Vibration Guided Automatic Vision for Enhanced Security

Ipsita Sanyal, K. R. Dhavana,Kailash T. V., Kruthika R., Bhavanishankar K.

Abstract: ***The existing security systems are secure but are not smart enough to handle arbitrary scenarios leading to many false triggers of the alert system. Furthermore, these systems require constant human intervention which isdifficult to achieve.They are also vulnerable as they contain many loopholesand the sensors used are easily manipulatable. The proposed system tries to solve this problem in an efficient and a smart way by the use of sensors, AI and IoT which makes the system robust and resistant againstattacks. The system implements advanced face detection via Single Shot Detection and face recognition via Inception Neural Network for recognition of object in a fast and accurate way. This helps the system act according to the situation, thus preventing any damage to theregion which implements this system. In this work the proposed system is implemented and tested as a Home Security System. The system can also be extended to work in other areas like banks, data hubs, museums etc.The overall accuracy of the system was recorded to be 97.95%.***

Keywords: CNN, Inception Neural Networks,Internet of Things, security systems, recognition.

I. INTRODUCTION

In the current era where the activities of terrorism and theft are prevalent the need for an efficient and unassailable security system for both highly sensitive and low sensitive areas is indispensable. With the connectivity and technological advancements there prevails a risk of misusing the power of these technologies in unethical task like theft, etc., hence there is an absolute need for implementing security in low sensitive areas like hospitals, apartments, business organizations, datacenters, homes, etc. Surveillance is instrumental in the implementation of an efficient security system.The process of closely monitoring an activity, person or an area is known as surveillance. The state-of-the-art monitoring system calls for constant monitoring and human monitoring is not efficient enough hence video surveillance and automatic analysis can be cited as the best solution. The present-day video surveillance systems do have limitations owing to the fact that many a times in order to have a 360° view of an area innumerable Closed-CircuitTelevision (CCTV) are deployed which invariably increases the cost of implementation.

Revised Manuscript Received on December 15, 2019.

Ms. Ipsita Sanyal, Department of Electronics and Communication Engineering, RNSIT Bangalore India. Email: ipsita.san1@gmail.com

Ms. K. R. Dhavana, Department of Electronics and Communication Engineering, RNSIT Bangalore India. Email: dhavana98@gmail.com

Mr. Kailash T. V., Computer Science Engineering, RNSIT Bangalore India. Email: kailashtalanki@gmail.com

Ms. Kruthika R., Department of Electronics and Communication Engineering, RNSIT Bangalore India. Email: kruthika2599@gmail.com

Dr. Bhavanishankar K., Assistant Professor, Department of Computer Science, RNSIT Bangalore. Email: bsharsh@gmail.com

There is significant amount of research done to improve the security systemsin order to make them invincible. Though the aforementioned issues can be addressed by the spynel 360° thermal sensor, but it does have limitations such as the cost of the camera is exorbitant and the data recorded needs to be processed and analyzed which calls for a significant computational power, thus for low sensitive areas such high cost and computationally expensive solutions are non-essential or optional.

Through this workthe aforementioned issues were addressed by implementing a solution for security, surveillance and automatic analysis with the help of Artificial Intelligence (AI) and Internet of Things (IoT), which is an economically viable solution. A brief description of the proposed work is as follows, Vibration sensors respond to repetitive mechanical motion. If a movement is detected in the area of surveillance, the vibration is sensed by the sensor producingan output. The microcontroller is programmed to monitor the sensor. The servo motor is mapped corresponding to the coordinates where the vibration sensor is implanted. Based on the direction of the sensed vibration, the servo motor is rotated. A Raspberry Pi camera is mounted on the servo motor, through which an android application fetches live stream video and displays it. This live stream is achieved through Wireless Fidelity. The captured video is analyzed in real time. Face detection is done and the obtained face data is then sent through face recognition and based on the results, alert system is activated.

II. RELATED WORKS

In the home security system implemented in [1],Arduino and Infrared sensors were used to detect any intruder and GSM module used to dispatch alerts. Work in [2] uses the CCTV footage for monitoring, implemented using Arduino, GSM and Camera. The model in [3] recorded an accuracy of 90.6%. Light Detecting Resistors were installed on the walls to detect a crossing across the region. If so, CCTV camera turn towards it and the alert system is triggered. Arduino Uno was used in Security Model [4] built for house monitoring which made use of fingerprint sensor and pass code system to access internal doors and a mobile application for the main door.

A live video feed is sent over to the user on detecting any motion and is also saved with a timestamp for future reference in [5]. A 2seconds to 6seconds delay was recorded on high network traffic. A PIR sensor and microcontroller TI-CC3200 were used in [6], which assist in the detection of human motion and alerts via Voice calls and alarm system which turns off after a certain delay.

Necessary actions can also be taken by the user using dial pads which map to specific actions. The camera captures the images of the subject on detecting the motion in [7], on recognizing the unauthenticated person the alert is sent through mail and alarm system is activated. Here, Yolo algorithm is used for intruder detection. The model [8] embarks Haar Cascade algorithm for face detection and recognition. The live video feed is recorded and based on the recognition results, the alert system is activated. The approach in [9] enables the user to rotate and move the camera according to his will. The PIR sensor is used for motion detection and a GSM module is instrumental in dispatching the alert to the user.

In [6] there are no usage of cameras which is a limitation because the user will not be able to perceive whether the person at the door is a known or unknown person. In the works [2][3][5][9] the camera is instrumental only in capturing or live streaming of videos, but there is no provision of automatic monitoring and analysis of the recorded videos henceforth requiring human intervention. In [1][3][5] the sensors used in the security systems are not capable of distinguishing the stimulus caused by the humans and other entities ,and hence responds to their respective stimulus irrespective of whether it is caused by humans or other entities like stray animals(including cats, dogs, etc.),which eventually leads to the false alert generation .In the works[1][2][3][4][7][8][9] ,in situations where there is a potential case of theft or suspicious activities the system merely alerts the user by sending an alert notification or by activating an alarm but is not smart enough to act accordingly to prevent such circumstances by itself and waits for the user's actions or decisions. The aforementioned issues are by and large addressed by the security model proposed in this paper.

III. PROPOSED SYSTEM

The block diagram of the proposed work is as in Fig 1. The working principle of each block is explained below.

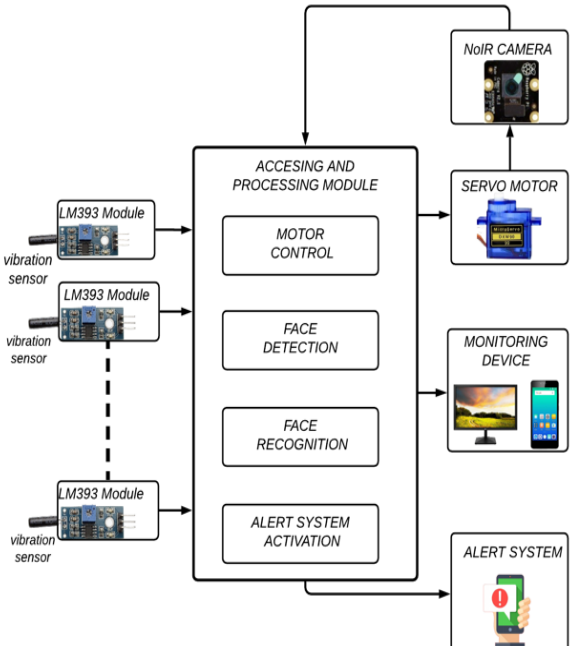

Fig. 1 Proposed system.

Fig. 2 Vibration sensor.

The Vibration sensor (SW18010p) consists of a small spring inside a metal casing. The sensor is integrated with a module which predominantly consists of a potentiometer and LM393 IC. The IC consists of two operational amplifiers (Op-amp), Op-amp1 and Op-amp2. The potentiometer is connected to the non-inverting terminal of the Op-amp1, which is instrumental in providing the reference voltage. It is also used to vary the sensitivity of the sensor. The output voltage of the sensor is given to the inverting input terminal of the Op-amp1 through a voltage divider circuit.The voltage division formula is as shown in (1).

$$V_{out} = \frac{R_2 \times V_{in}}{R_1 + R_2} \qquad (1)$$

Where,

$R_2$is the resistor across which the output voltage is to be measured

R_1 is the resistor in parallel to R_2

V_{in} is the input voltage

The output voltage of the op-amp is calculated by the formula represented in (2).

$$V_{out} = V_+ - V_- \qquad (2)$$

Where,

V_+ is the input voltage at the inverting terminal

V_- is the input voltage at the non-inverting terminal.

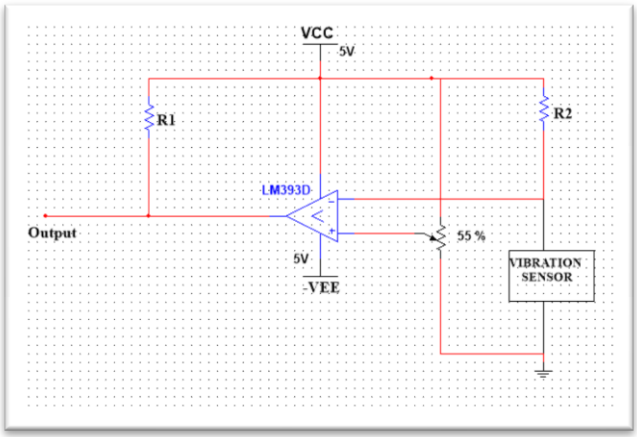

Fig.3 Circuit diagram of LM393 interfaced with SW-18010p.

Initially, when there is no stimulus to trigger the vibration sensor the voltage across the inverting terminal is higher than the reference voltage at the non-inverting terminal and hence the output of the op-amp is in the negative voltage range. When an appropriate stimulus is given to the sensor,

the voltage at the inverting terminal is lesser than that of the non-inverting terminal hence the output is in the positive range.

The analysis and processing module used in the proposed work is Raspberry-pi. It continuously monitors the vibration sensors, acquires data and processes it.

The camera module used is Pi NoIR Video Camera which is capable of performing efficiently during both low and high illumination conditions. The face detection and recognition algorithm is implemented using Convolution Neural Networks (CNN). CNN is implemented through transfer learning due to the limited size of the availableDataset. The reason we propose Deep Neural Networks (DNN) for face detection and recognition is that, through DNN more complex features of the face can be extracted [10]. The face detection technique in the proposed work is implemented by using Single Shot Detection (SSD) [11] usingVGG-16. The SSD model is a feed forward convolution neural network which is instrumental in generating bounding boxes around the object and scores of the presence of an object in the detected bounding boxes and the ultimate decision is made by the non-maximum suppression.

The face recognition implemented in this work is based on the FaceNet [13], the CNN architecture used is the inception model, also known as GoogleNet [14]. The inception net model involves the parallel processing of multiple layers of convolution and pooling. The outputs from the various parallel layers are concatenated which is the output pertaining to a single layer of the inception model. The inception model provides for a drastic reduction in the number of trainable parameters [13]. This feature enables the usage in applications where the computational power and resources are limited. Since the proposed work uses Raspberry-pi which has a limited computational power,hence it justifies that the implementation of the inception model for face recognition is an unparalleled and a perfect selection for the proposed work.

The FaceNet model is pre-trained on the Labeled Faces in Wild(LFW) dataset consisting of 13000 face images of 1680 persons [14].The weights that are learnt by the parameters on the LFW dataset is unaltered. Since the generated dataset is very limitedhence only forward propagation is performed to extract the encoded feature vector of size (128,1).The feature vectors are extracted for all the images present in the dataset and is compared with the feature vector extracted from the face image which is captured by the camera. The feature vectors of the dataset and the face image feed from the camera are then compared by computing the Euclidean distance, as follows:

$$E(b,a) = \sqrt{\sum_{i=1}^{n}(b_j - a_j)^2} (3)$$

Where,

b_j is the j^{th}weight in the feature vector b,

aj is the jth weight in the feature vector a.

The GoogleCloud IoT platform is used to establish communication between the Raspberry-pi and the actuators of the lock systems. The data uploaded to the cloud is modified to device configuration and then communicated to the device over the Message Queuing Telemetry Transport (MQTT) network protocol hence triggering the necessary actions.

IV. METHODOLOGY

When the sensor detects vibration, a voltage is generated when its spring comes in contact with the outer wall. The generated output voltage is sent to Raspberry-pi for further course of actions to take place. The Raspberry-pi activates the servo motor and the Pi-Camera module helping it to point to the region of motion.The feed from the Pi camera is continuously monitored and split up into multiple frames which are extracted at therate of 5 frames per second. The face detection algorithm is implemented on each of the extracted frames and the frame containing the maximum count of faces is selected for further processing. Here each face detected is cropped which is then subjected to face recognition.

The flow of events is represented as in the Fig 4.

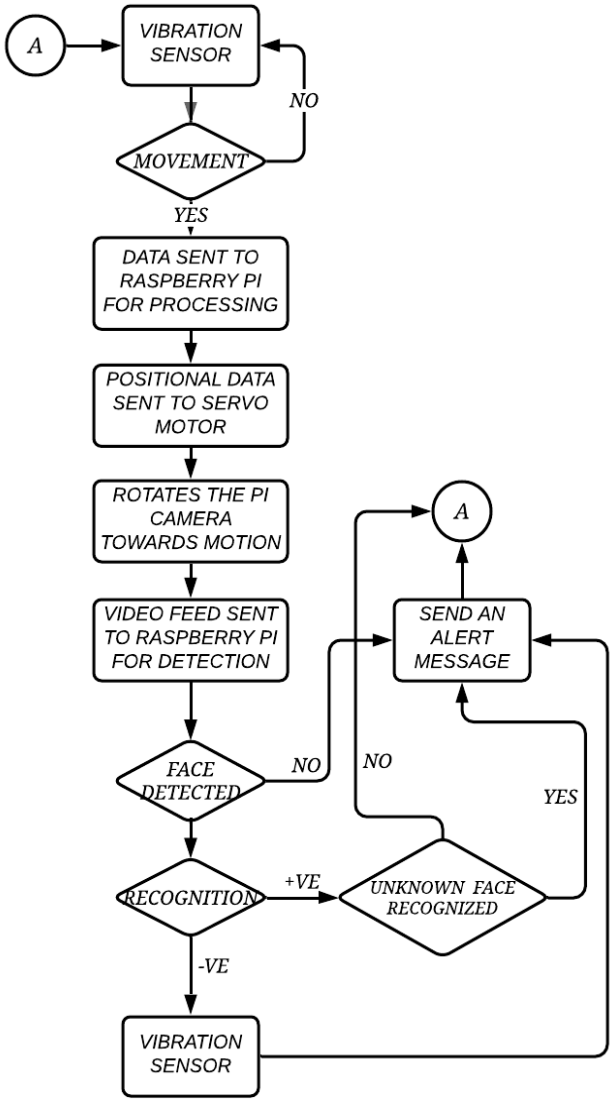

Fig 4 Flow chart of the proposed methodology.

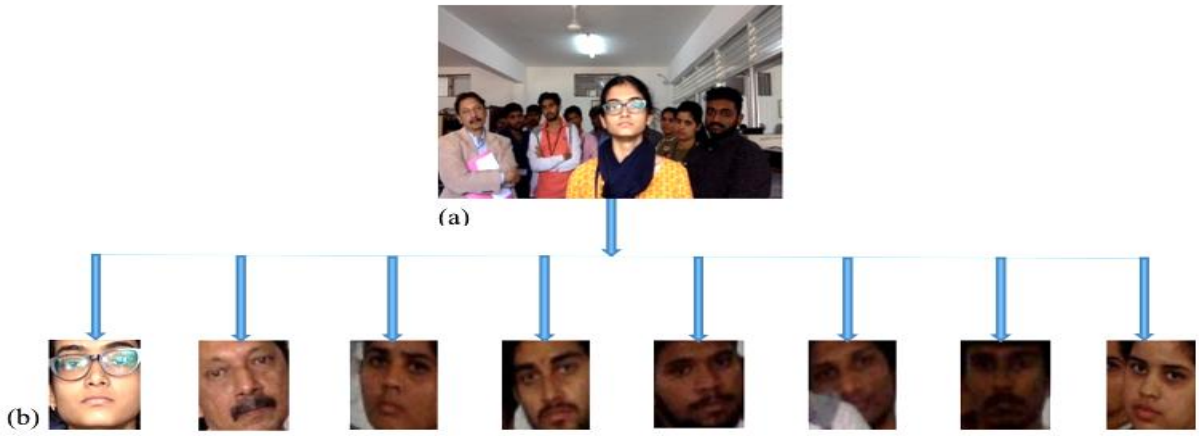

Fig 6 (a) Image captured through the Pi Camera, (b) Detected faces.

The SSD model implemented in this work is pre-trained on the COCO database. The Stochastic Gradient Descent (SGD) optimization algorithm is used to train the SSD model. The algorithm involves the learning rate to be 0.0001, momentum to be 0.98 and a weight decay of 0.0006. A mini batch size of 64 images is used for the training of the model.

The FaceNet model used for recognition is pre-trained by the SGD optimization algorithm. The hyperparameters are maintained same as that of the SSD. The threshold value of 0.48 is set for recognition. Ifthe calculated difference is less than the threshold, then the person is recognized elsethe person is labeled as "unknown" which triggers the alertsystem.

The alert system mainly consists of two aspects, the first being an alarm that activates when an unknown person or an entity is causing any movement in the surveilled area. The system also dispatches the intrusion information via a Short Message Service to the concerned user. The second aspect to the alert system is that, on detecting any intruder, a complete lockdown of the area is done by causing all possible exits inaccessible. The locks of the doors and the windows are bestowed with ability to communicate with the Raspberry-pi which on detecting an unknown person will activate all the lock systems.

V. EXPERIMENTATION AND RESULTS

An exemplary implementation of the proposed work was cognized and demonstrated. Different scenarios were exercised. The misconstrued conditions such as environmental changes or small gradual movements caused by external factors or entities other than humans, are effectively eradicated by the proposed work. Infrared cameras were used since they can perform in low-light conditions as well. The sensors were typically placed at easily accessible windows as well as doors that lead to and from the house. Standard face databases of family and friends were created. Three pictures of each person were collected in differently illuminated conditions as shown in Fig 5. Face detection followed by feature extraction were applied to the images of the generated database, which was further used to extract the similarity scores of the images collected from the feed of the PI Camera by computing the Euclidean distance.

Fig5Images of a known person captured at three different illumination conditions for database creation.

To test the working and the accuracy of the system, four experiments were conducted under varied scenarios, when the vibration sensors were triggered with a stimulus.

Scenario-1: Here the Pi-Camera is activated and points to the specific region of stimulus, recording the live video feed. This scenario is as represented in the Fig 6(a), where the captured image contains multiple persons. The face detection algorithm extracts the face detected as shown in Fig 6(b) which is then sent through for the face recognition. One among the detected faces was classified as "known" and the other images were classified as "unknown". Though there is a known face, due to the presence of unrecognized faces, the system sends a notification alert to the respective user of the system.

Scenario-2:Here from the feed of the Pi-Camera only known faces were recognized which were present in the created database. The system is smart enough to differentiate the situation to be a normal one and takes no action, thuspreventing any false triggering of the alert system.

Scenario-3: Here the faces detected and recognized were not present in the created database, thus indicating an intrusion. The systemintelligently recognizes this and initiates a lockdown of all the exits out of place, capturing the intruders. The system also sends an alert about the situation to the registered user. The user is accessible to a live stream of the video through IoT.

Scenario-4: Here the system detects no faces. This was because the trigger was due to an animal. The system on detecting no faces considers the trigger to be caused by a non-human entity and also sends a notification alert to the registered user warning them about the same.

The overall accuracy of the proposed system depends on the sensitivity of the vibration sensor, the classification accuracy of the face detection and recognition model and the network latency. Based on the distance the sensor senses the movement from, it outputs different voltage values as recorded in Table I.These voltage values are received by the Raspberry-pi in digital form, which are then converted to analog values.Based on these values a threshold value of 0.586 is set as seen in Fig 7, which identifies a genuine or false trigger i.e. when the output value is above this threshold indicating a genuine trigger. The system recorded an accuracy of 97% to distinguish between a genuine and false trigger.

Table I Output value of sensor corresponding to the distance of the stimulus.

Distance from sensor(m)	Digital output	Analog outputfrom the sensor (V)
0.1	455	2.222
0.2	377	1.842
0.3	331	1.632
0.4	247	1.207
0.5	120	0.586
0.6	53	0.259
0.7	26	0.063

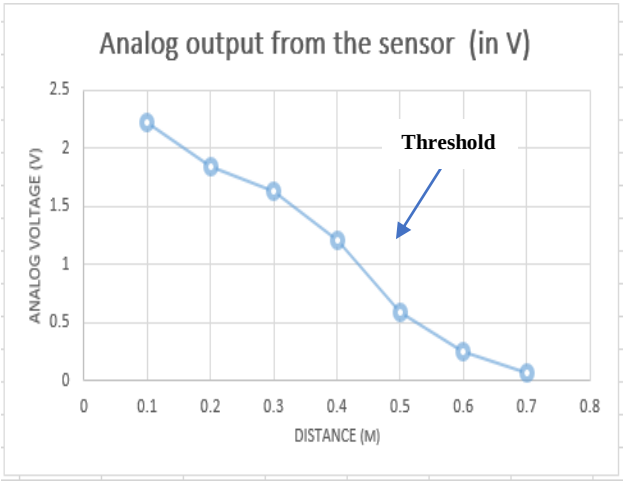

Fig. 7 Distance vs Analog output.

In order to determine the accuracy of the face detection and recognition modelfour scenarios were carried out recurrently under varying illumination conditions, varied poses of the subject, and the condition when the face was partially or completely occluded, the accuracy of the proposed security system under these conditions is as shown in Table II. It was observed that the model performed quite proficiently irrespective of varying illumination conditions with an accuracy of 99.41% and when there were diverse poses of the entity during image capture an accuracy of 98.1% was recorded, but there was a significant drop in the accuracy in the instance of occluded faces. The model witnessed an accuracy of 100% under ideal conditions i.e. at perfect illumination conditions and when the image captured is in the same orientation of the image as recorded in the dataset.

Table II Face detection and recognition accuracy under dissimilar circumstances.

Experimentation	ACCURACY (%)		
	Varied Illumination	Varied Poses	Occluded Faces
Scenario-1	98.8	97.63	35
Scenario-2	98.63	96.12	34.67
Scenario-3	98.32	98.1	-
Scenario-4	99.41	96.59	-

The overall performance of the detection and recognition model is determined by computing the average accuracy from Table II. Comparison between the ideal, existing and proposed detection and recognition model is as shown in Fig 8.

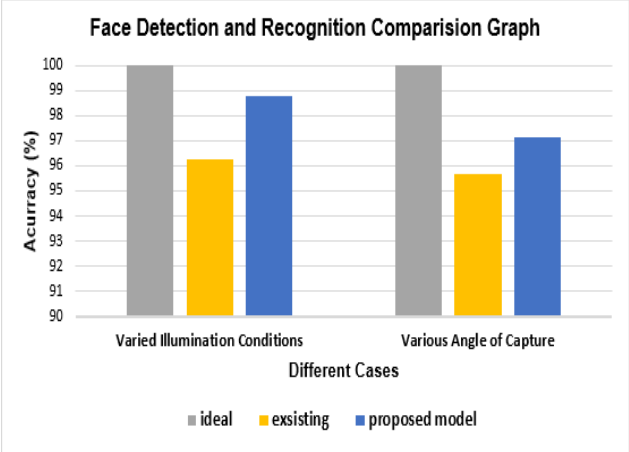

Fig. 8 Comparison of face detection and recognition model.

The overall accuracy of the proposed security system is determined by computing the average of the performances of the vibration sensor, the face detection and recognition model. Hence the overall performance of the proposed model is recorded to be 97.95% taking into consideration all the aforementioned scenarios.

VI. CONCLUSION

The issues discussed in section 2 of this paper were successfully bridged by the proposed work. The proposed security system recorded an average accuracy of 97.95 %. The experimentations carried out in this work was based on home environment but can also be customized to extend its applications to other commercial areas.Since the security system's accuracy exhibited a drastic fall in case of occluded faces of known people hence to improve upon it, is the future scope of the proposed work.

REFERENCES

1. Kaur, Surinder, Rashmi Singh, Neha Khairwal, and Pratyk Jain. "Home automation and security system." *Advanced Computational Intelligence 3*, no. 3 (2016): 17-23.
2. Memon, Kainat Fareed, Javed Ahmed Mahar, Hidayatullah Shaikh, Hafiz Ahmed Ali, and Farhan Ali Surahio. "GSM based Android Application: Appliances Automation and Security Control System using Arduino." *International Journal Of Advanced Computer Science And Applications* 8, no. 2 (2017): 206-210.
3. Çavaş, Mehmet, and Muhammad Baballe Ahmad. "Design and Simulation of Four Walls Crossed Security System against Intrusion Using Pic Microcontroller."*American Journal of Engineering Research* (AJER) Volume-7, Issue-12, pp-233-244
4. Hate, Mandar, Manjunath Gowda, and Kaustubh Kubal. "IoT baseWireless Security System."*International Journal of Research inEngineering, Science and Managemen*t Volume-2, Issue-4, April-2019.
5. Das, Somak R., Silvia Chita, Nina Peterson, Behrooz A. Shirazi, and Medha Bhadkamkar. "Home automation and security for mobile devices." In 2011 *IEEE International Conference on Pervasive Computing and Communications Workshops* (PERCOM Workshops), pp. 141-146. IEEE, 2011.
6. Kodali, Ravi Kishore, Vishal Jain, Suvadeep Bose, and Lakshmi Boppana. "IoT based smart security and home automation system." In 2016 International Conference on Computing, communication and automation (ICCCA), pp. 1286-1289. IEEE, 2016.
7. Kumbhar, Deepak S., Shubhangi M. Taur, and Shubhangi S. Bhatambrekar. "IoT Based Home Security System Using Raspberry Pi-3." (2018).
8. Vadivukarasi, K., and S. Krithiga. "HOME SECURITY SYSTEM USING IoT." *International Journal of Pure and Applied Mathematics* 119, no. 15 (2018): 1863-1868.
9. K.Pranathi, Nihar Munagala, D.Venkatesh, N.Pavankumar, A.Phaneendra. "Designing of Multifunctional Surveillance and Securitysystem". *International Research Journal of Engineering and Technology* ,04 volume,03 issue (2017).
10. Li, Haoxiang, Zhe Lin, Xiaohui Shen, Jonathan Brandt, and Gang Hua. "A convolutional neural network cascade for face detection." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5325-5334. 2015.
11. Liu, Wei, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. "Ssd: Single shot multibox detector." In *European conference on computer vision*, pp. 21-37. Springer, Cham, 2016.
12. Schroff, Florian, Dmitry Kalenichenko, and James Philbin. "Facenet: A unified embedding for face recognition and clustering." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 815-823. 2015.
13. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. CoRR, abs/1409.4842, 2014.
14. http://vis-www.cs.universityofmassachusetts.edu/lfw

AUTHORS PROFILE

Ipsita Sanyal is B.E.in Electronics and Communication from RNS institute of Technology, Bengaluru.She has been a activity coordinator of IEEE student chapter @ RNSIT. Her Area of interest includes AI, biometrics, pattern recognition, embedded systems.

K. R. Dhavana is B.E.inElectronics and Communication at RNS institute of Technology, Bengaluru. Her area of interest is embedded systems and IOT development and secured first place for implementing a work on "Li-Fi" during Open-House Expo conducted in the College.

Kailash T. V. obtained B.E. in Computer Science at RNS institute of Technology, Bengaluru. His domain of interest is full stack WEB/App development and has implemented cross platform apps like "BMI Calculator", "Multiple Currency Convertor" andiOS games like "ADD 1".

Kruthika R. is B.E.inElectronics and Communication at RNS institute of Technology. Her area of interest is VLSI, sensors and IOT development and secured first place for implementation of "A smart WatchBot".

Dr. Bhavanishankar K.– Working as an Asst. Prof. at RNS Institute of Technology in Department of CSE. His research area of interests includes image processing, machine learning, artificial intelligence, algorithms and data structures. He has published 8 research articles in International Journals. He is passionate about guiding creative technical projects.