oct. 2019, Art. no. 042005.
- [17] L. M. Dang, S. I. Hassan, S. Im, and H. Moon, "Face image manipulation detection based on a convolutional neural network," *Expert Syst. Appl.*, vol. 129, pp. 156–168, Sep. 2019.
- [18] S. Nayak, S. L. Happy, A. Routray, and M. Sarma, "A versatile online system for person-specific facial expression recognition," in *Proc. IEEE Region 10 Conf. (TENCON)*, Oct. 2019, pp. 2513–2518.
- [19] M. R. Haque, A. A. Saud, Y. B. Annajiat, and S. Hossain, "Solving multi-class classification tasks with classifier ensemble based on clustering," M.S. thesis, United Int. Univ., Dhaka, Bangladesh, 2019.
- [20] Q. Wang, L. Gao, H. Wang, and X. Wei, "Face detection for privacy protected images," *IEEE Access*, vol. 7, pp. 3918–3927, 2019.
- [21] Z. Mahmood, N. Bibi, M. Usman, U. Khan, and N. Muhammad, "Mobile cloud based-framework for sports applications," *Multidimensional Syst. Signal Process.*, vol. 30, no. 4, pp. 1991–2019, Oct. 2019.
- [22] L. Mao, F. Sheng, and T. Zhang, "Face occlusion recognition with deep learning in security framework for the IoT," *IEEE Access*, vol. 7, pp. 174531–174540, 2019.
- [23] J. Kim, Y. Sung, S. M. Yoon, and B. G. Park, "A new video surveillance system employing occluded face detection," in *Proc. Int. Conf. Ind., Eng. Other Appl. Appl. Intell. Syst.*, Berlin, Germany: Springer, Jun. 2005, pp. 65–68.
- [24] J. Kim, J. K. Suhr, H. G. Jung, and J. Kim, "Face occlusion detection by using B-spline active contour and skin color information," in *Proc. 11th Int. Conf. Control Autom. Robot. Vis.*, Dec. 2010, pp. 627–632.

[25] C.-M. Chen, B. Xiang, Y. Liu, and K.-H. Wang, "A secure authentication protocol for Internet of vehicles," *IEEE Access*, vol. 7, pp. 12047–12057, 2019.

[26] X. Tan, S. Chen, Z.-H. Zhou, and F. Zhang, "Recognizing partially occluded, expression variant faces from single training image per person with SOM and softk-NN ensemble," *IEEE Trans. Neural Netw.*, vol. 16, no. 4, pp. 875–886, Jul. 2005.

[27] M. J. Beevi, "A fair survey on Internet of Things (IoT)," in *Proc. Int. Conf. Emerg. Trends Eng., Technol. Sci. (ICETETS)*, Feb. 2016, pp. 1–6.

[28] D. Dragomir, L. Gheorghe, S. Costea, and A. Radovici, "A survey on secure communication protocols for IoT systems," in *Proc. Int. Workshop Secure Internet Things (SIoT)*, 2016, pp. 47–62.

[29] F. Samie, V. Tsoutsouras, S. Xydis, L. Bauer, D. Soudris, and J. Henkel, "Distributed QoS management for Internet of Things under resource constraints," in *Proc. 11th IEEE/ACM/IFIP Int. Conf. Hardw./Softw. Codesign Syst. Synth. (CODES)*, 2016, pp. 1–10.

[30] J. Liu, Y. Mao, J. Zhang, and K. B. Letaief, "Delay-optimal computation task scheduling for mobile-edge computing systems," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 1451–1455.

[31] S. Maheshwari, D. Raychaudhuri, I. Seskar, and F. Bronzino, "Scalability and performance evaluation of edge cloud systems for latency constrained applications," in *Proc. IEEE/ACM Symp. Edge Comput. (SEC)*, Oct. 2018, pp. 286–299.

[32] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1499–1503, Oct. 2016.



**FADI AL-TURJMAN** (Member, IEEE) received the Ph.D. degree in computer science from Queen's University, Kingston, ON, Canada, in 2011. He is currently a Full Professor and a Research Center Director of Near East University, Nicosia, Cyprus. He is a leading authority in the areas of smart/intelligent, wireless, and mobile networks' architectures, protocols, deployments, and performance evaluation. His publication history spans over 250 publications in journals, conferences, patents, books, and book chapters, in addition to numerous keynotes and plenary talks at flagship venues. He has authored and edited more than 25 books about cognition, security, and wireless sensor networks' deployments in smart environments, published by Taylor & Francis, Elsevier, and Springer. He received several recognitions and best papers' awards at top international conferences. He also received the prestigious Best Research Paper Award from *Computer Communications Journal* (Elsevier), from 2015 to 2018, in addition to the Top Researcher Award for 2018 at Antalya Bilim University, Turkey. He has led a number of international symposia and workshops in flagship communication society conferences. He serves as an Associate Editor and the Lead Guest/Associate Editor for several well reputed journals, including the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS (IF 23.9) and *Sustainable Cities and Society* Elsevier (IF 5.7).



**ABDALLA ALAMEEN** received the B.Sc. degree in computer science from the International University of Africa, Sudan, in 1997, and the M.Sc. and Ph.D. degrees in computer science from the Department of Computer Science and Information Technology, AL Neelain University, Sudan, in 2010 and 2013, respectively. He is currently working as the HOD and an Associate Professor at the Computer Science Department, College of Arts and Science, Prince Sattam Bin Abdulaziz University, Saudi Arabia. His research interests include information security, cloud computing, E-government, database, and web application.



**MOSLEH M ABUALHAJ** received the Ph.D. degree in multimedia network protocols from the University of Malaya. He is currently a Senior Lecturer with Al-Ahliyya Amman University. His research interests include Volpe, multimedia networking, congestion control, and cloud computing.

...



**MANJUR KOLHAR** received the Ph.D. degree in heterogeneous IP communication systems from the National Advanced IPv6 Centre of Excellence, Universiti Sains Malaysia. He is currently an Associate Professor of computer science and information with Prince Sattam Bin Abdulaziz University, Saudi Arabia. His research interests include computer science, computer networks, cloud computing, machine learning, and resource management.