A Multimodal IoT-Based Home Intrusion Detection System

Bhavesh Bhuckory and Sameerchand Pudaruth*

***Abstract*—The rise in crime rates has significantly proven the need for security in houses. Although various types of surveillance systems are already being used, most make use of single modes of detection which are insufficient to identify intrusions. This paper presents an affordable IoT based home security system that combines different ways to detect intruders. It has been implemented using a Raspberry Pi 4 device, an IP camera, a microphone, and sound and motion sensors. When people enter a monitored area, they are detected either through their movement, voice, or while knocking on the door. The number of individuals present is counted, followed by face recognition. The system can also recognise masked faces. The system uses a servo motor as a physical door lock, which is unlocked for authorized users. A speaker identification module is responsible for authenticating the voice of a person via the camera whenever the latter is heard but not seen. When a potential intrusion is detected, the footage is uploaded to the cloud and event details are transmitted to a cloud database. Owners are also immediately notified through messages and can access all the intrusion details on a web application.**

***Index Terms*—Security, object detection, face recognition, masked face recognition, speaker identification, smart lock, immediate notifications**

I. Introduction

Security is one of the most important issues when it comes to the safety of one’s family, properties, and valuable belongings. In the early days, people relied on the presence of watchmen and security guards to monitor their places and watch out for burglars. However, due to the evolution of technology, devices such as cameras and alarm systems can now be used, and thus the task of surveillance has become easier as it can now also be done remotely.

We are witnessing a significant increase in the use of such systems, due to the rise in crime rates [1]. In Mauritius, the overall offense rate (excluding road contraventions) per 1000 population has increased from 35.6 in 2019 to 43.8 in 2020. A total of 2818 burglary cases were recorded between 2019 and 2020 [2]. On a global basis, a burglary can occur about every 18 s, which amounts to 4800 cases daily [3]. Hence, with such alarming statistics, it can be inferred that home security is becoming a crucial requirement for homeowners.

Homes with security systems are 300 times less likely to get broken into than the ones without a security device [4]. The effectiveness of these systems is always in doubt, however, due to the number of crimes still occurring. In many burglar alarm systems, false alerts are often caused due to small movements by pets, changes in lighting, or even by strong winds [5]. This often results in false alarms being sent to security services. Thus, the trust in home security systems could be increased if the accuracy for the real-time detection of unauthorized individuals could be improved.

Another challenge faced by people is to continuously monitor areas to identify intruders through surveillance systems. Closed Circuit Television (CCTV) is one of the most used surveillance systems around the world. Currently, there are 770 million operational CCTV cameras worldwide, while 54% of them are situated in China [6]. Having the functionality of recording movements, such systems have contributed to the identification of many criminals. Many high-profile cases were aided by CCTV systems, including the Charlie Hebdo attack in 2015, where the suspects were caught on recorded footage [7].

However, in a home security system, camera surveillance would be more efficient if it could detect intruders automatically, rather than through passive monitoring. Even though smart security cameras are used in some places, most are based solely on motion detection [8]. Additionally, in many cases, there have been large gaps between the time of occurrence of a burglary and the time at which it was reported to the owners or authorities. This usually occurs when buildings are monitored remotely. Hence, a mechanism is required for automatic and immediate reporting of events.

According to a survey based in the US, 30.1% of homeowners use a security system without video, while 43.9% rely solely on security cameras [9]. Hence, only 26% of them used combinations of different home security devices. By combining various security devices together, the effectiveness and the degree of security in the house increases, and it also provides several ways to detect intruders or strangers. Incorporating noise detection using a microphone may alert the owner when intruders break-in. Moreover, additional sensors may aid the system to detect movements. Thus, it is important to make use of a multi-faceted security system rather than relying on a single security device.

The main objective of this work is to develop a home security system that automatically detects potential intruders using several devices and methods. This paper proceeds as follows. Section II provides an overview of related works on home security systems. The design and methods used for implementation are described in Section III. The results and their evaluation are discussed in Section IV. Finally, Section V concludes the paper.

II. Related Works

This section shows similar systems produced in the past along with the different devices and technologies used. Tanwar *et al.* [10] developed a system which implements a real time email alert mechanism. It uses Passive Infrared (PIR) sensors and cameras connected to the USB ports and General Purpose Input/Output (GPIO) pins of a Raspberry Pi device. The system detects motion by comparing the input signals of the PIR sensor to its previous values. If they differ, it means

Manuscript received September 28, 2022; revised November 23, 2022; accepted May 9, 2023.

The authors are with the Information and Communication Technology Department, Faculty of Information, Communication and Digital Technologies, University of Mauritius, Mauritius.

*Correspondence: s.pudaruth@uom.ac.mu (S.P.)

DOI: 10.7763/IJCTE.2023.V15.1340

that motion is detected. The camera then captures an image and stores it locally. The image is attached to an email and sent to the homeowner. This system is cost effective, but it can cause false alerts since it cannot differentiate between authorized and unauthorized individuals.

A home security system based on wireless data transmission was designed by Sahoo and Pati [11]. It is made of two circuits, namely a sensor and a central circuit. The sensor circuit forms a wireless sensor network by associating PIR sensors with Zigbee transceivers and Arduino nano processors. The central circuit is made of an Arduino Mega board, a Zigbee transceiver, a Wi-Fi module, and a GSM module. Every time motion is detected, the transceiver of the sensor circuit sends a signal to the transceiver of the central circuit which notifies the homeowner through his/her cell phone via text messages. Additionally, an IP camera is connected to the Wi-Fi module which records the event upon intrusion detection and uploads the footage to a cloud server, using ThingSpeak software. Thus, a user can access the images and videos anytime by logging into his/her account.

In many systems, computer vision has been used to increase the accuracy of intruder detection, as it aids to identify unauthorized individuals in monitored areas. Amato *et al.* [12] proposed a system which exploits facial recognition using deep learning to determine whether the detected persons' faces match a set of authorized faces from a database. The prototype is composed of a wireless camera associated with a Raspberry Pi 3 model B microcontroller. The camera captures a frame every 2.9 seconds, while the system continuously performs face detection using the Single Shot Detector RestNet-10 model [13]. For each detected face, deep features are extracted using the OpenCV DNN module. The distance between the queried face deep feature and all the deep features in the training is calculated to perform the k-Nearest Neighbour (kNN) classification. If the value returned by the classifier is below a given threshold, an email notification is sent to the owner which indicates a potential intrusion. The disadvantage of this system is that it requires high processing power and is also quite slow.

See and Lee [14] developed a camera surveillance system which consists of two separate mechanisms namely, a face recognition module, and a motion detection module. The face recognition is always performed actively on captured frames and, in case of any failure or malfunction, the motion detection module is activated. On the other hand, Pudaruth *et al.* [15] developed a system having both features working together. To achieve this, an Infrared (IR) camera is used in the system. Whenever the motion level exceeds a threshold of 15%, the face detection module is triggered. The motion level is simply determined by the difference between the current and previous frame. If the frame contains a person, the face recognition module is activated. The Eigen face recognition algorithm is used which compares the captured face to the images stored in the database. If a face is not matched, the homeowner is informed via an email attached with the captured image.

Another vision-based security system was implemented by Khodadin and Pudaruth [16] which uses an IP camera for intruder detection. Video frames are continuously generated and processed by several modules. Firstly, the YOLO algorithm is used for object detection. If a person is detected, the system triggers a person counting module. At the same time, a facial recognition module is performed first by using the Dlib toolkit to extract facial landmarks. It then uses the Histogram of Oriented Gradients (HoG) as its feature descriptor. The person's facial features are then compared to the stored ones. If an unknown person is detected, motion detection and object displacement modules are triggered. The object displacement mechanism allows the homeowner to know which objects might have been stolen or displaced. The event is recorded and stored on a cloud server, which the user can access remotely. Text messages, voice calls, email and WhatsApp messages are sent to the homeowner, attached with appropriate information.

Face recognition mechanisms have also been included in door security systems. Paul *et al.* [17] proposed a smart home security prototype which connects a door lock to a Raspberry Pi 3 device, along with a Pi camera. The camera continuously reads frames and performs face detection using the Haar cascade classifier. When a person appears in front of the door, his/her face is detected, and a face recognition module is triggered. The system uses a combination of neural networks and the Local Binary Patterns Histogram (LBPH) model to have a good accuracy level. The neural network is trained for several authorized persons. If a person is recognised, the Raspberry Pi signals the door to be unlocked. The person is then allowed to enter the house. If a stranger is detected, the homeowner is notified accordingly via his/her mobile phone. Moreover, the face recognition model can authenticate faces from different angles and emotions, and masked faces also. The LBPH algorithm requires low computing power and is thus suitable for Raspberry Pi microcontrollers.

Shahid *et al.* [18] developed an intelligent home security system which makes decisions based on data from sensors. The prototype includes an ultrasonic sensor, a PIR sensor, and a sound sensor, connected to a NodeMCU. An Arduino Uno is also added to the input module to provide an interface to an ArduCAM camera. Every second, sensor data, including human presence, sound levels and change in object distance are uploaded to a real time cloud server. Meanwhile, the server processing module, which is a separate Raspberry Pi, fetches the data from the database and uses logistic regression to predict whether an intruder is detected or not. The system presents a 97.5% accuracy for detecting intruders. It was trained using a large dataset of raw sensor values recorded when people were detected. Hence, if human presence is detected, a flag value is set. Following this update event, the Arduino MCU signals the camera to record a video, and face recognition is performed using a Support Vector Machine (SVM) classifier. In case the detected face is not identified, an email or SMS alert is sent to the owner.

Park *et al.* [19] designed a hybrid sensor-based surveillance system that combines an image and sound field sensors to detect intrusion within an indoor environment. The sound field sensor technology generally consists of a sound generator, a microphone, a speaker, and a signal processor [19, 20]. A multi-tone sinusoidal sound is firstly generated, and then the microphone measures this sound. At the same time, real sound in the surrounding environment is also recorded. To detect an intrusion, a Signal-to-Noise Ratio (SNR) algorithm is applied, where the signal represents the difference between the multi-tone sound and the real sound pressure level, and the noise represents the maximum deviation during the multiple measurements of the generated sound pressure level. If the SNR is higher than a specified threshold, the system turns on a light, as well as the camera, so that the object can be tracked. The correlation, filter-based Minimum Output Sum of Squared Error (MOSSE) is used for the tracking purpose. Once the person has been detected, the system triggers an alarm. The problem with this system is the high cost of the sound field sensors. Also, if it detects high pitch noises, such as wind or birds chirping, it might create false alerts and thus cause the camera to operate unnecessarily. Moreover, a frequency-based system was also designed by Goncalves [21], whereby intrusion was detected if a generated wave was obstructed by a person. The system uses multiple millimetre wave (mmWave) radars to achieve this functionality. However, this mechanism is only suitable for highly restricted areas, as it is very sensitive to movements.

Home security systems have evolved from those requiring manual intervention to automated ones. Traditional systems were mostly based on CCTV cameras with local storage. Newer systems have emerged which combine cameras and sensors, as well as different storage options, such as cloud servers. By introducing computer vision techniques, surveillance systems have become more sophisticated in terms of avoiding false alarms and recognizing authorized users. Furthermore, the combination of these intelligent camera systems, along with microcontrollers and sensors is usually more efficient in terms of power consumption and security. Some prototypes are also associated with microphones to exploit noise levels in the monitored area to detect movement or strangers. However, we have seen that there are a limited number of systems which combine different sensors and entities together intelligently to produce the best solution. Thus, this work presents a system which detects potential intruders through their movement, voices, or by facial authentication.

III. Methodology

This section describes the design and architecture of the system to be implemented. Fig. 1 shows the flow of execution of the home security system, from the moment of monitoring until the delivery of notifications.

The system runs on a Raspberry Pi 4 device, whereby motion, noise and voice detection are performed continuously. When a user is detected through any of these methods, an image is captured and processed to identify the person as well as recognize his/her face to prevent any false alerts. The system also caters for a masked face recognition process in case the person is wearing a mask. The system also caters for situations where a voice is detected but the person is not present in the range of the camera. In this case, a speaker verification module is triggered to authenticate the voice. Whenever an intrusion occurs, the footage is sent to the cloud and the owner is informed immediately. All the event details are accessible on a web application as well. The cost of setting up this system is about 200 USD which includes the cost of running the web services on the cloud.

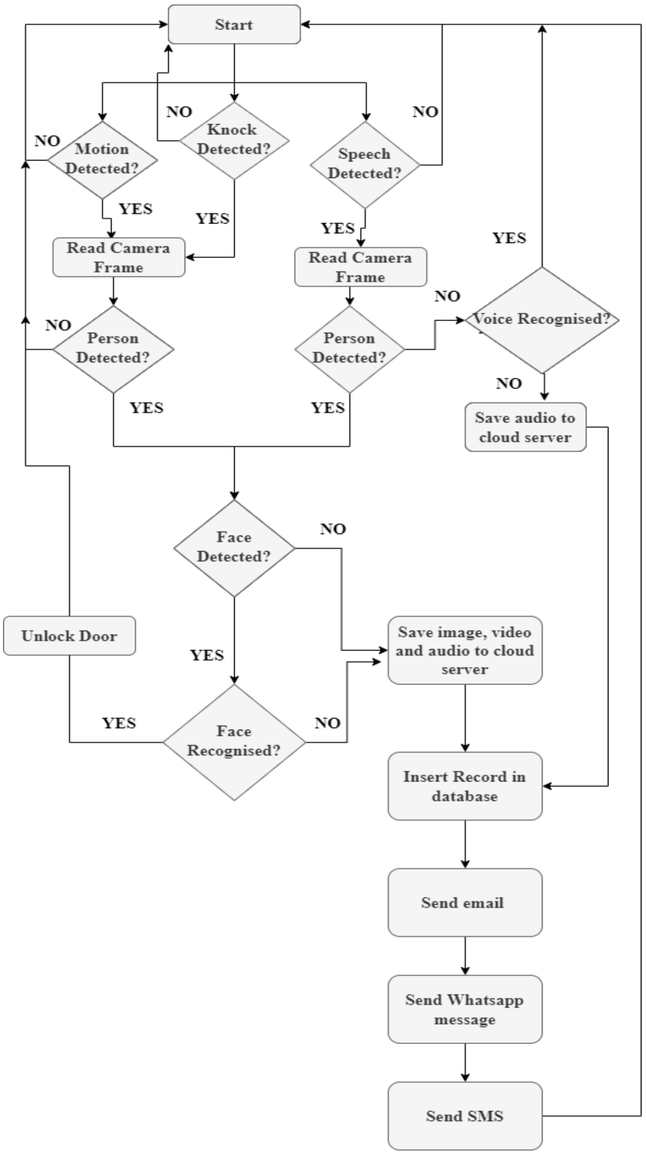


Fig. 1. Methodology of the home security system.

A. Motion Detection

To detect movement in the monitored area, Passive Infrared (PIR) sensors are used. This sensor detects infrared radiation from objects within its area using a pyroelectric sensor. Since the Raspberry Pi device is purely digital, a series of 0s and 1s are read, where the 1s indicate movement and 0s indicate no movement.

B. Obstacle Detection

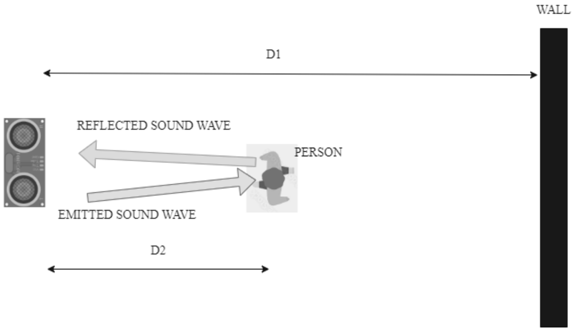


Fig. 2. New distance recorded in the presence of a person.

An ultrasonic sensor is used to detect obstacles in the monitored area. It is an electronic device that emits sound waves using piezoelectric crystals and absorbs the reflected ones on its receiver. Using the time taken and speed of ultrasound, the distance to the closest obstacle can be measured. It has a range of 3 meters. The ultrasonic sensor must be calibrated in an empty room in the beginning, and Fig. 2 shows how the distance value changes when a person appears in front of the sensor.

C. Knock Detection

A piezoelectric sound sensor has also been used in this system. It is placed on the door surface to detect vibrations. When someone knocks on the door, the sensor detects it. It operates similarly to the PIR sensors, as it outputs a series of 0s and 1s whereby the 1s represent vibrations. It has been calibrated to detect surface noises only.

D. Voice Detection

By only detecting noises, it is not enough to determine whether humans are passing by the monitored area or not. Hence, to make the noise detection more accurate, the system should be able to differentiate between speech and non-speech sound. An alert is thrown only if a speech sound is detected. For this feature, the pretrained Voice Activity Detection (VAD) module of the SpeechBrain library has been used [22]. Fig. 3 shows the flow of execution of the model when an audio file is passed to it.

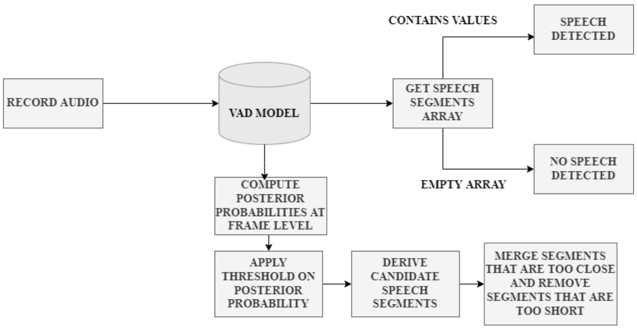

RECORD AUDIO
VAD MODEL
GET SPEECH SEGMENTS ARRAY
CONTAINS VALUES
SPEECH DETECTED
EMPTY ARRAY
NO SPEECH DETECTED
COMPUTE POSTERIOR PROBABILITIES AT FRAME LEVEL
APPLY THRESHOLD ON POSTERIOR PROBABILITY
DERIVE CANDIDATE SPEECH SEGMENTS
MERGE SEGMENTS THAT ARE TOO CLOSE AND REMOVE SEGMENTS THAT ARE TOO SHORT

Fig. 3. Voice activity detection process.

E. Person Detection

The system first checks whether a person is present in the surroundings as the MobileNet SSD model is already trained to detect several objects including people [23]. Fig. 4 shows the Single Shot Detection (SSD) process which is applied to the captured frame. For each detected person, a count is incremented.

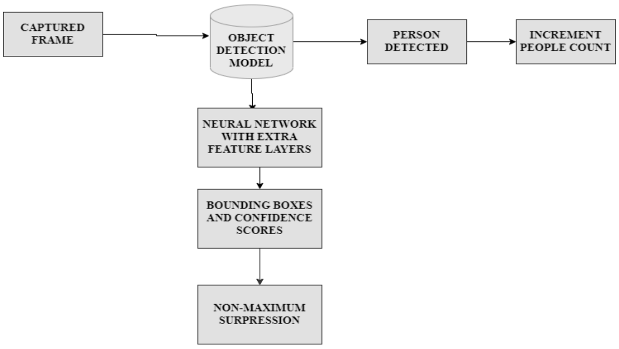


Fig. 4. Object Detection process.

F. Face Recognition

To prevent false alerts, the system should be able to differentiate between the homeowners and intruders. Thus, a face recognition module has been implemented. Generally, it consists of two parts, one of which also caters for masked faces. Fig. 5 shows how the processes are carried out.

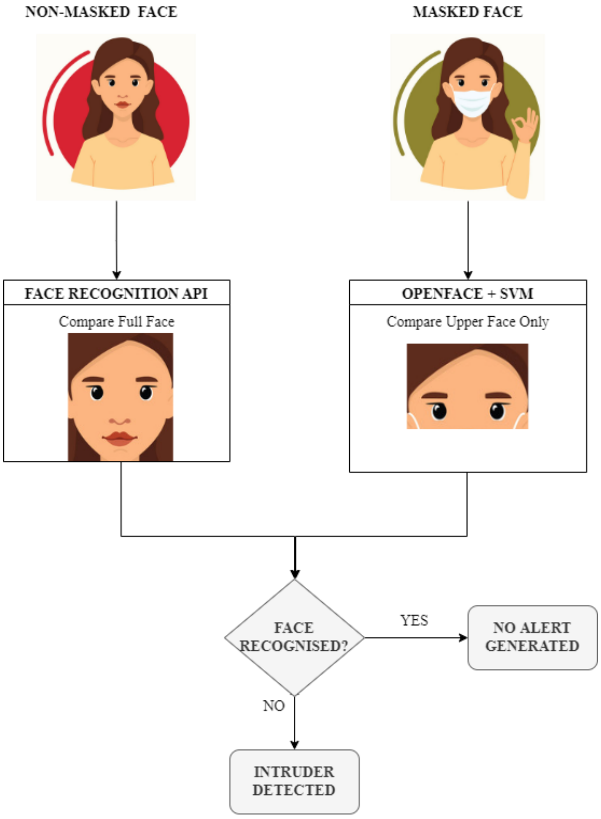


Fig. 5. Face recognition system.

When full faces are detected, the face recognition API uses Euclidean distance to compare the detected face to the stored ones. Using the cosine distance, it determines the degree of similarity between them. Moreover, the dataset also contains a set of masked faces, so that the users can be recognized while wearing their face masks. Since they cannot be detected using the face recognition API, the OpenFace toolkit is used in combination with the SVM classifier to formulate a supervised learning solution.

The OpenFace toolkit was used to extract the upper face features of the dataset [24]. SVM was adapted to masked face recognition by formulating the binary classifier using the extracted features against the names of the individuals. The Scikit-Learn Python library was fed with folders of known users' images, labelled with their names, while the folder of unknown users was filled with 100 random faces. In this way, the model was trained to follow a specific pattern for everyone, whereas there was no fixed pattern for the unknown ones. Therefore, whenever a masked face does not match a known person's class, it automatically falls in the category of unknown individuals.

G. Door Unlocking

Whenever a person is verified as an authorized user, a servo motor rotates 90 degrees, indicating that the door is unlocked.

H. Speaker Verification

If people are not detected by the camera after the microphone detects voices, the audio is passed through a pretrained Time Delay Neural Network (TDNN) model provided by the SpeechBrain library to verify if the speaker is unknown to the system [22]. It basically calculates the cosine similarity between the stored audios and the new audio.

I. Cloud Storage

When an intrusion is detected, the captured image, a short video and a recorded audio are uploaded to the cloud. The Dropbox API is used for this purpose. In addition, the details are stored in the MongoDB cloud database.

J. User Notifications

Users are immediately notified through SMS, email, and WhatsApp messages. The use of multi channels for notifications makes sure that the user sees at least one of them, regardless of where he/she is.

IV. Results and Evaluation

This section aims to discuss the implementation and testing of the home security system. Firstly, the different components of the systems, hardware and software requirements and additional tools are described. Then, the components are tested and evaluated to validate their functionalities. The system was implemented using the hardware specified in Table I.

TABLE I. Hardware Specifications

Device	Raspberry Pi 4 Model B+
RAM	2GB
Storage Capacity	8GB
Camera	Full HD 1080p, Infrared IP camera (2.4 GHz Wi-Fi based)
Sensors	2 × PIR sensors (7-Metres range)
	1 × Sound sensor used with least sensitivity (30 cm range)
	1 × ultrasonic sensor (3-Metres range)
Additional tools	1 × Passive buzzer (2 V), LED bulb, OLED Screen, Servo Motor
Sound Recorder	USB Microphone (4 metres range and 192kHz/16BIT resolution)

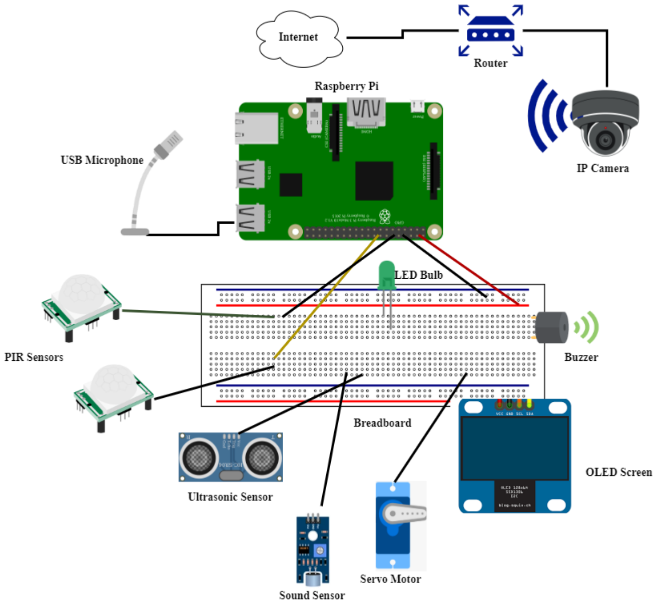


Fig. 6. System architecture.

After combining and assembling all the devices, a well-functioning system was obtained as shown in Fig. 6.

All the components of the system have been tested in different scenarios. For the motion and obstacle detection module, four different people were asked to walk in front of the sensors from various directions. As shown in Table II, all of them were detected either by movement, or as obstacles.

TABLE II. Motion and Obstacle Detection Testing

Walking Direction	**Motion/Obstacle Detection**			
	Person 1	**Person 2**	**Person 3**	**Person 4**
Towards Sensor	True	True	True	True
Away from sensor	True	True	True	True
Left to right	True	True	True	True
Right to left	True	True	True	True

The noise detection system was tested with various types of sounds including those containing human speech and door knocking noises. The Voice Activity Detection (VAD) model performed well as it could identify human voices easily and was able to discard other types of noises. The sound sensor was also activated whenever the door was being knocked. Table III shows the results of the tests carried out.

TABLE III. Speech Detection Testing

Sound	**Speech Detected**	**Door Knock Detected**
Talking	True	False
Singing	True	False
Shouting	True	False
Dogs barking	False	False
Cats meowing	False	False
Birds chirping	False	False
Background noise	False	False
Door Knocking	False	True

The camera was set to deliver 20 frames per second (fps). Since the camera operates wirelessly, there is a small delay in receiving the frames. Hence, some frames are skipped in the image queue to obtain the most recent one. The person detection model was tested by different users and in different light conditions as well. As shown in Fig. 7, the model works well even in dark places. It has been seen that the maximum distance for detection is 3 m to 4 m away from the camera.

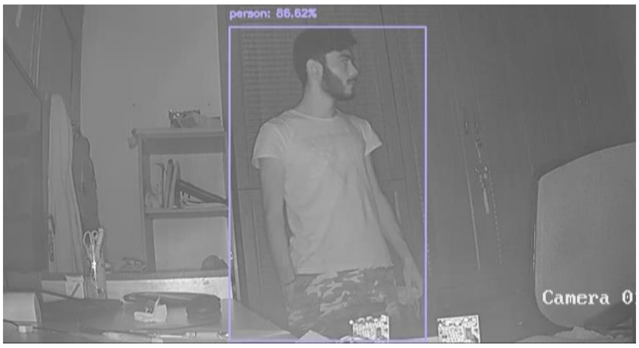


Fig. 7. Person detection.

The face recognition system was tested with images of the authorized users as well as with unknown ones, under different lighting conditions. Since this system requires heavy resources to work, the processing was done on a local server. The Raspberry Pi device sends images wirelessly to the server, which, in turn, applies the models and returns the result. It performed well as it could properly recognize the homeowners. In addition, the intruders were detected since their faces were unknown to the system. Masked faces were also correctly recognised by the system as shown in Fig. 8.

However, the system fails to recognize the authorized user if he/she is too far from the camera, or his/her face is not facing the camera directly. The maximum distance for faces to be detected by the model is 1 m to 2 m away from the camera.

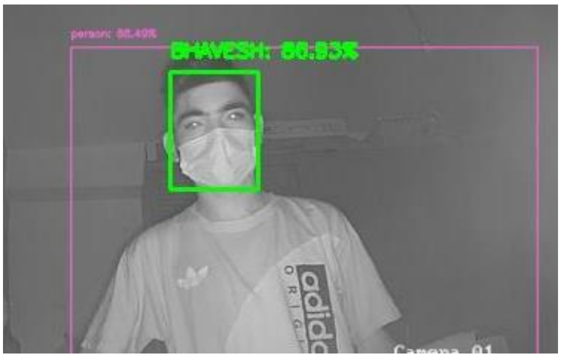


Fig. 8. Face recognition with mask.

TABLE IV: FACE RECOGNITION TESTING

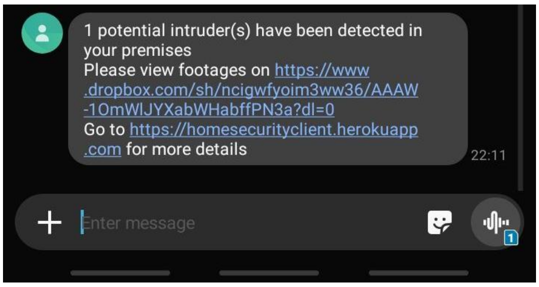
Model	Masked/ Unmasked faces	Description of Person	Final Output	Accuracy (%)
Python Face Recognition API	Unmasked	Known user 1	Recognised	90.1
		Known user 2	Recognised	88.3
		Known user 3	Recognised	87.8
		Known user 4	Recognised	89.2
		Unknown user 1	Unrecognised	88
		Unknown user 2	Unrecognised	85
		Unknown user 3	Unrecognised	86.5
		Unknown user 4	Unrecognised	85.6
OpenFace + SVM	Masked	Known user 1	Recognised	90.1
		Known user 2	Recognised	83.3
		Known user 3	Recognised	81.8
		Known user 4	Recognised	83.2
		Unknown user 1	Unrecognised	81
		Unknown user 2	Unrecognised	83
		Unknown user 3	Unrecognised	81.5
		Unknown user 4	Unrecognised	82.6

To test the accuracy of the face recognition module, four homeowners and four strangers walked in front of the system, at around 1m to 2 m from the camera. Table IV shows the results obtained. All the faces were detected, but the identification was not perfect in about 20% of cases. The speaker recognition system was tested with the voice of the owner as well as the unknown individuals, while all of them were away from the camera's field of view. The system successfully identified the owner and did not recognise the unknown people. The system sometimes fails to recognise the owner a few times whenever he/she is far from the microphone or is talking in a low voice. The maximum distance for voices to be recognised correctly from the microphone is 2 m. When using the voice together with face recognition, the accuracy in identifying the owner jumps from 80% to about 90.1%. Thus, the system is much more effective in identifying an intruder. With this combination, there are less false alarms and confidence in the system will increase.

Table V shows the cosine similarities of the detected voices against the stored voices of the system user.

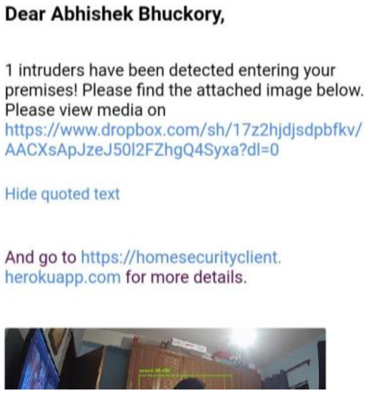
TABLE V: THE ARRANGEMENT OF CHANNELS

Speaker	Voice Recognised	Average cosine similarity with stored voice audios
System User	True	0.2614
Unknown 1	False	−0.056
Unknown 2	False	−1.234
Unknown 3	False	0.151


1 potential intruder(s) have been detected in your premises
Please view footages on https://www.dropbox.com/sh/ncigwfyoim3ww36/AAAW-1OmWlJYXabWHabffPN3a?dl=0
Go to https://homesecurityclient.herokuapp.com for more details

22:11

Fig. 9. SMS.

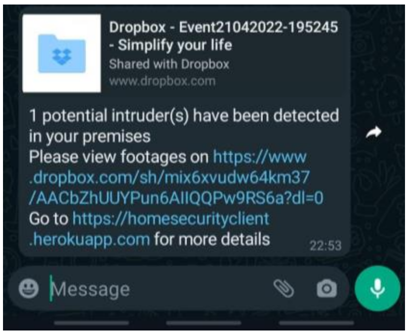

Dear Abhishek Bhuckory,

1 intruders have been detected entering your premises! Please find the attached image below. Please view media on https://www.dropbox.com/sh/17z2hjdjsdpbfkv/AACXsApJzeJ50l2FZhgQ4Syxa?dl=0

Hide quoted text

And go to https://homesecurityclient.herokuapp.com for more details.

Fig. 10. Email.


Dropbox - Event21042022-195245 - Simplify your life
Shared with Dropbox
www.dropbox.com

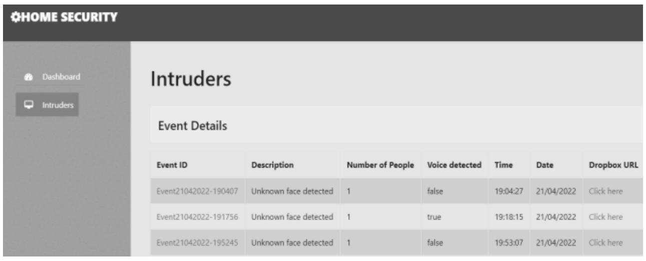
1 potential intruder(s) have been detected in your premises
Please view footages on https://www.dropbox.com/sh/mix6xvudw64km37/AACbZhUUYPun6AIlQQPw9RS6a?dl=0
Go to https://homesecurityclient.herokuapp.com for more details

22:53

Fig. 11. WhatsApp text message.

When an intrusion is detected, the images, video and the recorded audios are uploaded to the Dropbox cloud storage. Event details are then inserted in the MongoDB cloud database. The notification system was also tested while the homeowner was at a remote location. Messages, including SMS, emails and WhatsApp texts were successfully received along with the Dropbox links, as shown in Figs. 9–11, respectively.

A web application was also developed using the Angular and ExpressJS frameworks to represent a dashboard where the homeowner can view all the event details. Fig. 12 shows the web page which displays the information in a tabular form.



Intruders

Event Details

Event ID	Description	Number of People	Voice detected	Time	Date	Dropbox URL
Event21042022-190407	Unknown face detected	1	false	19:04:27	21/04/2022	Click here
Event21042022-191756	Unknown face detected	1	true	19:18:15	21/04/2022	Click here
Event21042022-195245	Unknown face detected	1	false	19:53:07	21/04/2022	Click here

Fig. 12. Dashboard.

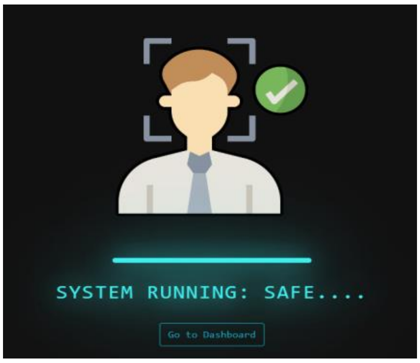


Fig. 13. Safe mode.

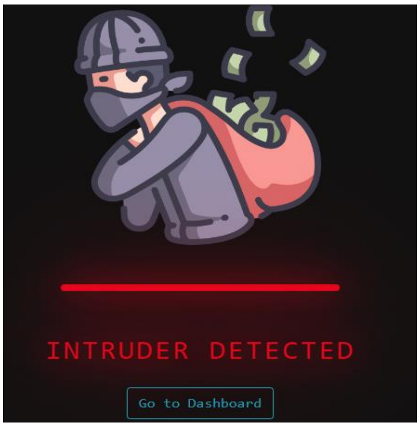


Fig. 14. Alert mode.

Another web page was also created which indicates the live status of the system. It stays on a safe mode when there is no intrusion and turns into an alert mode if the system detects an unknown person, as shown in Figs. 13–14 respectively.

The system can detect intrusions in a multimodal manner, using different hardware. As soon as the relevant sensors are activated, it takes a maximum time of 30 seconds for the alert messages to reach the owner. Despite having different modules running in the system, the program still performs well in the Raspberry Pi device. Moreover, working with the server for processing helps the Raspberry Pi to work efficiently. The system can easily provide surveillance for a larger area. Since IP cameras are used, there are no restrictions on the number of connection ports, unlike wired cameras. Hence, the Raspberry Pi can easily connect to many cameras having the same network address. Moreover, the Dropbox and MongoDB subscriptions may be upgraded to a larger extent for storage of more footage.

While users are connected to the internet, they can remotely access emails and WhatsApp messages, as well as the images stored on the cloud. Yet, even if they have no internet access, SMS are still received. The implemented program is easy to run on the Raspberry Pi and is fully automated. In addition, the web interface provides detailed information and is user friendly. Since it was deployed on a cloud server, it can be accessed anytime and anywhere.

V. Conclusion

Rising crime rates have shown that existing security systems are not effective enough. Intruders can easily break into domestic households during the absence and even during the presence of the homeowners. Thus, in this work, we have developed a home security system using a Raspberry Pi. Several other components have been used to identify potential intrusions. A motion detection module is used to detect movement, while a noise system detects human voice. Surface noises, such as door knocks, are also detected. An object detection model counts the number of people. To prevent false alarms, two face recognition systems were utilized, whereby one uses Euclidean distance to compare new faces with the stored ones, while the other uses an SVM classifier targeted to recognise masked faces. When a user is verified, the door unlocks automatically. When a potential intrusion is detected, notifications are delivered to the homeowners' mobile phones, and the event details are uploaded to a cloud platform. The notifications take an average delay of 30 seconds per event. All the stored details are then displayed on a web application, providing a user-friendly interface to represent the data.

Conflict of Interest

The authors declare that there is no conflict of interest.

Author Contributions

Sameerchand Pudaruth conceptualized the study and contributed to the writing, reviewing, editing and formatting of the paper. Bhavesh Bhuckory conducted the literature review and implemented the software for this system. Bhavesh Bhuckory also wrote the first draft of the paper and contributed to the editing of the paper. All authors have approved the final version.

References

[1] Smart Home Security Global Market Report 2021: COVID-19 growth and change to 2030. (September 30, 2021). The Business Research Company. Report ID: 6151586. [Online]. Available: https://www.reportlinker.com/p06151586/Smart-Home-Security-Global-Market-Report-C

[2] Crime, Justice, and Security Statistics, 2020. (Police, Prosecution, Judiciary, Prisons and Probation). (June 30, 2021). Statistics Mauritius. [Online]. Available: https://statsmauritius.govmu.org/Documents/Statistics/ESI/2021/EI1595/CJS_Yr20_300621.pdf

[3] 10 Surprising Facts About Burglary and Security Systems. (2021). Millennium Fire & Security. [Online]. Available: https://millennium-fire.com/10-surprising-facts-burglary-security-systems/

[4] C. Tak. (August 4, 2019). Surprising statistics about home break ins. Rent Blog. Crime & Safety Families. [Online]. Available: https://www.rent.com/blog/surprising-statistics-about-home-break-ins

[5] Richmond Alarm Company. What Causes False Alarms on Home Security Systems? [Online]. Available: https://richmondalarm.com/security-tips/causes-false-alarms-home-security

[6] P. Bischoff. (2022). Surveillance Camera Statistics: Which City has the Most CCTV Cameras? Comparitech. [Online]. Available: https://www.comparitech.com/vpn-privacy/the-worlds-most-surveilled-cities

[7] BBC News. (December 16, 2020). Charlie Hebdo: Fourteen guilty in 2015 Paris terror attacks trial. BBC News. [Online]. Available: https://www.bbc.com/news/world-europe-55336094

[8] A. Vigderman and G. Turner. (November 16, 2021). Do I Need a Home Security System? Security.org. [Online]. Available: https://www.security.org/home-security-systems/do-i-need-one/

[9] Home Security Statistics in 2021. SafeHome. [Online]. Available: https://www.safehome.org/data/home-security-statistics/

[10] S. Tanwar, P. Patel, K. Patel, S. Tyagi, N. Kumar, and M. S. Obaidat, "An advanced internet of thing-based security alert system for smart home," in *Proc. 2017 International Conference on Computer, Information and Telecommunication Systems (CITS),* IEEE, 2017, pp. 25–29.

[11] K. C. Sahoo and U.C. Pati, "IoT based intrusion detection system using PIR sensor," in *Proc. 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, IEEE, 2017, pp. 1641–1645.

[12] G. Amato, F. Carrara, and F. Falchi, "Facial-based intrusion detection system with deep learning in embedded devices," in *Proc. the 2018 International Conference on Sensors, Signal and Image Processing*, 2018, pp. 64–68.

[13] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multi-box detector," in *Proc. ECCV 2016: 14th European Conference on Computer Vision*, 2016.

[14] J. See and S.W. Lee, "An integrated vision-based architecture for home security system," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, pp. 489–498, 2007.

[15] S. Pudaruth, F. Indiwarsingh, and N. Bhugun, "A unified intrusion alert system using motion detection and face recognition," in *Proc. 2nd International Conference on Machine Learning and Computer Science (IMLCS)*, Malaysia, 2013, pp. 17–20.

[16] F. Khodadin and S. Pudaruth, "An intelligent camera surveillance system with effective notification features," *International Journal of Computing and Digital Systems*, vol. 9, no. 6, pp. 1251–1261, 2020.

[17] J. Paul, R. S. Bhowmick, B. Das, and B. K. Sikdar, "A smart home security system in low computing IoT environment," in *Proc. 2020 IEEE 17th India Council International Conference (INDICON)*, 2020, pp. 1–7.

[18] A. Shahid, M. Ali, M. Eijaz, S. Minhas, and N. Sabahat, "Shield: An intelligent and affordable solution for home security," in *Proc. 2019 11th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, IEEE, 2019, pp. 1–5.

[19] H. Park, J. Park, H. Kim, S. Q. Lee, K. H. Park, and J. Paik, "Hybrid sensor network-based indoor surveillance system for intrusion detection," *Symmetry*, vol. 10, no. 6, p. 181, 2018.

[20] K. H. Park and S. Q. Lee, "Early-stage fire sensing based on audible sound pressure spectra with multi-tone frequencies," *Sensors and Actuators A: Physical*, vol. 247, pp. 418–429, 2016.

[21] E. S. Gonçalves, F. C. Teixeira, D. F. Albuquerque, and E. F. Pedrosa, "Asynchronous mmWave radar interference for indoor intrusion detection," in *Proc. Robot 2019: Fourth Iberian Robotics Conference: Advances in Robotics*, Springer International Publishing, 2020, vol. 2, pp. 367–378.

[22] M. Ravanelli, T. Parcollet, P. Plantinga, A. Rouhe, S. Cornell, L. Lugosch, C. Subakan, N. Dawalatabad, A. Heba, J. Zhong, and J. C. Chou, "SpeechBrain: A general-purpose speech toolkit," arXiv preprint, arXiv:2106.04624, 2021.

[23] ArcGIS. (2020). How single-shot detector (SSD) works? ArcGIS Developer. [Online]. Available: https://developers.arcgis.com/python/guide/how-ssd-works/

[24] A. Fydanaki and Z. Geradts, "Evaluating OpenFace: An open-source automatic facial comparison algorithm for forensics," *Forensic Sciences Research*, vol. 3, no. 3, pp. 202–209, 2018.