International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 10, No.1 (Aug-2021)
http://dx.doi.org/10.12785/ijcds/100179

Smart Campus Monitoring Based Video Surveillance using Haar Like Features and K-Nearest Neighbour

Ali H. Hamad[1]

[1] *University of Baghdad, Department of Information and Communication Engineering, Baghdad, Iraq*

Received 02 Dec. 2020, Revised 03 Apr. 2021, Accepted 10 Apr. 2021, Published 5 Aug. 2021

Abstract: Intelligent video surveillance systems based on Internet of Things (IoT) technology have proven to be major primary components for security in many areas, such as smart cities. These systems are important because they provide messages that transfer information about people on campus among camera nodes, thereby providing real-time video surveillance monitoring. The proposed system consists of several cameras and an intelligent processing system, represented by a raspberry pi. The cameras are distributed in different locations in the university campus. Each camera node is connected to the internet and can communicate and share information with other nodes, as well as communicate with a central monitoring (server) via Message Queuing Telemetry Transport (MQTT) IoT protocol. The cameras can extract information in real time from video, and identify everyone as either students, teachers and/or employees using computer vision algorithms. Two methods of face detection and recognition techniques are applied: a feature-based technique that uses the Haar cascade, and an image-based technique that uses k-nearest neighbour (kNN). Face detection and recognition based on the Haar cascade classifier is more suitable for resources with embedded limited systems since it requires less computation, while kNN is more accurate and shows better results in a dynamic environment. All programs were written using open-source Python under a Linux operating system and by using OpenCV library.

Keywords: Smart Campus, Video surveillance, Haar Like Features, IoT, MQTT, KNN

1. INTRODUCTION

People typically use surveillance cameras to create safe areas in cities, campuses and home, using the cameras to monitor and record videos for later processing. Surveillance means monitoring human activities, behaviour and other information for protection purposes by applying remote monitoring devices such as Closed-Circuit Television (CCTV) cameras and electronic interception of stream transmissions (e.g. internet, smartphone). Intelligent video surveillance systems convert traditional surveillance from data acquisition and analysis to smart information analytics using real time imaging and video processing tools. Video surveillance systems can be applied in real time and extract useful information at a higher resolution [1].

Conventional video surveillance systems are either simple video recording, or streamed video observed by people without additional technological assistance. This makes installation and operation of such systems expensive. Modern systems use advanced technologies, including object identification, detection, tracking and pattern recognition, to identify objects as human, vehicle, or other [2]. Video surveillance systems include tools to capture, record, store and display video to users. These systems also have a mechanism to transfer video information between devices. These systems can have many issues, including the following:

- Requirements: An intelligent camera uses image processing to analyze and extract information from a video in order to make an appropriate decision. Also, communication between cameras can provide a wider monitoring area. Communication channels must also be secure for two reasons; video surveillance for security, and privacy.
- Benefits: Surveillance cameras have added new features for the protection and security of businesses and homes. Today's cameras have a range of options that encourage deployment in wide areas and multiple locations, and development in security camera systems has increased the security monitoring levels [3].
- Challenges: Large scale video surveillance analysis is the most important concern for the future of smart cities. Security systems can provide diverse functionality, including object detection, tracking, recognition, identification and feature extraction [4].

E-mail: ahamad@kecbu.uobaghdad.edu.iq

Video surveillance systems used in public locations must deal with large object traffic density, as well as heterogeneous, variable and occlusions of moving objects. Such systems must be able to detect complex and sudden events [5], and observe the activities, behavior and other immediate information regarding objects. Conventional video surveillance needs a lot of storage space, since it records and captures everything observed by a surveillance camera, and this requires much of the storage resources of the system devices. Extracting useful and meaningful information from recorded video is time consuming because of a long time of the recorded videos. These limitations disregard the advantages of traditional video surveillance systems. Recorded videos are analyzed for abnormal scenes, and the limitation of this approach in real-time is it reduces the utility and security of the system. With the advance of new technology and emergence of intelligent identification techniques, it is possible to instill traditional video surveillance with intelligent decision capabilities according to the required scenarios. It can also monitor scenes by classification, detecting, tracking and behavior analytics. Intelligent Video Surveillance systems (IVS) based on image and video processing are employed widely to prevent crime and provide security.

Computer vision requires intelligent tools and techniques to recognize different objects. Various algorithms and techniques, including fuzzy technology, neural network, genetic algorithm and support vector machines have been used in recent years. All have been tested experimentally and achieved promising results [6, 7]. Figure 1 shows how intelligent video surveillance can detect and recognize different objects of different sizes and types.

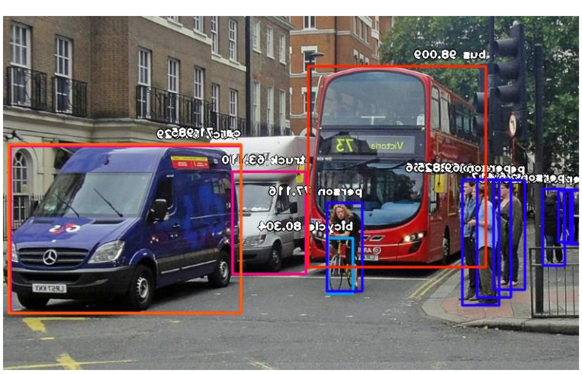


Figure 1. Object detection and recognition using computer vision

2. Literature Survey

Video surveillance with the internet of things is an important and promising technology that has attracted the attention of the researchers around the world. In 2017, Stergiou et al. [8] proposed an innovative topology paradigm that could better use IoT technology in video surveillance systems, with the aim of improving their use and enhance video data transmission through the network. Also in 2017, Patel et al. [9] introduced a traffic monitoring system based on the IoT for accident detection which uses raspberry pi, and cameras for live video streaming. The system generates information regarding traffic, such as number of vehicles, emergency accident situations, and improper projectile of vehicle. This information can be used to manage and divert live traffic as needed to avoid road traffic problems. In 2017 as well, Aydin et al. [10] introduced an IoT that worked with computer vision to detect the human faces. The proposed system uses raspberry pi and Passive Infrared Sensor (PIR) to detect movement. In 2018, Othman and Aydin [11] proposed a real-time system for face detection and recognition that can process images at a high speed. The objective of this work was to protect homes and offices by recognizing people, and using the PIR sensor to detect movement in specific areas. The Raspberry Pi can then capture the images, and the faces in the captured image will be detected and recognized. In 2018, Premavathi et al. [12] proposed a classification and feature descriptor for face recognition systems, using kNN with various proximity calculations and the results are assessed in terms of reliability and accuracy. Shao et al. [13] proposed an intelligent processing technique in 2018 that used messages from front-end smart cameras. And in 2017 Deshmukh et al. [14] presented a comparative study of two Machine-to-Machine (M2M) protocols: the constrained application protocol (CoAP) and message queuing telemetry transport protocol (MQTT) in terms of network latency and bandwidth consumption. In 2018 Oliveira et al. [15] compared WebSocket and MQTT protocols by using ESP8266 and a Node.js server. They measured the latency using a local network and a device with memory allocated.

This paper will contribute to the design and implementation of an intelligent video surveillance monitoring and system-based object tracking system. Cameras are distributed randomly over a wide area, and they share information extracted from different videos. The extracted information is then saved and processed in the cloud server. The purpose of this work is to design and implement an IoT video surveillance system for monitoring and tracking people on a university campus, and transfer the extracted information between nodes using the MQTT protocol. Moreover, it applies face detection and recognition with two techniques: Haar like features and k-nearest neighbour.

The balance of the paper is organized as follows: Section 3 presents the preliminary work for both the MQTT protocol and the face detection and recognition techniques. Section 4 considers the overall architecture of the proposed algorithms and examines the techniques.

Section 5 presents a case study based on results of the work in both normal and search modes. And Section six presents the conclusions of the work.

3. PRELIMINARIES

A. MQTT PROTOCOL

MQTT is a lightweight data oriented IoT protocol designed for limited resource devices such as IoT processes. The protocol uses publish-subscribe patterns for message transfer between IoT devices, and the devices that use MQTT protocol are not aware of devices on the other side. Some devices publish messages to a broker server, which can manage message delivery to the subscribers. MQTT uses the publish/subscribe (pub/sub) method for communication. The publisher publishes about a topic while the subscriber subscribes in the same topic the publisher used. Some devices publish messages to a broker server, and the broker delivers the messages to the subscriber devices. MQTT architecture, as shown in Figure 2, consists of three main components: a subscriber, a publisher and a broker. Each component has an independent work mechanism [16,17].

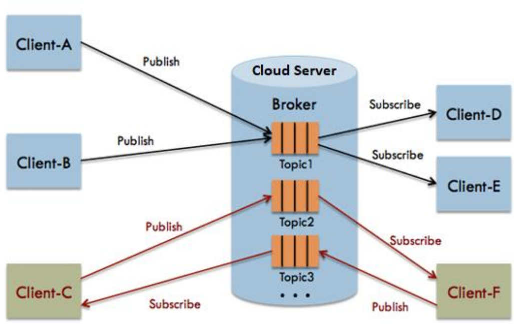


Figure 2. MQTT IoT protocol architecture

B. FACE DETECTION AND RECOGNITION TECHNIQUES

Face detection and recognition is a challenging problem in computer vision technology and image analysis. It attracts researchers because there are many different types of applications based on diverse methods and algorithms. Face detection has become a biometric sign in last few years, due to its importance in security applications. The first step in face analysis systems is real-time face detection.

The Haar cascade classifier is often used because it is fast, accurate and can detect multiple faces simultaneously. The most important feature of this method is that it can be trained off-line, while the process can take several hours to be trained on only one object. When well trained, it can work quickly with less computing functionality and with many types of embedded microprocessors, such as raspberry pi. A second technique based on a k-nearest neighbor algorithm has also been used for face detection and recognition. The kNN algorithm can recognize specific faces from a video with multiple student faces that were captured on video using a camera. The Histogram of Oriented Gradient (HOG) feature has been used as face detection algorithm [18] because it is robust and is not sensitive to light or geometric changes. In addition, the computational complexity of the HOG feature is far less than that of the original data.

i. HAAR-LIKE FEATURE

'Viola and Jones' is a widely used algorithm, and it was used in this work. It is an object detection method that reduces computation time while maintaining optimal accuracy detection. It is considered a powerful and fast face detection method that is 15 times faster than existing methods with 95% accuracy [19]. The algorithm uses Haar-like features to extract target features, which are then evaluated rapidly. A Haar wavelet is a mathematical approach that generates square waves to identify signs with sudden transformations. Figure 3 shows an example of combining several wavelets. A cascade can be created to identify edges, circles and lines with different color intensity.

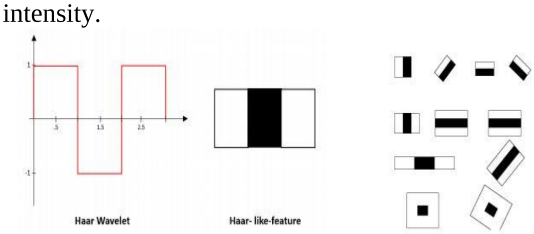


Figure 3. A Haar wavelet and resulting Haar-like features

ii. K- NEAREST NEIGHBOUR LASSIFIER (KNN)

k-nearest *neighbor* (kNN) is a machine learning technique that is widely used as an object classifier for face detection and recognition. Images are classified by the majority vote of their neighbours. The shortest Euclidian distance between a testing image pixel and a trained image pixel can be determined by a distance matrix [20,21]. Figure 4 shows an example of a kNN classifier with the square object representing the sample to be classified. It first needs to be classified into a star or a triangle. If k=5 for example, the probability of the square object is closer to the triangular object, while if k=10 the probability of the square object is closer to the stars. K-Nearest Neighbor (kNN) is a data classification algorithm used for face recognition. Each pixel in a digital image of the face represents information. In recognition, pixel matrix of face image should be reshaped into a vector before classification.

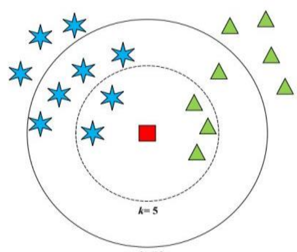


Figure 4. Principle of the K-NN classification algorithm.

4. PROPOSED SYSTEM DESIGN

The proposed system is a data transfer for a video surveillance system in an IoT infrastructure; the schematic diagram is shown in Figure 5. The system consists of several cameras distributed over a wide area in The University of Baghdad campus, and each camera captures a video stream and extracts the required information using a microcontroller embedded system. The system is designed using a raspberry pi camera module and an algorithm written in python code that uses an open cv library. In this paper, each pi camera is connected to a raspberry pi and distributed widely in areas such as university campuses. Figure 6 shows the raspberry pi devices with their cameras distributed in different places on the campus. The proposed system is designed so it can train on the faces of the students and staff offline to create a database. The system is then trained for this database for normal operation. The proposed system operation involves two phases: offline and online. Thus, two techniques have been used for face detection and recognition: the Haar cascade and k-nearest neighbor process.

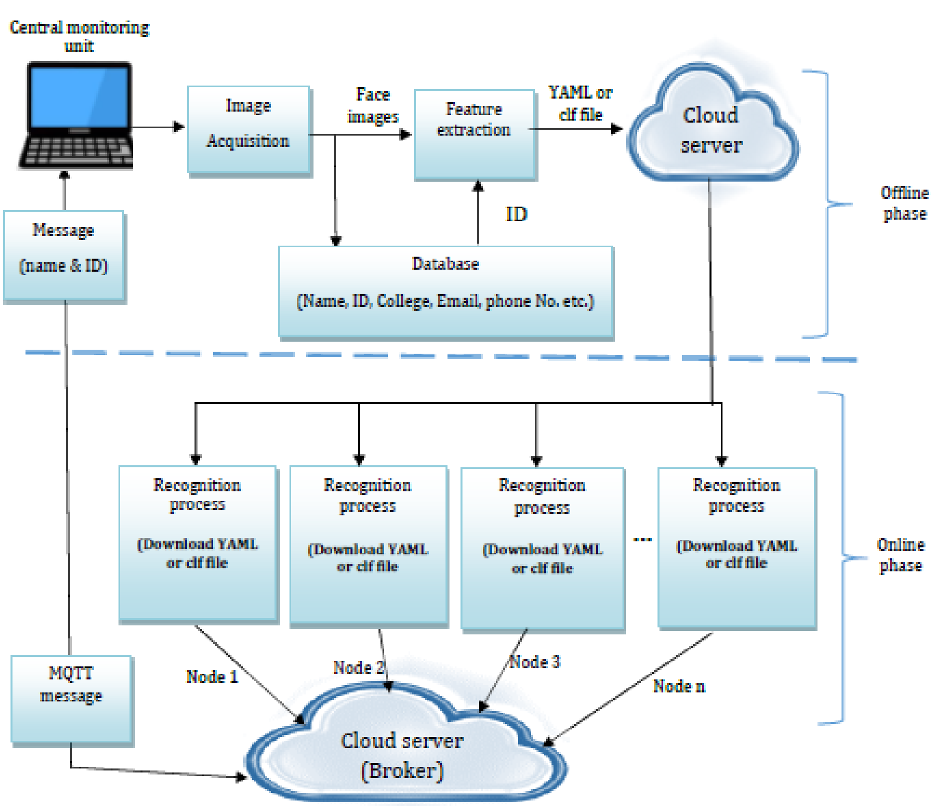


Figure 5. Proposed video surveillance system description

Figure 6. Distributed monitoring cameras in the campus location

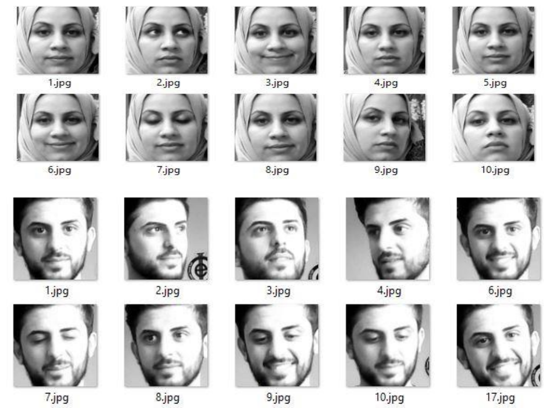


Figure 7. Face detection and training using Haar cascade

A-Offline phase using the Haar cascade classifier

In the face detection process, several images are taken of each person. The images are then customized to incorporate different aspects, including distances, places and lighting. All the images are processed in the training program to identify the area that represents a face only and isolate it from the rest of the image. Figure 7 illustrates a practical implementation of the face detection process. In this phase, the system analyses the faces of 60 people on the campus, including students, employees and professors. For every person, a set of 10 frames of their faces, and different postures such as sight intensity and emotions are used, with various environments such as sunlit or dark taken into consideration. This is done with a face detection program written in python that uses an opencv library. The size of the YAML file was 200 kB, which can be increased if more people are involved.

B- Online phase using the Haar cascade classifier:

The training process generates a YAML file that contains a database of face images and the name and ID of each individual. Face detection is then applied in the face recognition process, again to define a face in an image with different objects. The detected face is compared with the faces database in the YAML file, and if the face is identical to a known face (with confidence ratio of 90% selected) the system recognizes that person. In this work, the LBP algorithm was used for the recognition process. The system works in real time in this phase; each node first downloads the trained YAML file generated in the off-line phase, then the file is decrypted. This means each node has the same database for everyone on the campus with their ID.

C-Offline phase using the kNN algorithm

In the face detection process, several images are taken of each person, and all the images have different variables such as distances, places and lighting. All images are processed in the training program to identify the area of the face only and isolate it from the rest of the image. A ball tree algorithm is used in the training process; thus every node defines a D-dimensional hypersphere (i.e. ball) that contains a subset of the points to be searched [22]. Figure 8 shows the practical implementation of the face detection process.

D-Online phase using the kNN algorithm

Images of student faces have been registered previously with the students' names and identification information in the database. The system is trained on these images to generate a clf file (classifier), and uploads the file to each camera node in the system through the cloud server. The face recognition program is then executed to recognize any individual that passes near any of the cameras; if the person is in the database their name is returned to the admin, otherwise the system indicates 'unknown person'. This proposed work was implemented with an open CV library in python to ensure fast image processing. The main process recognizes faces continuously by capturing an image from the attached pi camera. The kNN classifier is first trained on a set of labeled (known) faces, and can then predict the person in an unknown image by finding the K most similar faces (images with closest face features under the Euclidian distance) in its training set, and performing a majority vote (possibly weighted) on their label.

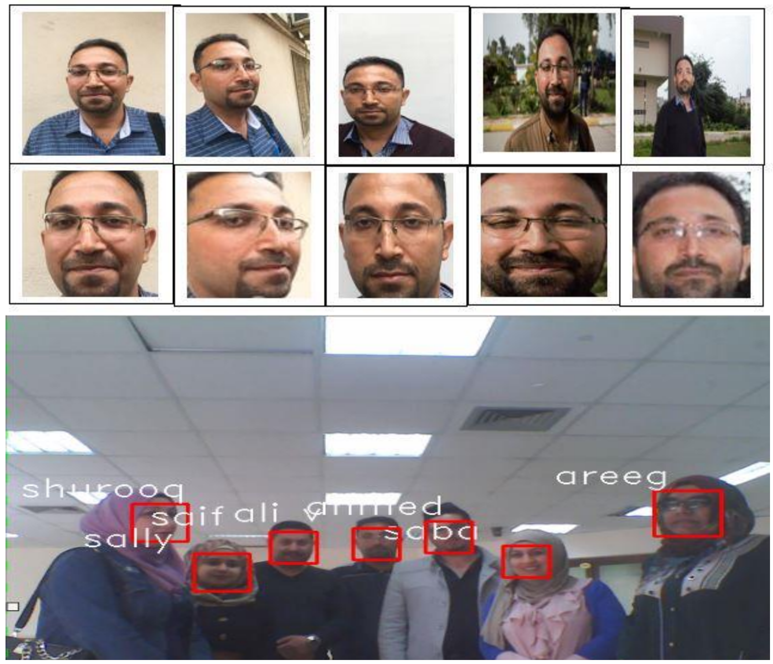


Figure 8. Face recognition using the kNN algorithm

5. CASE STUDY

i. Case One: Normal mode (monitoring mode)

All nodes initially work in this mode of operation, when nodes start working it first downloads the trained YAML file from the cloud server. This ensures all nodes have the same database of all the individuals in the campus. The recognition process then starts and updates the tracking database. In this mode, if any person is recognized by one of the cameras on the campus, it will update the database by sending the person's information to the central monitoring system. The information includes the person's ID, name, date, and location when he/she was recognized, as shown in Table 1. In monitoring mode, the system can search for any person in the database to find his location at a specific time; this can be used later for reports. Thus, in this case any person identified can be tracked; Table 2 shows a report of tracking a person within the campus. This surveillance video monitoring can work continuously 24/7. In this mode of operation, nodes work as publishers and send their information via the MQTT protocol to the central monitoring unit, which is a subscriber.

Table 1. Database for monitoring process process

No.	ID	Name	Date	Node
1	5	ali hussein	22-10-2018 20:41:17	Communications Dept.
2	35	dina jamal	23-10-2018 12:57:42	Network Dept.
3	35	sami abdullah	23-10-2018 13:02:26	Network Dept.
4		rana tariq	23-10-2018 13:05:53	Network Dept.
5	35	rada saad	02-11-2018 13:39:30	Network Dept.
6		ali hayder	02-11-2018 13:42:00	Communications Dept.
7	5	hana soud	02-11-2018 13:43:07	Communications Dept.
8		ghasaq saad	02-11-2018 13:43:09	Communications Dept.
9	38	suha fadhil	02-11-2018 13:43:10	Communications Dept.
10	5	lana saad	02-11-2018 13:43:14	Communications Dept.

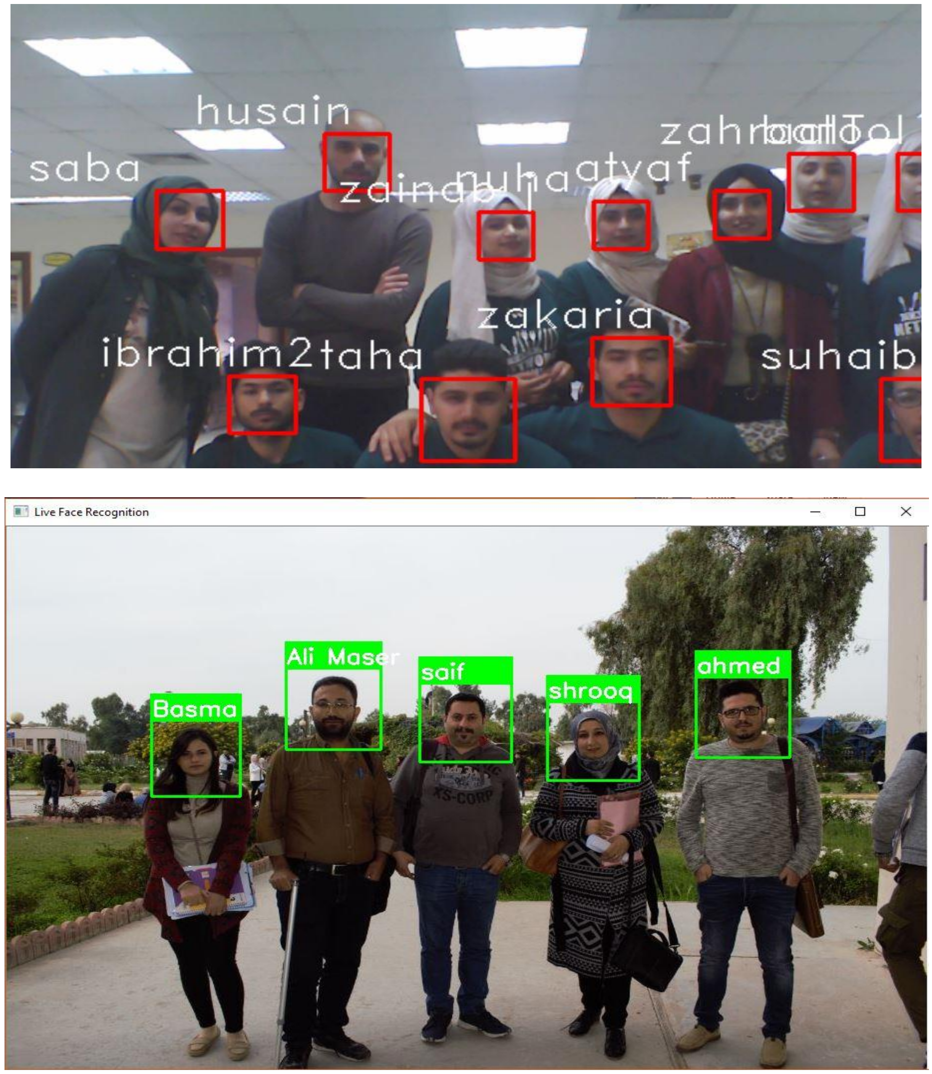


Figure 9. Search mode operation algorithm

Table 2. Database for tracking report process

No.	Id	Name	Date	Node
1	5	saba saad	22-10-2018 20:41:17	Communications Dept
2	5	saba saad	02-11-2018 13:39:30	Network Dept
3	5	saba saad	02-11-2018 13:43:07	.Communications Dept.
4	5	saba saad	02-11-2018 13:43:14	Communications Dept.
5	5	saba saad	02-11-2018 13:43:28	Communications Dept.

ii. Case 2: Search mode

When searching for a person who was trained previously by the system, and needs to know where the person is on the campus, the admin puts the person's name in the search box, and conducts the following processes (Figure 9 illustrates the search mode operation):

1- The program searches for the ID of the person in the MySQL database;

2- When the ID is found, it is broadcast as a message to all cameras using the MQTT protocol, with the central monitoring computer acting as the publisher;

3- A camera node, acting as a subscriber, will switch to the search processes after receiving the message from the broker, and then run a face recognition program to find the person;

4- If any node on the campus finds a match, it becomes a publisher and sends a message to all nodes (including the central computer) informing them that it has found the person; and,

5- All nodes then stop searching and return to normal mode.

CONCLUSION

The proposed system includes intelligent and secure video surveillance, with the video stream processed within the nodes rather than uploaded to the cloud server for later processing. This eliminates the need for high bandwidth, since the required features are extracted using an intelligent node (raspberry pi). The following data can be acquired doing this:

- Online video streaming to a cloud is limited by the internet bandwidth, which can cause a delay in processing;
- The MQTT protocol is a simple and fast method of message transfer among nodes;
- The Internet of things is a promising technology for enhancing previous technologies, including video monitoring.
- The Haar cascade classifier is more suitable for limited resources embedded in systems such as raspberry pi, while the kNN algorithm is more accurate.

REFERENCES

[1] Vasileios A. Memos, Kostas E. Psannis, Yutaka Ishibashi, Byung-GyuKim, B.B Gupta, "*An Efficient Algorithm for Media-based Surveillance System (EAMSuS) in IoT smart city framework*", Elsvier, Future Generation Computer Systems, pages 619–628, 2018.

[2] Qasim M. Rajpoot and Christian D. Jensen, *"Security and Privacy in Video Surveillance: Requirements and Challenges"*, Springer, SEC 2014, IFIP AICT 428, pp. 169–184, 2014.

[3] Kulkarni S., Bagul M., and Dukare A., " *Face Recognition using IoT*", IEEE, International Journal of Innovations & Advancement in Computer Science, (IJIACS), Volume 7, Issue 3, March 2018.

[4] Michał Z, Katarzyna G., Dariusz F., and Radosław H., "*Intelligent video surveillance systems for public spaces – a survey*", Journal of Theoretical and Applied Computer Science Vol. 8, No. 4, pp. 13-27, 2014.

[5] Nana Tang, "*Motion image detection algorithm for juvenile obesity based on video surveillance and internet of things*", Microprocessors and Microsystems, Volume 81, 2021,

[6] Santiago Felici-Castell, Miguel García-Pineda, Jaume Segura-Garcia, Rafael Fayos-Jordan, Jesus Lopez-Ballester, "*Adaptive live video streaming on low-cost wireless multihop networks for road traffic surveillance in smart cities*", Future Generation Computer Systems, Volume 115,Pages 741-755, 2021

[7] Juan Isern, Francisco Barranco, Daniel Deniz, Juho Lesonen, Jari Hannuksela, Richard R. Carrillo,Reconfigurable, "*cyber-physical system for critical infrastructure protection in smart cities via smart video-surveillance*", Pattern Recognition Letters, Volume 140,Pages 303-309,2020

[8] Hanbin Luo, Jiajing Liu, Weili Fang, Peter E.D. Love, Qunzhou Yu, Zhenchuan Lu, "*Real-time smart video surveillance to manage safety: A case study of a transport mega-project*", Advanced Engineering Informatics, Volume 45, 2020,

[9] Patel R., Dabhi V. K. and Prajapati H. B., "*A survey on IoT based road traffic surveillance and accident detection system (A smart way to handle traffic and concerned problems*)", IEEE in Innovations in Power and Advanced Computing Technologies (i-PACT), pp. 1-7, 2017.

[10] Aydin I. and Othman N. A., "*A new IoT combined face detection of people by using computer vision for security application*", IEEE in International Artificial Intelligence and Data Processing Symposium (IDAP), pp. 1-6, 2017.

[11] Othman N. A. and Aydin I., "*A face recognition method in the Internet of Things for security applications in smart homes and cities,*" IEEE, 6th International Istanbul Smart Grids and Cities Congress and Fair (ICSG), pp. 20-24. 2018.

[12] Premavathi C., Thangaraj P., "*Efficient Hand-dorsa Vein Pattern Recognition using KNN classification with Completed histogram CB in TP Feature Descriptor*", International Journal of Recent Technology and Engineering (IJRTE), Volume-7 Issue-4S, November 2018.

[13] Shao Z., Cai J., and Wang Z., "*Smart Monitoring Cameras Driven Intelligent Processing to Big Surveillance Video Data*," in IEEE Transactions on Big Data, vol. 4, no. 1, pp. 105-116, 1 March 2018.

[14] Deshmukh P. R. and Bhalerao D., "*An implementation of MQTT through the application of warehouse management system for climacteric fruits and vegetables,*" IEEE in 2nd International Conference on Communication and Electronics Systems (ICCES), pp. 844-849, 2017.

[15] Oliveira G. M. B., Danielly C. M. Costa, Ricardo J. B. V. M. Cavalcanti, "*Comparison Between MQTT and WebSocket Protocols for IoT Applications Using ESP8266*", IEEE Workshop on Metrology for Industry 4.0 and IoT, pp. 236-241,2018.

[16] Cruz-Piris, L., Rivera, D., Marsa-Maestre, I., De La Hoz, E. and Velasco, J., " *Access control mechanism for IoT environments based on modelling communication procedures as resources*" Sensors, vol. 18, pp. 917, 2018.

[17] Gündoğan, C., Kietzmann, P., Lenders, M., Petersen, H., Schmidt, T.C. and Wählisch, M., " NDN, CoAP, and MQTT: A Comparative Measurement Study in the IoT," *arXiv preprint arXiv:1806.01444*, 2018.

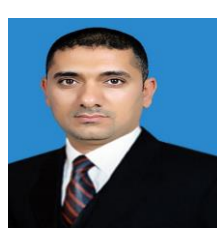
[18] Kachouane M., Sahki S., Lakrouf M. and Ouadah N., "HOG based fast human detection", IEEE, 24th International Conference on Microelectronics (ICM), Algiers, pp. 1-4 , 2012.

[19] Kumar A., Kaur A. and Kumar M., "*face detection techniques: a review*", Springer Nature B.V. 2018

[20] Setiawan E. and Muttaqin A., *"implementation of k-nearest neighbors face recognition on low power processor",* TELEKOMNIKA, Vol. 13, No.3, PP.949-954, September 2015.

[21] Wang J., Wang S. and Ji Q., "*Early Facial Expression Recognition Using Hidden Markov Models*", IEEE 2nd International Conference on Pattern Recognition, pp. 4594-4599, 2014.

[22] Yanhui Guo, Siming Han, Ying Li, Cuifen Zhang, Yu Bai, K-Nearest Neighbor combined with guided filter for hyperspectral image classification, Procedia Computer Science, Volume 129, Pages 159-165,2018.



Ali H. Hamad: Faculty of Information and communication engineering department university of Baghdad since 2003. He received his Ph. D in control and systems engineering university of Basrah Iraq. M. Sc. and B. Sc. from university of Technology Iraq 2000 and 1997 respectivly. Research iterests include IoT, network security, wireless sensor network, cryptography etc.