Article

Smart Home Anti-Theft System: A Novel Approach for Near Real-Time Monitoring, Smart Home Security and Large Video Data Handling for Wellness Protocol

S. Pandya [1],*, H. Ghayvat[2], K. Kotecha[3], M. H. Yep[4] and P. Gope[5]

[1] Computer Department, P P Savani University, Gujarat, India; sharnil.pandya84@gmail.com
[2] Computing Department, Hong-Kong Polytechnic University, Hong Kong, China; ghayvat@gmail.com
[3] Computer Department, Symbiosis International University, India;drketankotecha@gmail.com
[4] Manchester Metropolitan University, Manchester, UK; M.Yap@mmu.ac.uk
[5] Trust, Centre for Research in Cyber Security of Singapore University of Technology and Design, Singapore, Singapore; gope_prosanta@sutd.edu.sg
* Correspondence: sharnil.pandya84@gmail.com; sharnil.pandya@ppsu.ac.in; Tel.: +91-9429130543

Abstract: The proposed research methodology aims to design a generally implementable framework for providing a house owner/member with the immediate notification of an on-going theft (unauthorized access to their premises). For this purpose, a rigorous analysis of existing systems was undertaken to identify research gaps. The problems found with existing systems were that they can only identify the intruder after the theft, or cannot distinguish between human and non-human objects. Wireless Sensors Networks (WSNs) combined with the use of Internet of Things (IoT), Cognitive Internet of Things, Internet of Medical Things, and Cloud Computing are expanding smart home concepts and solutions, and their applications. The primary objective of the present research work was to design and develop IoT and cloud computing based smart home solutions. In addition, we also propose a novel smart home anti-theft system that can detect an intruder, even if they have partially/fully hidden their face using clothing, leather, fiber, or plastic materials. The proposed system can also detect an intruder in the dark using a CCTV camera without night vision facility. The fundamental idea was to design a cost-effective and efficient system for an individual to be able to detect any kind of theft in real-time and provide instant notification of the theft to the house owner. The system also promises to implement home security with large video data handling in real-time.

Keywords: smart anti-theft system; intruder detection; unsupervised activity monitoring; smart home; partially/fully covered faces

1. Introduction and Related Work

In the modern era, security and surveillance are important issues. Recent acts of theft/terrorism have highlighted the urgent need for efficient video surveillance and on-the-spot notification of ongoing thefts to house owners and other household members. A number of surveillance solutions are currently available on the market, such as CCTV cameras and digital video recorders (DVRs) that can record the unauthorized activities of a trespasser, but cannot distinguish between human and non-human objects [1-2]. In recent times the ratio of theft has increased tremendously due to a lack of awareness and low availability of smart-gadgets. The task of face detection and the recognition of an intruder become very difficult when the intruder hides their face partially or fully using some type of material, such as plastic, leather, or fabric.

Legacy systems cannot provide real-time theft notification to the house owner nor detect partially of fully obscured faces. It is also challenging for old systems to detect the intruder in the dark using a CCTV camera without night vision capability. The major flaw with this kind of arrangement is that it demands the 24/7 availability of a house owner or member, or manual video surveillance, which is almost impossible [3-4]. In addition, it is a tedious task to go through all the recorded video clips after a possible theft is becomes known. It might happen that the storage server contains a large amount of family member footage, which is of no use in identifying trespassers.

The proposed approach can be applied to an IoT-based smart home monitoring system in near real-time. As shown in Figure 1 and 2, a smart home was designed and developed based on an integrated framework of sensors, cameras, and customized hardware to analyze unauthorized access. The system operates at two different levels: through a hardware interface and through a software interface. At the hardware interface level, an intelligent sensing node is deployed, connected to a central sensing node which acknowledges the data and sends it to the storage server. The software modules are further subdivided into several further levels, including data logging, data retrieval, and storage. The main objective of the software is to detect and report unsupervised human activity using large data handling techniques as close to real-time as possible.

The work is presented as follows. The initial section has provided an introduction on legacy systems, potential issues, and their impact on society. Section 2 describes the necessity of the present research. It also describes the intelligent features of the system that can detect obscured faces using graphical and statistical methods. Subsequently, Section 3 presents the design and experimental setup of the system. It also describes two critical features: (i) the detection of obscured faces; and (ii) detection of an intruder in the dark. Section 4 presents the methodologies in three different phases with graphical analysis and statistical results. The final section (Section 5) discusses the research outcomes.

2. Necessity of a New IoT-Based Theft System

Nowadays, intruders have become more technologically aware and have carried out burglaries using smart gadgets like gas-cutter, smart anti-lock systems and many more. For such intruders, it is straightforward to disconnect a CCTV camera surveillance, which has an indirect connection to the digital video recorder and a database server residing at home. Therefore, there is a need to modify existing systems [5-11][13-18] and propose an intelligent approach that can not only provide unsupervised human activity monitoring, but can also stop an on-going theft by notifying the house-owner at the earliest opportunity. All legacy systems work on the premise of object detection, object motion detection and tracking. Such systems are prone to false alerts or notifications, which might result in sending false emergency notifications to the house owner/member, the escape of the intruder after the theft, and unnecessary disruptions to the residents. To resolve these issues, a novel human activity monitoring detection, recognition, and home security approach is presented in the remaining sections. The list of used terminologies is described in Table 1.

The overall arrangement of the sensing units is as follows. The smart home monitoring and control framework is applied on two unique levels: equipment and programming. The equipment framework contains the sensor arrangements. This equipment framework is further classified into four areas: body sensor setup (BSS), ambient sensor setup (ASS), crisis sensor setup (CSS), and other sensors setup (OSS). The BSS is enabled with an impact sensor. A remote BSS provides observations of inhabitants in different physiological states. The BSS framework incorporates physiological checking gadgets that are capable of recording the daily activities of the smart home residents without disturbing their daily routine.

The second equipment system, ASS, contains a temperature sensing unit, motion sensing unit, and pressure sensing unit. The CSS is equipped with numerous manual push buttons, for example there is a panic push button for emergency situations such as a house fire which activates security and alarm systems. The final setup, OSS, offers the utilization checking and control of electrical home devices through the electrical and electronic sensing unit.

OSS additionally incorporates the contact sensing unit. The OSS framework is in charge of information accumulation, mining, and storage into the database server. Finally, server information is collected and handled by machine learning and information mining models to deliver helpful data for website and output action. The arrangement of the wireless sensing units is reported in the following section.

The smart home monitoring and control framework has been persistently run since May 2015 in a vintage house, built in 1945. Figure 1 depicts the house where the smart home monitoring and control framework is operating and the residents are living out their daily lives [12]. A schematic of the home the installed sensors are shown in Figure 2. It indicates distinctive remote sensing units for the observation of natural parameters and diverse household objects. The smart home monitoring and control framework contains a heterogeneous sensor setup, which monitors temperature, force, contact, smoke, motion, and use of electronic and electrical (E and E) appliances. Figure 3 shows the outside temperature detecting unit, Figure 4 shows the use of the E and E unit with household machines, the force detecting unit is shown in Figure 5, and Figure 6 presents the motion detecting unit. Furthermore, Figure 7a demonstrates the contact sensing unit associated with the refrigerator and Figure 7b pictures the manual push buttons.

Figure 1. Image of a smart building where a smart home monitoring and anti-theft system has been installed.

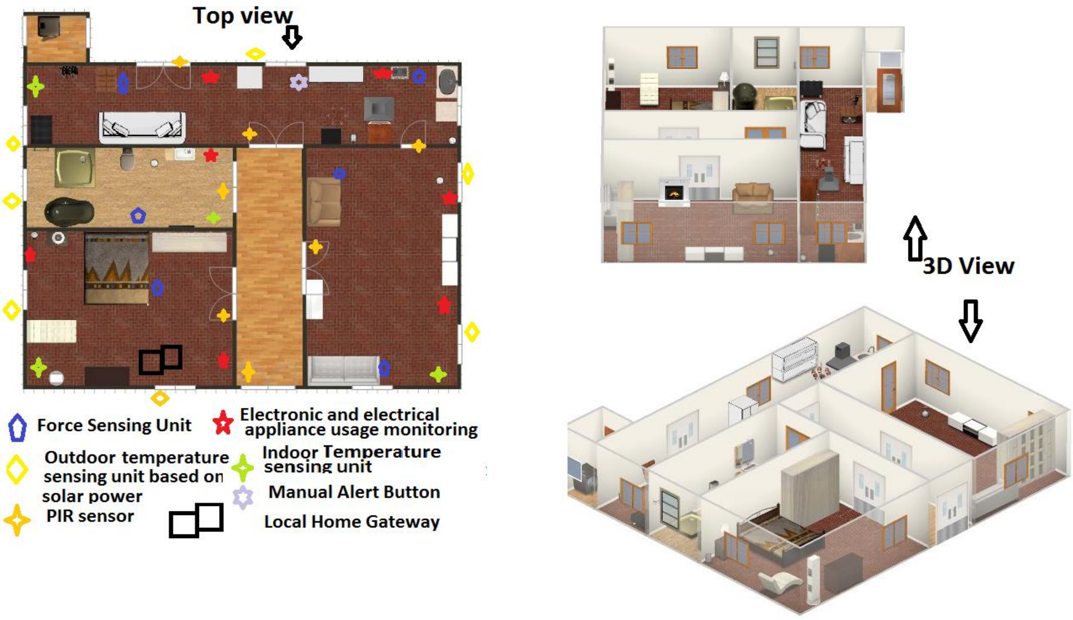

Top view
3D View
Force Sensing Unit
Electronic and electrical appliance usage monitoring
Outdoor temperature sensing unit based on solar power
Indoor Temperature sensing unit
Manual Alert Button
PIR sensor
Local Home Gateway

Figure 2. 3D Black box view of the house with Top View.

Figure 3. Outdoor temperature measurement sensing unit.

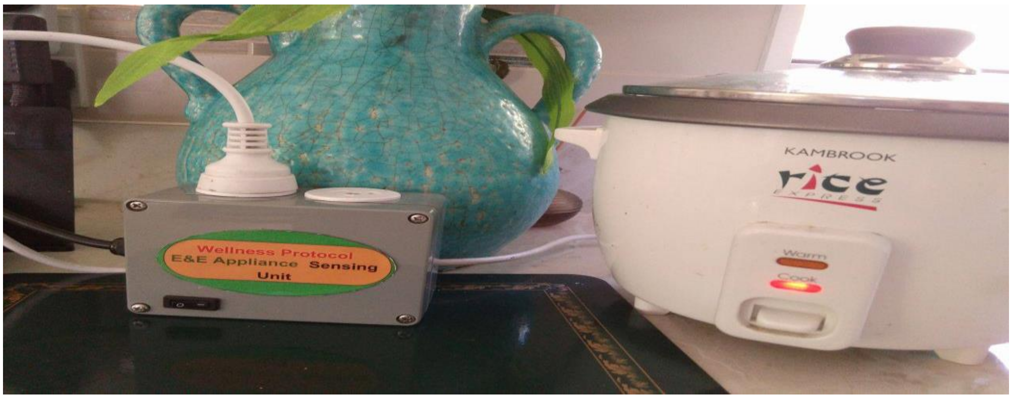


Figure 4. Electronic and electrical (E & E) sensing unit monitoring and controlling rice cooker usage.

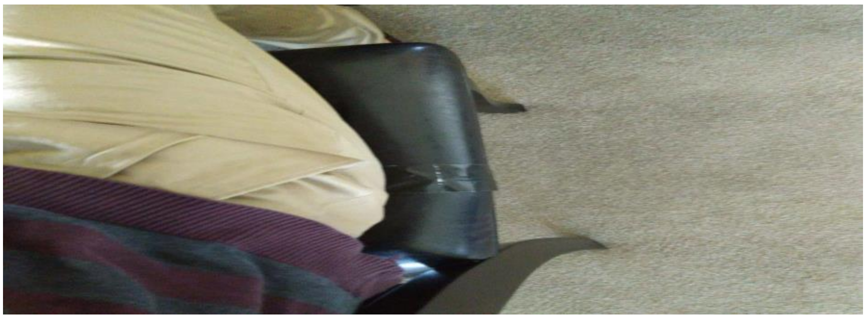


Figure 5. Force sensing unit deployment to monitor chair usage.

Figure 6. PIR sensing unit deployment to monitor entry and exit of bedroom.

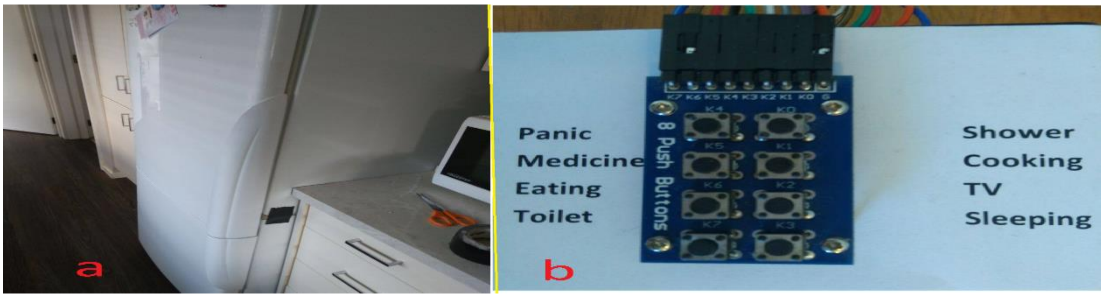


Figure 7. (**a**) Contact sensing unit connected to fridge door and (**b**) Manual push button indicator.

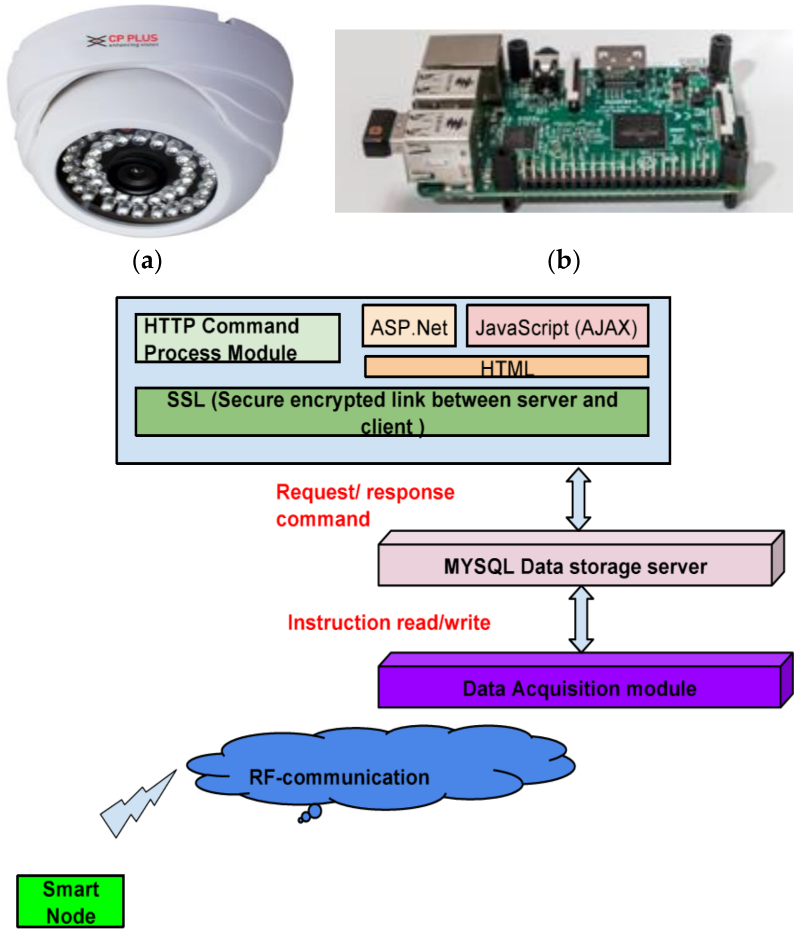
Table 1. List of used terminologies.

Notations	Acronyms
A	Present novel system for surveillance, detection and instant notification of intrusion for preventing theft
I	Image Capturing Module
M	Identification Module
Me	Eye and Face Detection
Mp	Pixel Processing
Mms	Motion Detection
Mc	Comparison
S	Storage Module
C	Controller Module
Cp	Processing Module
T	Transmission Module
D	Display Module

3. Design and Experimental Setup

As shown in Figure 8, a CP-PLUS analogue camera model 850 tvl, a Wi-Fi enabled customized hardware with 1 GB RAM and 256 GB memory card, and a Samsung Grand mobile device were used. The camera was placed in each of the smart home to protect the house from the intruders equipped with smart gadgets. Nowadays, intruders often understand that the camera would be connected to a laptop/computer/tablet and can disconnect it to stop the system functioning. The customized hardware used in the system has Wi-Fi capability, which allows placing the proposed system anywhere inside the smart home where Wi-Fi is available. Even in the case of power failure, the proposed system can still function if it is connected to a laptop/tablet with a hotspot internet connection. In addition, the customized hardware is covered by a plastic coating to protect it from the water inside the electricity conceal lines.

As shown in Figure 9, the proposed system starts functioning when an intruder comes into the monitored area. The movement of an intruder is captured by a face detection module, as shown in Figure 10. The main function of the face detection module is to differentiate human objects from non-human objects. When any human action is detected, the camera is activated at a rate of 15 frames per second. After the initial frame has been captured, it is immediately sent to the house owner on a specialized mobile app. The video processing module continues to capture the rest of the frames at 45 frames per seconds for up to 40 s and stores them in mp4 video format. The captured 40 s mp4 video is then compressed to approximately 10 s of fast-forward format video. The reason behind converting and compressing the video from 15 frames per second to 45 frames per second is to enable the video to be received by a mobile phone with slow internet connection. The compressed video is also sent to the homeowner. This step allows the house owners to make a rapid decision about whether to inform a neighbor or the police.


CP PLUS
HTTP Command Process Module
ASP.Net
JavaScript (AJAX)
HTML
SSL (Secure encrypted link between server and client)
Request/ response command
MYSQL Data storage server
Instruction read/write
Data Acquisition module
RF-communication
Smart Node

(**a**) (**b**)

(**c**)

Figure 8. Detailed workflow of the system. (**a**) CP Plus 850 tvl analogue camera; (**b**) Customized hardware in which the proposed research methodology is coded; (**c**) The layered architecture of an anti-theft system for an IoT-based smart home solution.

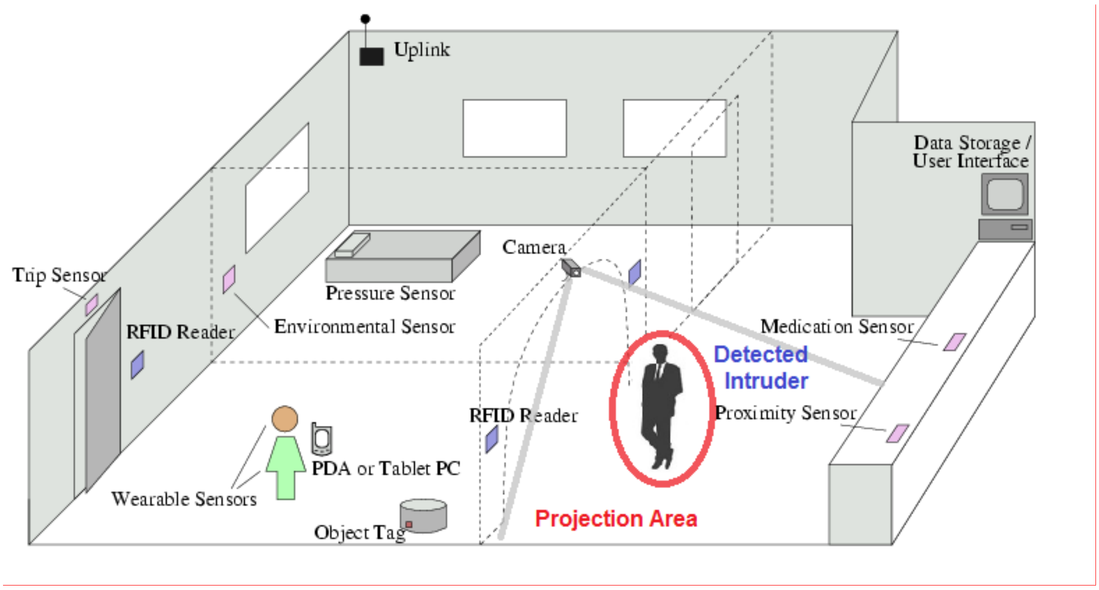


Figure 9. Projection area of the smart home anti-theft system.

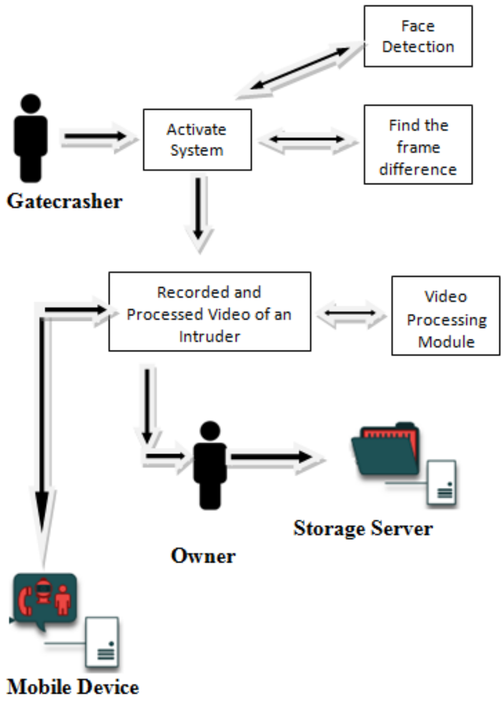


Figure 10. Detailed workflow of the system.

The detailed technical properties of the core modules of the proposed system are explained in the following:

I. Functioning Module/Processing Module: Captures the presence of an intruder in the surveillance area as shown in Figure 9 at the rate of 15 frames per second. After detecting an intruder, the functioning module sends the captured frames to the identification module (M) for detection and intruder identification.

II. Identification Module: When an intruder enters the surveillance area, the identification module (M) detects the presence and identifies whether it is human or non-human by assessing various regions of the face of the captured intruder using the eye and face detection module (Me). The identification module (M) has four sub-modules, as shown in Figure 11:

a. Eye and Face Detection Module: Captures the face of the intruder by distinguishing between human and non-human objects. If an intruder has partially covered their face, the module detects the brightest part of the intruder's face or the region of interest using the pixel processing module (Mp). Following this process, the eye and face detection module sends the captured images to the pixel processing module for further processing.
b. Pixel Processing Module: First, if the intruder has partially covered their face, this module detects the brightest region of the face such as the eyes, cheeks, or upper part of the head [19-20]. Next, it detects the motion of the captured intruder using the motion detection module (Mm). This eliminates the risks of false alarms from photographs of humans printed on the walls, magazines or newspapers.
c. Motion Detection Module: If a detected intruder is a human, the motion detection module detects the motion of an intruder; it captures the first frame of an intruder from video at a rate of 15 frames per second and captures the following frames at 45 frames per second. The motion of the captured intruder is based on the differences between the captured frames. The video is then sent to the comparison module (Mc) to mitigate the possibility of false alarms.
d. Comparison Module: To mitigate the possibility of false alarms, the comparison module (Mc) distinguishes between human and non-human objects using the modified Haar Cascade algorithm and sends near real-time notifications to the house owner/member.

III. Storage Module: This module stores the images of all the captured intruders. The detection and recognition of any unauthorized access in the smart home surveillance area is captured by the smart home anti-theft system in three phases: (a) primary face detection phase, (b) secondary face detection phase, and (c) a final phase where an intruder has partially or fully hidden their face or has been detected in the dark.

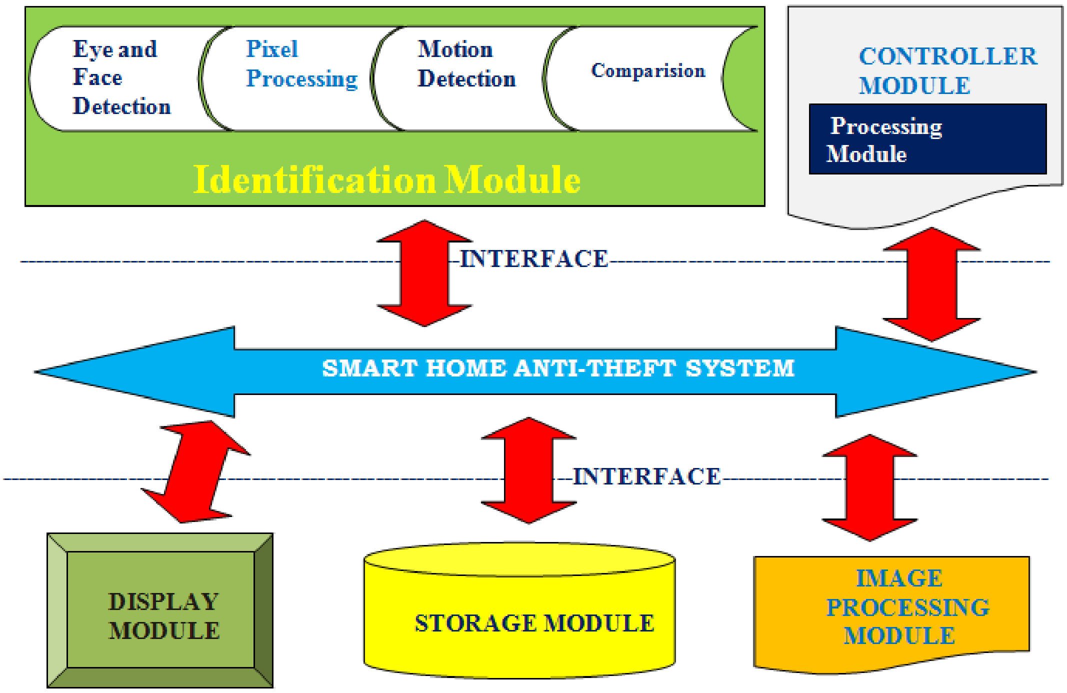


Figure 11. Communication block diagram of the smart home anti-theft system.

4. Methodologies and Results Analysis

The research methodology was classified into three stages with their corresponding results; these stages are as follows: (a) primary phase, (b) secondary phase, and (c) final phase.

Primary Phase:

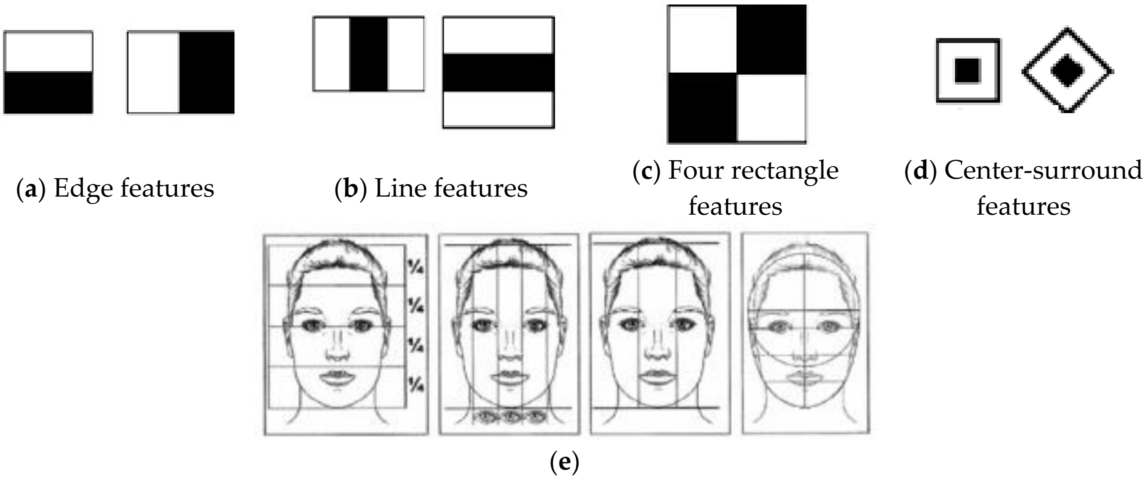
The eye and face detection module was programmed with the modified feature-based Haar cascade face detection classifier algorithm [15, 19, 21, 22]. The modified Haar cascade face detection classifier algorithm detects an edge, line, and center-surround features of an intruder object. The variety of Haar cascade classifiers is shown in Figure 12. As shown in Figures 12 and 13a, the rectangle features can be captured and computed based on the intermediate representation of the captured intruder's image. The results of the comparison analysis is given in Figure 13b, which shows that the success rate of the Haar cascade classifier was very high when compared to previous methodologies. This is the primary reason why the modified Haar cascade classifier was used in the present research experiments. However, to compete with the newly introduced methodologies, the proposed algorithm has been modified accordingly. The detected image of an intruder at a location a, b contains the sum of the pixels above and to the left of a, b, as described:

$$ii(a,b) = \sum_{a' \le a, b \le b'} i(a',b') \tag{1}$$

where *ii*(*a*, *b*) is the detected image of an intruder and *i*(*a*, *b*) is the original image. Using the following pair of recurrences:

$$s(a,b) = s(a,b-1) + i(a,b) \tag{2}$$

$$ii(a,b) = ii(\mathrm{a},b-1) + s(a,b) \tag{3}$$



(**a**) Edge features

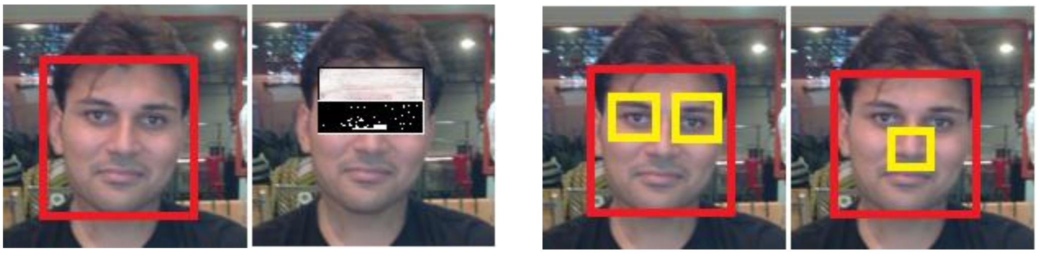
(**b**) Line features

(**c**) Four rectangle features

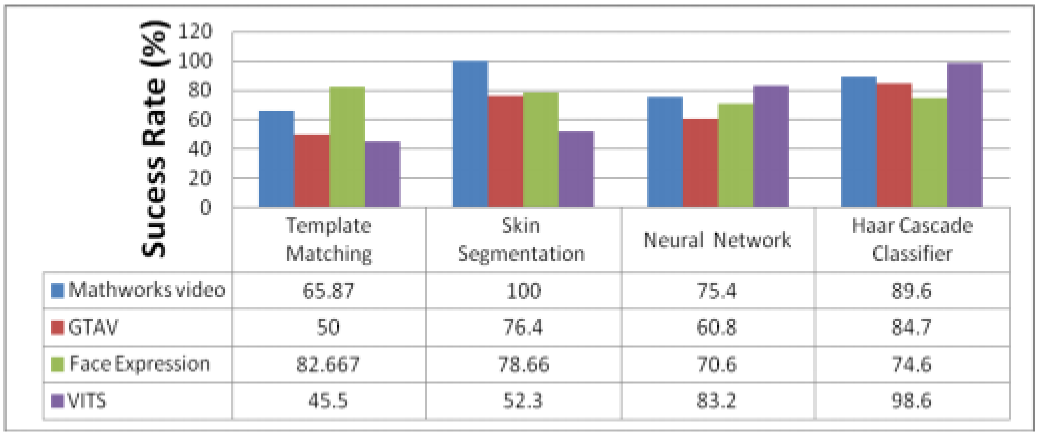
(**d**) Center-surround features

(**e**)

Figure 12. Haar cascade classifiers, (**a**) edges; (**b**) lines; (**c**) four rectangles; (**d**) center-surround features, and (**e**) face classifiers.



(**a**)



	Template Matching	Skin Segmentation	Neural Network	Haar Cascade Classifier
🟦 Mathworks video	65.87	100	75.4	89.6
🟥 GTAV	50	76.4	60.8	84.7
🟩 Face Expression	82.667	78.66	70.6	74.6
🟪 VITS	45.5	52.3	83.2	98.6

(**b**)

Figure 13. (**a**) Haar cascade image classifier samples (all of the shown human faces in the conducted experiments required copyright permission obtained on 19/03/2017); (**b**) Comparison analysis of the Haar cascade classifier with existing methodologies [21-22].

A cascade classifier calculates the centroid of the detected intruder's face, and a radius is calculated based on a scale value. According to the geometrical face model [23], with the cropped mouth region as the region of interest (ROI), we defined the top left coordinates as (*x*, *y*), the centroid point as (*cx*, *cy*), and the radius as *r*. The mouth region could then be roughly obtained as follows:

$$x = \left(cx - \left(\frac{2r}{5}\right)\right) \mathit{and}\ y = cy + \frac{r}{3} \qquad (2)$$

Wm = *r* and *Hm* = *r*/2 where mouth height is *Hm* and mouth width is *Wm*.

As shown in Figure 14, a positive result from the first classifier triggers the assessment of a second classifier which in turn triggers a third classifier, and so on. If during the whole process, a negative result is prompted, then it results in the termination of the sub-window. In this phase, numerous classifier stages are built using the modified Haar cascade to minimize false negatives.

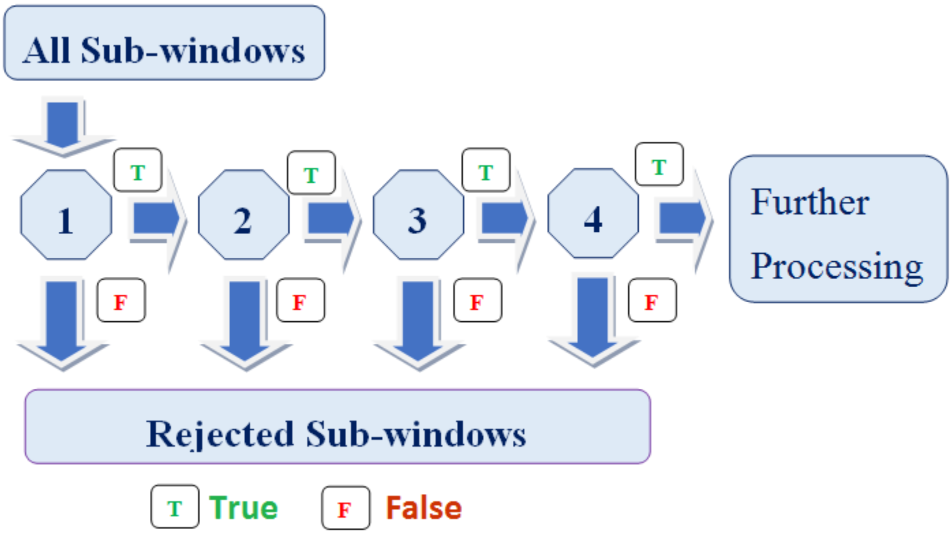


Figure 14. Schematic representation of cascade classifiers.

In comparison to the existing methodologies, this framework was 15 times faster than the other recognition frameworks [21]. As per the face location framework portrayed by Rowley et al., two location systems face location framework is the fastest algorithm among existing techniques. The proposed system could also capture a group of faces captured in a single frame using modified Haar cascade crowd classifiers, as presented in Figure 15.

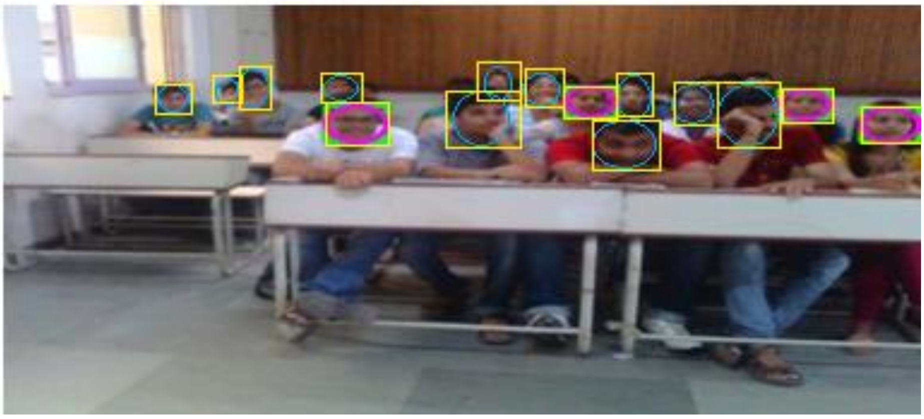


Figure 15. Crowd Cascade Image Classifier Sample (copyright permission for photograph obtained on 01/04/2017).

As shown in Figure 16, rigorous research experiments were conducted with 10,000 diverse face samples in scenarios where an intruder was detected in ordinary scenarios and where an intruder had not fully/partially covered his/her face. Figure 16 shows face detection rates for approximately 10,000 human samples. Accurate face detection was achieved in more than 97% of cases, as shown in Figure 16.

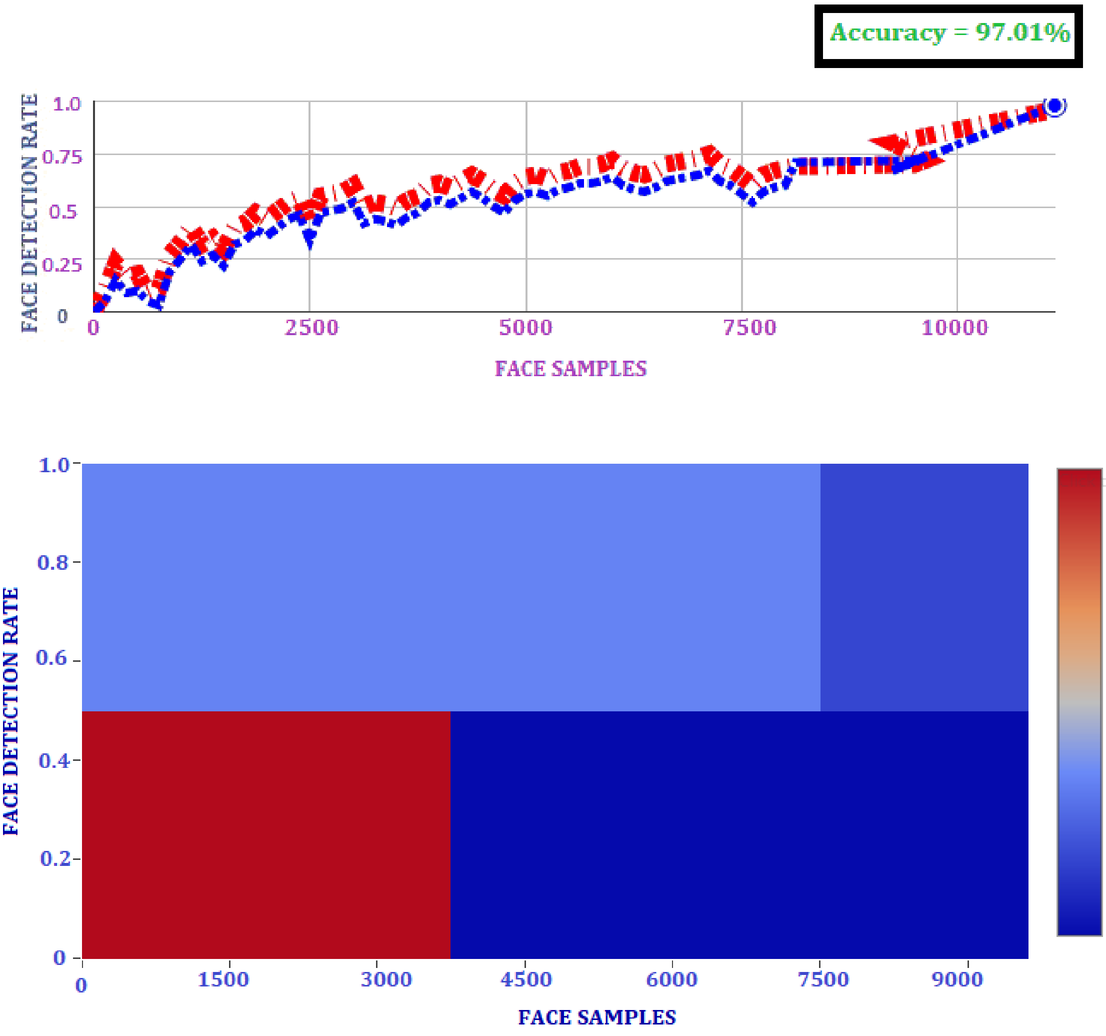


Figure 16. Face detection rate analysis of the proposed system with 10,000 face samples in an ordinary scenario where an intruder has not occluded their face, as in Figure 15.

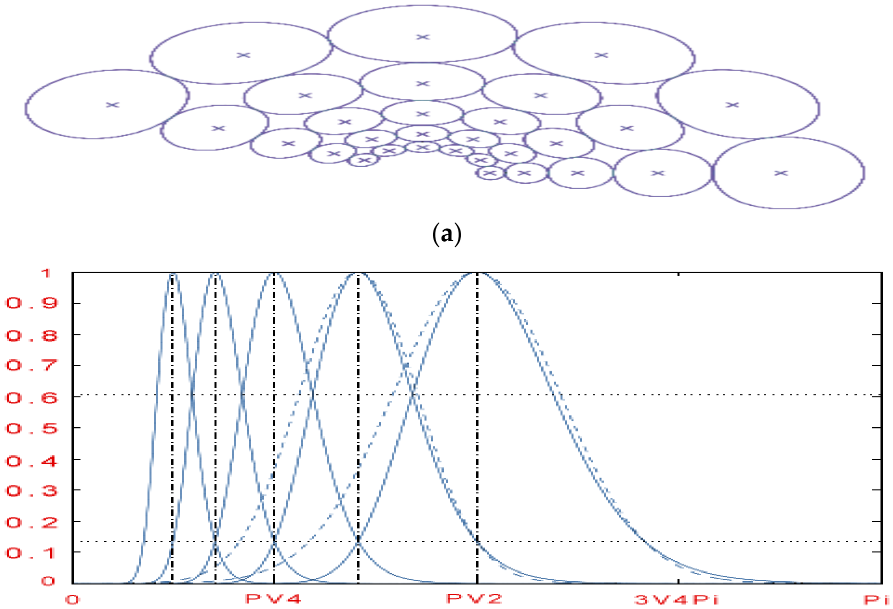
Secondary Phase:

After the initial face detection is complete, an intelligent secondary stage algorithm is designed to eliminate false positives which capture half-hidden human faces, based on the brightest features of the face [24-29]. A dynamic threshold was set to capture the brightest part of the face, even where part of the face was hidden. For detecting the retinal point and other bright facial features, a modified Gabor filter was used [30-32].

The nearby power range of the picture was inspected at every retinal point by the vector of Gabor channels that constituted the responsive field connected with that point. Gabor channels are ideal as they minimize the joint picture and frequency plane spread. At the point when only a small number of frequency channels are utilized, the Gaussian range of the channels results in over-dense coverage near the starting point of the frequency plane, while the high-frequency regions are poorly covered. This is because each Gaussian weights high and also low frequencies in a symmetric way, while the disintegration is coarser at high frequencies. To counter this effect, a set of changed Gabor channels were employed, which are characterized as Gaussians in the log-polar frequency plane [19]. Figure 17 shows the modified Gabor filter in the Fourier domain and a 2D plot of cross-sectional coordinates. Thus, a channel tuned to angular frequency can be defined as follows:

$$\omega o = \exp(\delta o)$$
$$G(\omega, \delta) = \frac{A \exp\{(\delta - \delta 0)^2\}}{2\sigma_\delta^2} \times \exp\left\{\frac{-(\omega - \omega o)^2}{2\sigma_\omega^2}\right\} \quad (3)$$

where A is a normalization constant and (δ, ω) are polar frequency coordinates.



(a)

(b)

Figure 17. (**a**) Gabor filters in the Fourier domain, level curves (**b**) cross-sectional plot.

Figure 18 depicts various classifier samples where an intruder has partially covered his/her face with plain, black, shaded, framed, and frameless glasses. It also shows scenarios where an intruder has partially covered their face with transparent, solid, plastic and/or leather materials. In addition, Figure 19 also presents a few classifier samples that were captured on a dark night with an ordinary analog camera.

Rigorous experiments were conducted with approximately 10,000 face samples in scenarios where a detected intruder had partially covered his/her face with some type of transparent, solid, plastic, or leather material. Figure 19 reports the correct face detection rate analysis for 10,000 human under these conditions. The correct face detection rate accuracy achieved was more than 84%, with similar rates achieved if more than 10,000 images were used.

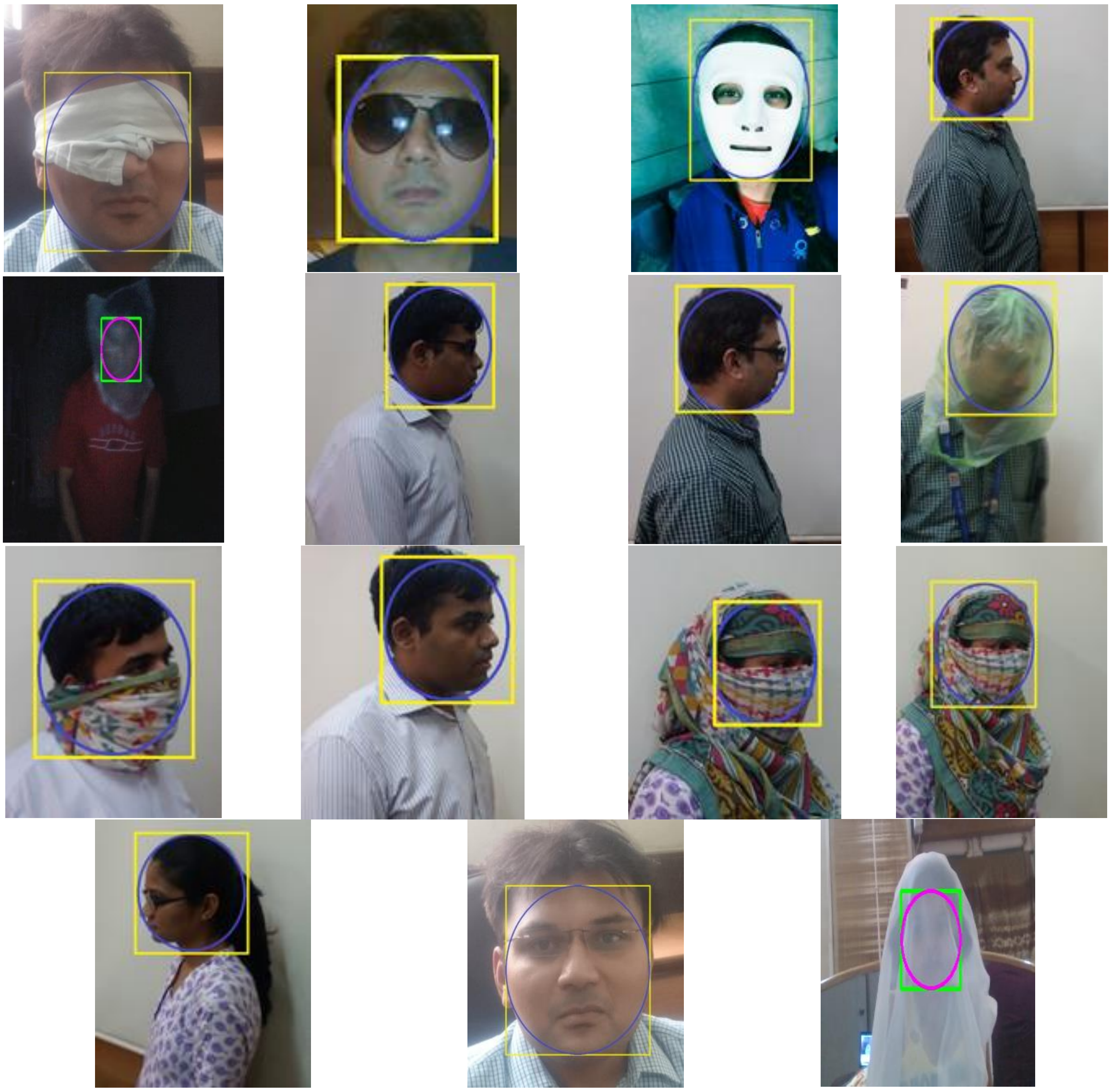
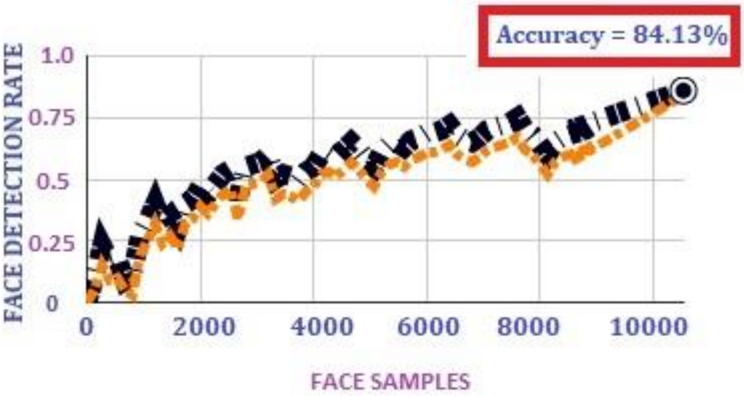


Figure 18. Intermediate stage classifier samples where an intruder has worn plain/black glasses, partially covered face with some type of transparent, solid, plastic, leather material, including the following: tinted glasses, plain-rimmed glasses, face covered with a white cloth, facial overlap with plastic, face covered in a transparent cloth (copyright permission for images obtained on 21/05/2017).



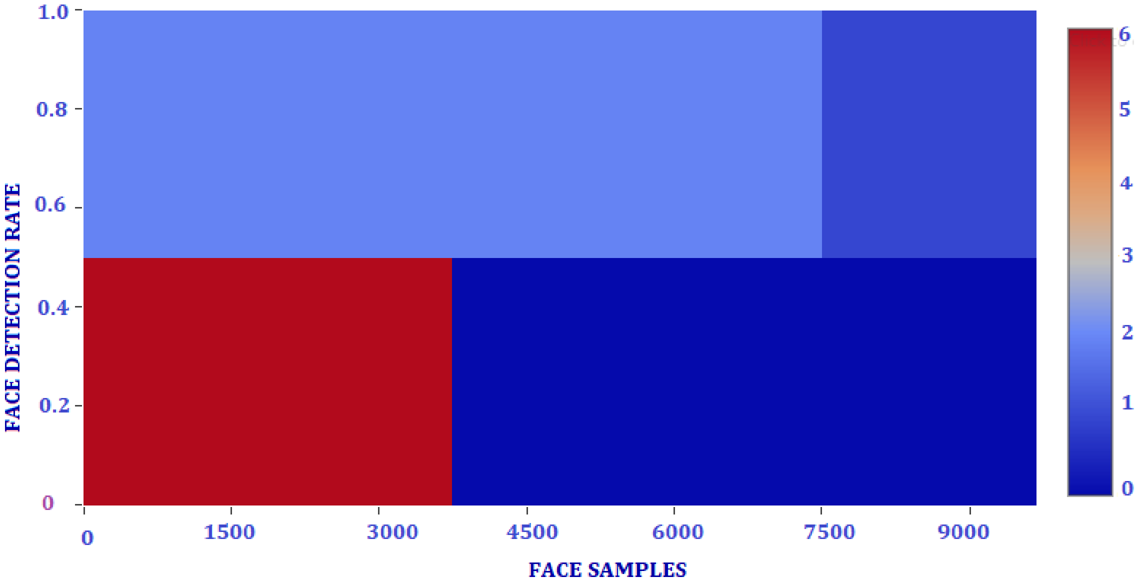


Figure 19. Face detection rate analysis of the proposed system with 10,000 face samples in scenarios where an intruder has partially covered their face with some type of transparent, solid, plastic, or leather materials.

Final Phase:

After the completion of the secondary phase, if the system has not been able to capture the facial features of the trespasser, it will execute the final phase. In this phase, if a detected intruder is a human, it will track the motion of the detected human object and notify the house owner about the presence of a stranger in the house by sending the latest captured image and a compressed video (at 45 fps). Furthermore, the system is intelligent enough to capture the intruder based on the variations of the face; a dynamic threshold has been set to capture the brightest part of the face even if an intruder has fully covered his/her face with some type of transparent, solid, plastic, and/or leather material. The system can also detect intruders at night or in poor light conditions. This requires an ordinary analog camera, which makes the system very economical. To detect the intruder's motion, reference images (a, b) and the captured video frames were obtained from the captured video sequence. A binary motion detection mask M (a, b) can be calculated as follows:

$$M(a, b) = \{1, \text{ if } |It\ (a, b) - B\ (a, b)| > T \quad \{0, \text{ if } |It\ (a, b) - B\ (a, b)| \leq T \tag{4}$$

Here, T is the pre-defined threshold which captures the moving objects in a captured video frame. If the outright contrast between a reference picture and an approaching video outline does not surpass T, the pixels of the identification cover are named "0," which implies it contains background. Generally, dynamic pixels are named "1," which assigns it as containing moving items. To capture complex scenes, a simple statistical difference method was applied, using the mean value and the standard deviation. This method computes the mean value for each and every pixel of the captured video frames. The pixel value of each pixel μxy has been calculated from a collection of all the previously captured frames in the given time interval $(t_0, t_k - 1)$. For every computed pixel, the threshold is give by a standard deviation σ_{ab} as follows:

$$\mu_{ab} = 1/K \sum_{k=0}^{k-1} I_k\ (a, b)$$

$$\sigma_{ab} = (1/K \sum_{k=0}^{k-1} (I_k(a, b) - \mu_{ab})2)1/2 \tag{5}$$

In order to achieve object motion detection, the difference between background of the image and captured video frame was calculated. Here, the predefined parameter is λ. If the exact difference between the captured video frame and image background is greater than $\lambda\sigma_{xy}$ or less than $\lambda\sigma_{ab}$, then the background pixel can be calculated as follows:

$$I(a, b) = \begin{cases} 1, \text{ if } |It(a, b) - B(a, b)| > \lambda\sigma_{xy} \\ 0, \text{ if } |It(a, b) - B(a, b)| \leq \lambda\sigma_{xy} \end{cases} \quad (6)$$

Figure 20 depicts various classifier images where an intruder has completely covered his/her face. It also shows classifier samples which have been captured at night with an ordinary analog camera. It is clear from the samples in Figure 20 that it is not necessary for an intruder to look directly at the camera in order to be captured. Once an intruder is in the projected area of the system, it will capture the intruder even if he/she is not directly looking at the camera.

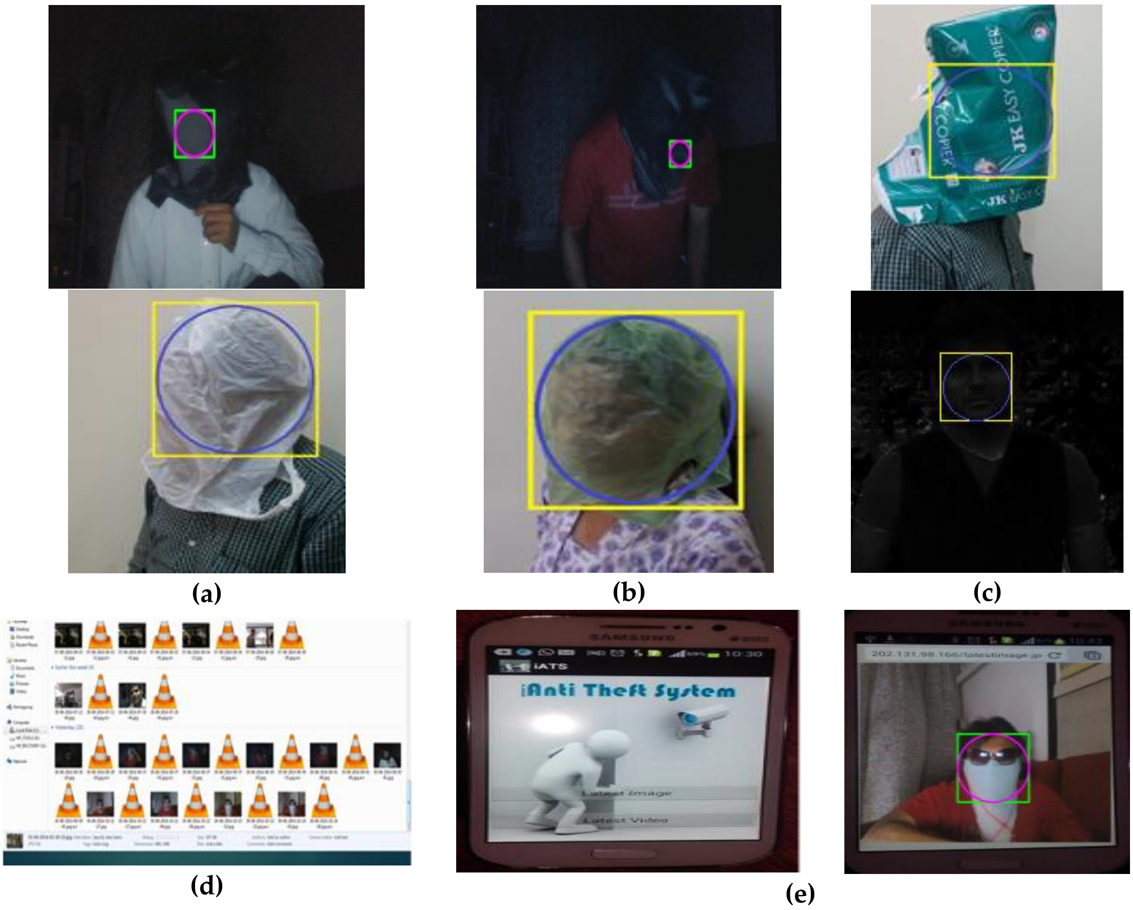


Figure 20. Final Stage Classifier Samples where an intruder has a fully covered face with some type of transparent, solid, plastic, leather materials and detection in the dark night. Face hidden with (**a**) black plastic, (**b**) black plastic and in motion; **(c)** face captured on a dark night ,**(d)** storage samples of captured intruder's image and a compressed video with a timestamp, **(e)** a customized mobile app with samples of on-going theft notification (copyright permission was obtained for all of the human faces shown on 11/05/2017).

As shown in Figure 21, rigorous research experiments were conducted with approximately 10,000 face samples in scenarios where a detected intruder had fully covered his/her face with some type of transparent, solid, paper, plastic, or leather material. The above figure depicts the face detection rate analysis for 10,000 human samples with these variations. Faces were detected accurately approximately in 78% of cases as shown in Figure 21.

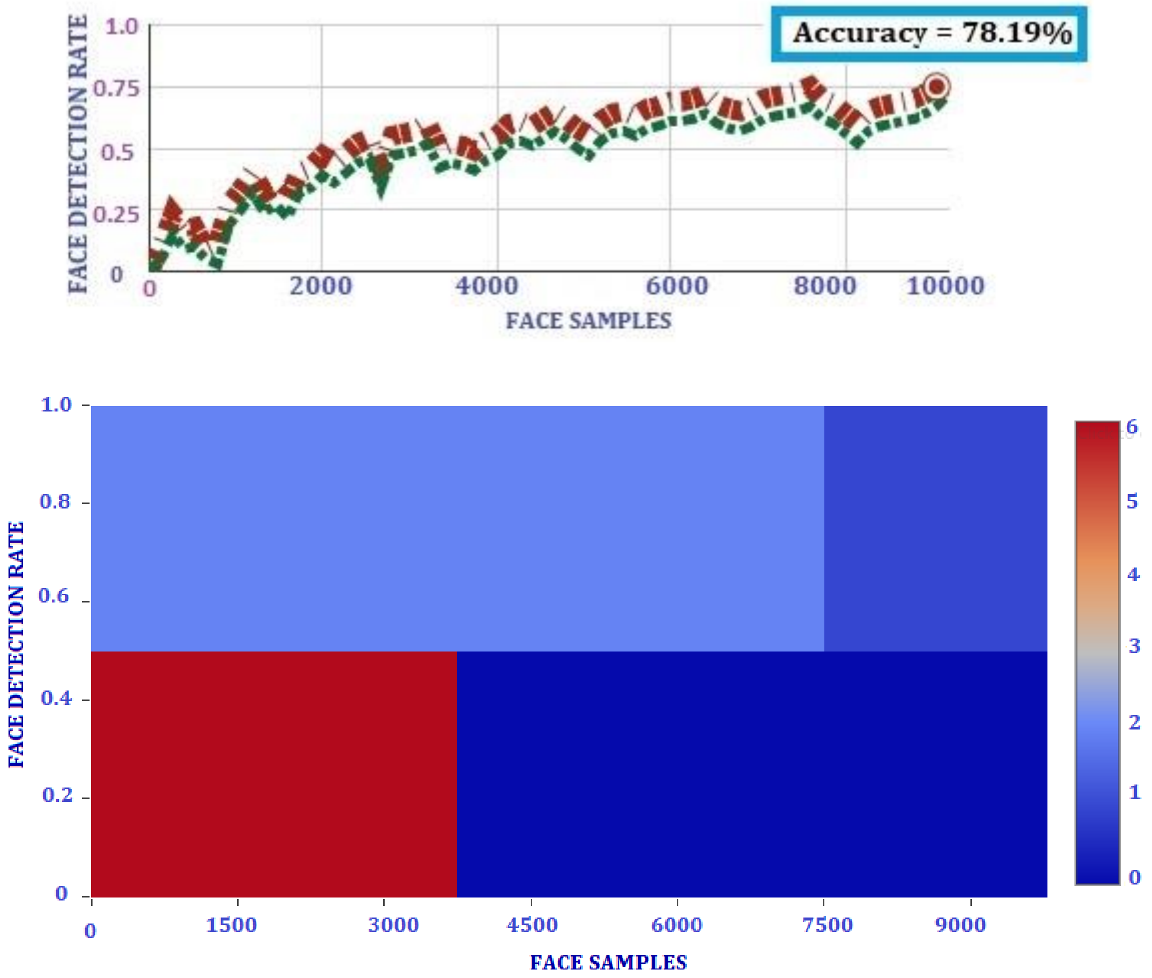
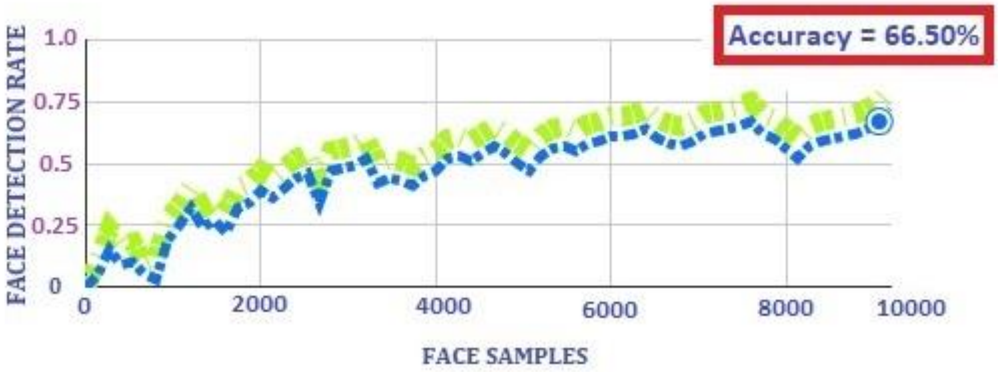


Figure 21. Face detection rate analysis of the proposed system with 10,000 face samples in scenarios where an intruder has a fully covered face, as per Figure 9.

Figure 22 depicts the correct face detection rate analysis against 10,000 images of human detected at night by an ordinary analogue camera without night vision capability. Faces were accurately detected in approximately 67% of cases, with similar accuracy achieved for high numbers of images.



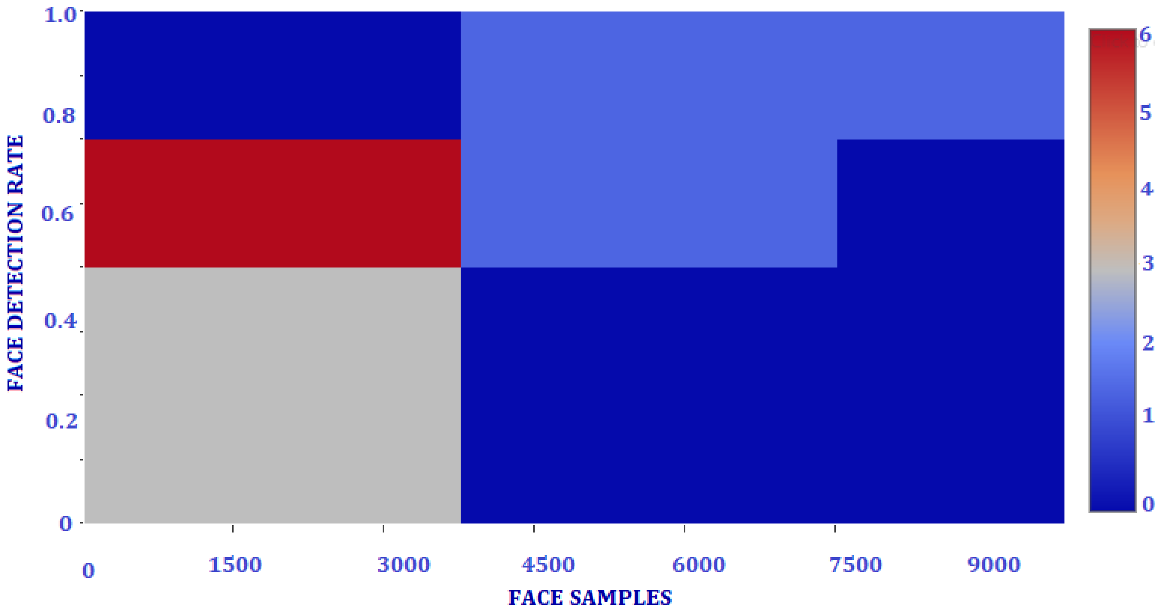


Figure 22. Face detection rate analysis of the proposed system with 10,000 face samples in scenarios where an intruder has been detected in the dark night.

Analysis of the male and female ratio used in the conducted rigorous research experiments in all the above-mentioned scenarios is presented in Figure 23, which depicts a variety of male and female samples in the following scenarios: (**a**) face not obscured, (**b**) face partially covered, (**c**) face fully covered, and (**d**) captured in the dark by an analogue camera. The total sample size used for all the research experiments was 10,000 images.

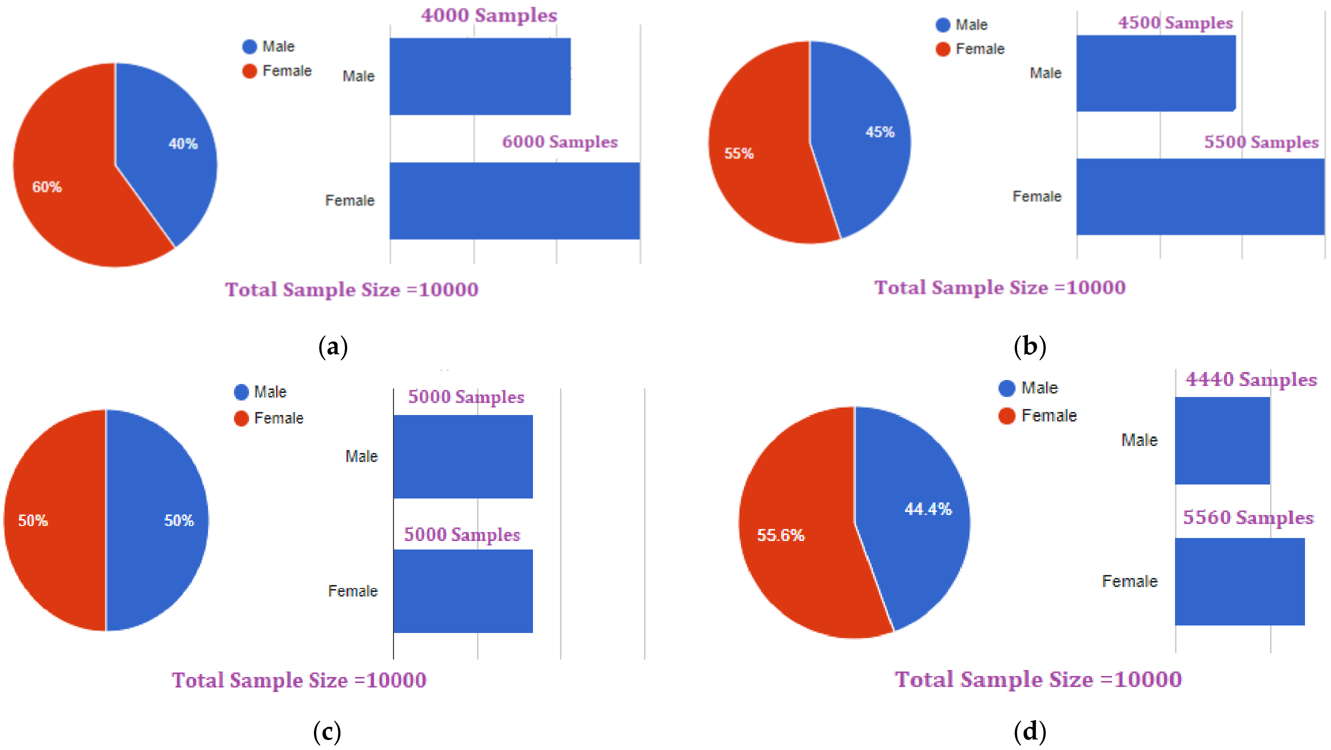


Figure 23. Analysis of male–female ratio in where the face is (**a**) face not obscured, (**b**) face partially covered, (**c**) face fully covered, and (**d**) captured in the dark by an analogue camera.

An analysis of various skin tones: dark, whitish brown, whitish, fair, very fair, and white is given in Figure 24. The proposed system achieved high accuracy for all varieties of skin tones and for all the above-mentioned scenarios, as shown in Table 2. It clearly depicts that the conducted experiments on a variety of skin tones such as white, very fair, fair, whitish brown, whitish and dark achieved above 80% accuracy when an intruder has partially hidden their face, above 70% when an intruder has fully hidden the face with some type of plastic, leather, or paper, and above 60% accuracy when an intruder was captured in the dark night with an ordinary analogue camera without night vision capability.

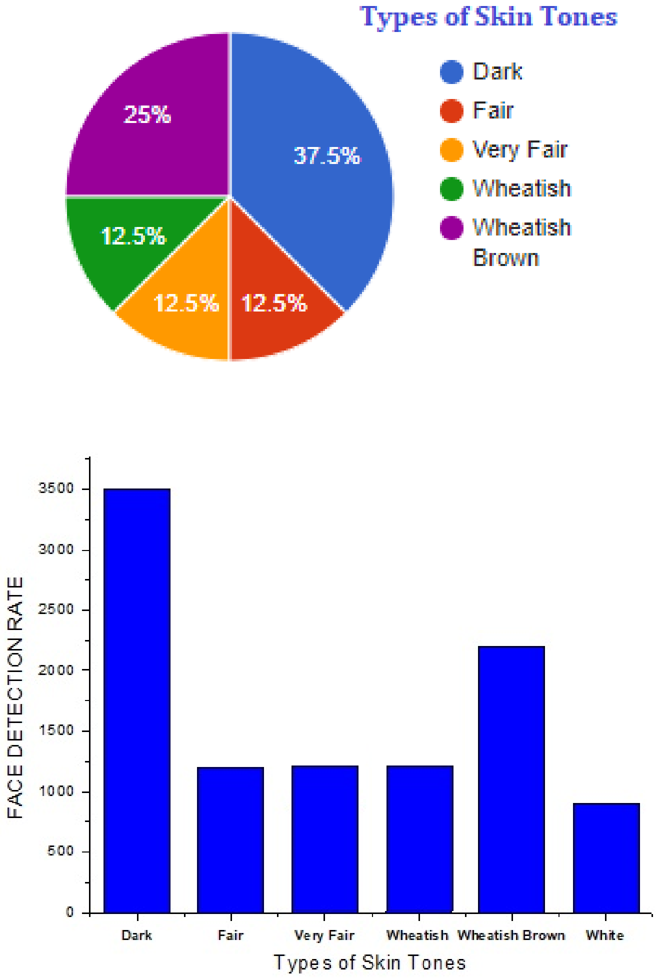


Figure 24. Analysis of various skin tones samples used in scenarios: (**a**) face not obscured, (**b**) face partially covered, (**c**) face fully covered, and (**d**) captured in the dark by an analogue camera.

Table 2. Accuracy analysis of variety of skin tones in different scenarios.

Skin Tones	**Accuracy % (Total Sample Size = 10,000)**			
	Scenario 1: Instruder with face not obscured	**Scenario 2: Intruder has partially covered face with some type of transparent, solid, plastic, leather materials (%)**	**Scenario 3: Intruder has fully covered face with some type of transparent, solid, plastic, leather materials (%)**	**Scenario 4: intruder is captured in the dark night or in bad light conditions (%)**
White	97.9	91.9	79.7	66.4
Very Fair	97.8	87.8	76.7	64.01
Fair	95.6	85.6	74.9	63.12
Wheatish	94.5	84.5	73.8	63.05
Wheatish Brown	94.3	84.3	72.6	62.75
Dark	93.7	83.7	71.9	62.12

As shown in Figure 25, a comparison was carried out between all the scenarios mentioned in Figures 16, 19, 21, and 22 for the correct face detection rate of 10,000 face samples. The comparison stated that the achieved correct face detection accuracy in scenarios where a detected intruder had not hidden his/her face, hidden his/her face partially, fully, and was detected in the dark were 97.01%, 84.13, 78.19%, and 66.5%, respectively. Figure 26 describes the age group analysis of 10,000 face samples captured in four scenarios: : (**a**) face not obscured, (**b**) face partially covered, (**c**) face fully covered, and (**d**) captured in the dark by an analogue camera.

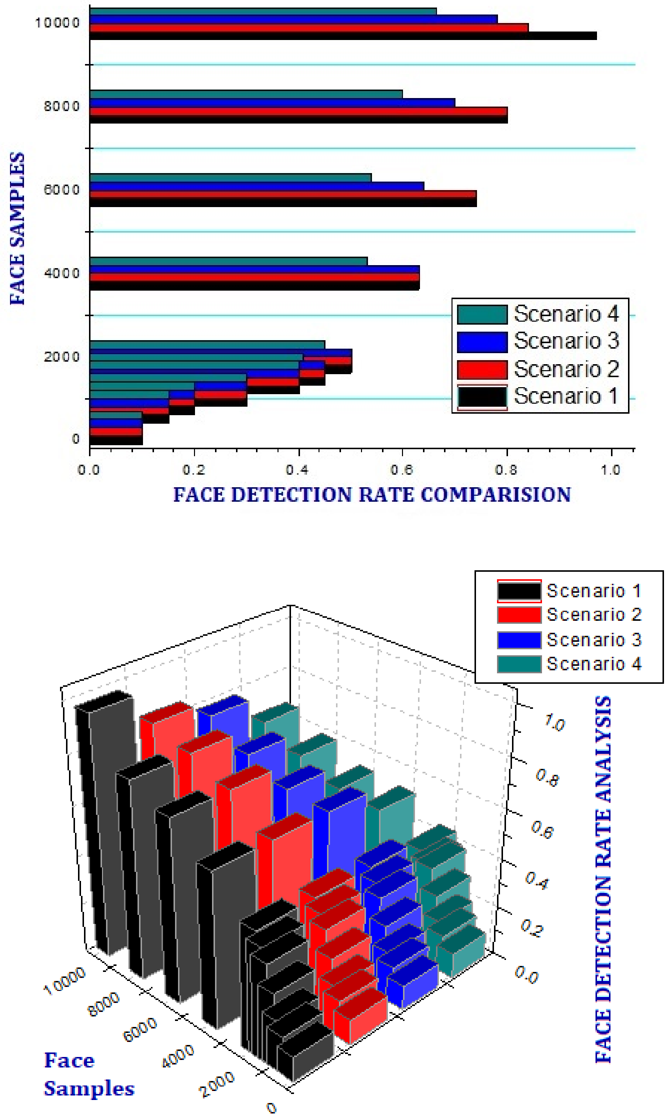


Figure 25. Comparison of face detection rate in various scenarios with 2000 face samples: (**a**) face not obscured, (**b**) face partially covered, (**c**) face fully covered, and (**d**) captured in the dark by an analogue camera.

Analysis of various age-groups used in the conducted rigorous research experiments in all the above-mentioned scenarios is presented in Figure 26, which depicts a variety of age-group samples: (i) <10 years (ii) between (10-15 years) (iii) between (15-20 years) (iv) between (20-35 years) (v) between (35-45 years) (vi) between (45-55 years) (vii) between (55-60 years) and (viii) >60 against corresponding face detection rate and face samples. The total sample size used for all the research experiments was 10,000 images.

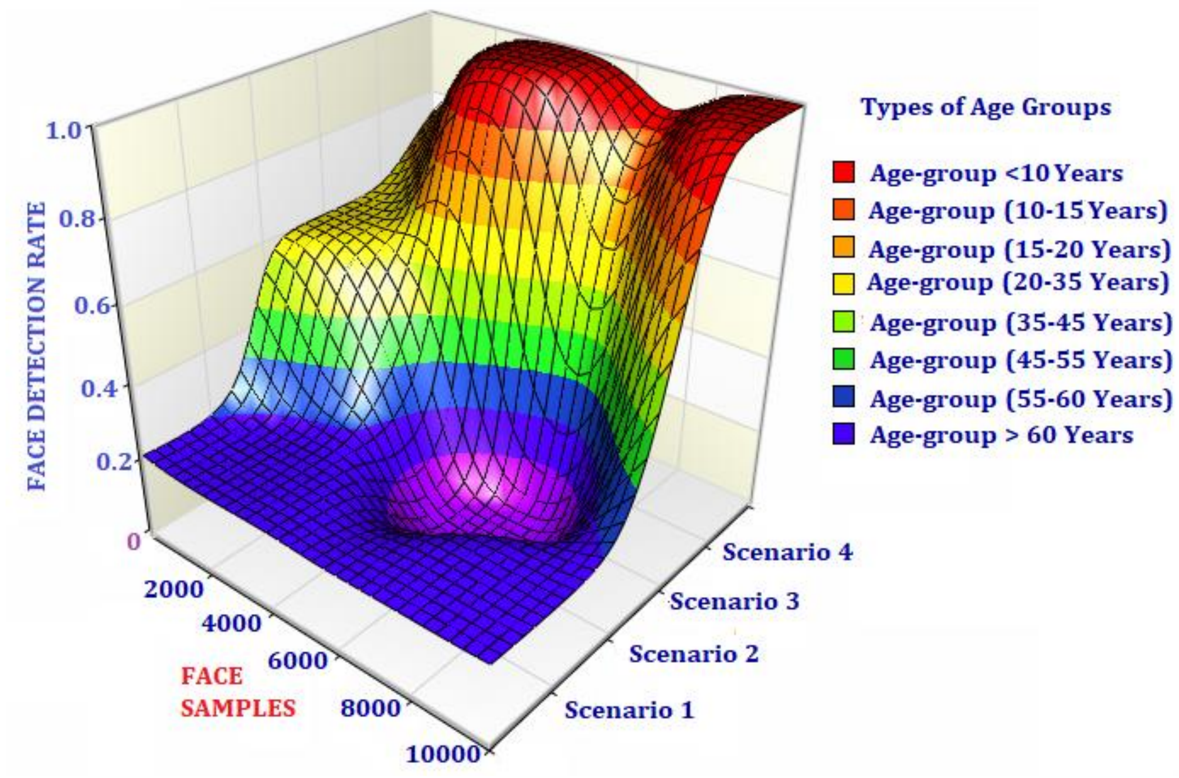


Figure 26. Analysis of face detection for various age groups in three scenarios: : (**a**) face not obscured, (**b**) face partially covered, (**c**) face fully covered, and (**d**) captured in the dark by an analogue camera.

5. Conclusions

This research paper presents an innovative method to prevent smart home theft by providing spontaneous notification of on-going intrusion. The research has provided a novel wireless sensing system for the surveillance and detection of a human intruder as well as instant notification of the intrusion to prevent theft. It eliminates the use of DVR for recording as well as the use of large amounts of memory for storage.

The system can effectively identify a human intruder and prevent false alarms when the intruder is a non-human, by distinguishing between human and non-human objects. All of these processes lead to the instant notification of intrusion by providing real-time notification about the potential theft. The theft system has detected the presence of the intruder even if the face of an intruder is partially or fully hidden with some type of plastic, leather, fiber, or other material. Moreover, this system does not need any supervision.

The information for the caregiver as well as authorized authority is uploaded to a website, either by the local home gateway server or cloud server. New research challenges of security and privacy have arisen due to an increase in products that connect the cyber and physical worlds. It is expected that these research problems will be further resolved in the upcoming future by fellow researchers.

Author Contributions: S. Pandya, H. Ghayvat and V. Patel conceived, designed and performed the experiments; M. H. Yep and P. Gope S.L. contributed to data analysis; S. Pandya analyzed the data and wrote the whole paper; all authors reviewed the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Z. Zhang, D. Yi, Z. Lei, and S. Z. Li, "Regularized Transfer Boosting for Face Detection Across Spectrum," IEEE Signal Processing Letters, vol. 19, no. 3, pp. 131–134, 2012.

[2] W. H. Alobaidi, I. T. Aziz, T. Jawad, F. M. F. Flaih and A. T. Azeez, "Face detection based on probability of amplitude distribution of local binary patterns algorithm," 2018 6th International Symposium on Digital Forensic and Security (ISDFS), Antalya, Turkey, 2018, pp. 1-5.

[3] Z. Jian, Z. Chao, Z. Shunli, L. Tingting, S. Weiwen and J. Jian, "Pre-detection and dual-dictionary sparse representation based face recognition algorithm in non-sufficient training samples," in Journal of Systems Engineering and Electronics, vol. 29, no. 1, pp. 196-202, Feb. 2018.

[4] T. Ahmed, S. Ahmed, S. Ahmed, and M. Motiwala, "Real-Time Intruder Detection in Surveillance Networks Using Adaptive Kernel Methods," 2010 IEEE International Conference on Communications, 2010.

[5] A. Sepas-Moghaddam, F. Pereira and P. L. Correia, "Light Field-Based Face Presentation Attack Detection: Reviewing, Benchmarking and One Step Further," in IEEE Transactions on Information Forensics and Security, vol. 13, no. 7, pp. 1696-1709, July 2018.

[6] H. Zhang, Q. Li, Z. Sun and Y. Liu, "Combining Data-Driven and Model-Driven Methods for Robust Facial Landmark Detection," in IEEE Transactions on Information Forensics and Security, vol. 13, no. 10, pp. 2409-2422, October 2018.

[7] "Identity Theft Research," Detection, Prevention, and Security Identity Theft Handbook, pp. 263–276, 2015.

[8] J. S. Kim, D. H. Yeom, Y. H. Joo, and J. B. Park, "Intelligent Unmanned anti-theft system using network camera," International Journal of Control, Automation and Systems Int. J. Control Autom. Syst., vol. 8, no. 5, pp. 967–974, 2010.

[9] V. V. Jog, D. Jain, R. Arora, and B. Bhat, "Theft prevention ATM model using dormant monitoring for transactions," 2013 IEEE Conference On Information And Communication Technologies, 2013.

[10] H. Li, "Design for home security video monitoring system based on SOPC," null JOURNAL OF ELECTRONIC MEASUREMENT AND INSTRUMENT, vol. 24, no. 3, pp. 294–300, Sep. 2010.

[11] M. Kankanhalli, "Multimedia Surveillance and Monitoring," 2006 IEEE International Conference on Video and Signal Based Surveillance, 2006.

[12] H. Ghayvat, S. Mukhopadhyay, X. Gui, and N. Suryadevara, "WSN-and IoT-Based Smart Homes and Their Extension to Smart Buildings," Sensors, vol. 15, pp. 10350-10379, 2015.

[13] T. Ahmed, S. Ahmed, S. Ahmed, and M. Motiwala, "Real-Time Intruder Detection in Surveillance Networks Using Adaptive Kernel Methods," 2010 IEEE International Conference on Communications, 2010.

[14] Y. Xin et al., "Multimodal Feature-Level Fusion for Biometrics Identification System on IoMT Platform," in IEEE Access, vol. 6, pp. 21418-21426, 2018.

[15] H. Pengfei, T. Huansheng , Q. Tie, X. Yue, L. Xiong, A. K. Sangaiah, "A unified face identification and resolution scheme using cloud computing in Internet of Things, Future Generation Computer Systems", vol. 81, pp. 582-592, 2018.

[16] Nannan Wang, Xinbo Gao, Dacheng Tao, Heng Yang, Xuelong Li, "Facial feature point detection: A comprehensive survey", Neurocomputing, vol. 275, pp. 50-65, 2018.

[17] O. Sharifi, M. Eskandari, "Cosmetic Detection Framework for Face and Iris Biometrics", Sensors, 2018.
D. Nguyen, T. D. Pham, N. R. Baek, K. R. Park, "Combining Deep and Handcrafted Image Features for Presentation Attack Detection in Face Recognition Systems Using Visible-Light Camera Sensors", Sensors, 2018.

[18] C. Yuan and W. Xu, "Multi-object events recognition from video sequences using an extended finite state machine," 2011 4th International Congress on Image and Signal Processing, 2011.

[19] R. Venkatesan, P. D. A. Raja, and A. B. Ganesh, “Unsupervised Learning Based Video Surveillance System Established with Networked Cameras,” Elsevier Journal of Advances in Intelligent Systems and Computing Advances in Signal Processing and Intelligent Recognition Systems, pp. 603–614, 2015.

[20] B. Yang, J. Cao, R. Ni and Y. Zhang, "Facial Expression Recognition Using Weighted Mixture Deep Neural Network Based on Double-Channel Facial Images," in IEEE Access, vol. 6, pp. 4630-4640, 2018.

[21] A. González-Briones, G. Villarrubia, J. F. De Paz, J. M. Corchado, “A multi-agent system for the classification of gender and age from images”, Elsevier Journal of Computer Vision and Image Understanding, 2018.

[22] B. K. Savas, S. Ilkin, and Y. Becerikli, “The realization of face detection and fullness detection in medium by using Haar Cascade Classifiers,” 2016 24th Signal Processing and Communication Application Conference (SIU), 2016.

[23] H. Yang and X. A. Wang, “Cascade classifier for face detection,” Journal of Algorithms & Computational Technology, 2016.

[24] Y. Xin et al., "Multimodal Feature-Level Fusion for Biometrics Identification System on IoMT Platform," in IEEE Access, vol. 6, pp. 21418-21426, 2018.

[25] C. Ding, D. Tao, “Pose-invariant face recognition with homography-based normalization”, Elsevier Journal of Pattern Recognition, vol. 66, pp. 144-152, 2017.

[26] W. Zhou, L. Dai, Y. Zou, X. Zeng, and J. Han, “A High Speed Reconfigurable Face Detection Architecture Based on AdaBoost Cascade Algorithm,” IEICE Transactions on Information and Systems IEICE Trans. Inf. & Syst., vol. E95-D, no. 2, pp. 383–391, 2012.

[27] P. Perera and V. M. Patel, "Efficient and Low Latency Detection of Intruders in Mobile Active Authentication," IEEE Transactions on Information Forensics and Security, vol. 13, no. 6, pp. 1392-1405, June 2018.

[28] J.-G. Park, “Face Detection Algorithm Using the Skin Color and Region of Interest,” jkiit The Journal of Korean Institute of Information Technology, vol. 13, no. 2, p. 53, 2015.

[29] J. Neugebauer, O. Kramer, and M. Sonnenschein, “Improving Cascade Classifier Precision by Instance Selection and Outlier Generation,” Proceedings of the 8th International Conference on Agents and Artificial Intelligence, 2016.

[30] K. Wang, T. Yu, Q. Y. Meng, G. K. Wang, S. P. Li, and S. H. Liu, “Edge Detection from High Resolution Remote Sensing Images using Two-Dimensional log Gabor Filter in Frequency Domain,” IOP Conf. Ser.: Earth Environ. Sci. IOP Conference Series: Earth and Environmental Science, vol. 17, p. 012191, 2014.

[31] W.-Y. Chiu and D.-M. Tsai, “Moving/motionless foreground object detection using fast statistical background updating,” The Imaging Science Journal, vol. 61, no. 2, pp. 252–267, 2013.

[32] N. Shone, T. N. Ngoc, V. D. Phai and Q. Shi, "A Deep Learning Approach to Network Intrusion Detection," in IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 2, no. 1, pp. 41-50, Feb. 2018.