applied system innovation

MDPI

*Article*

# Smart Home Anti-Theft System: A Novel Approach for Near Real-Time Monitoring and Smart Home Security for Wellness Protocol

**Sharnil Pandya [1], Hemant Ghayvat [2,*], Ketan Kotecha [3], Mohammed Awais [2], Saeed Akbarzadeh [2], Prosanta Gope [4], Subhas Chandra Mukhopadhyay [5] and Wei Chen [2]**

1 Computer Science & Engineering Department, Navrachana University, Vadodara 391410, Gujarat, India; sharnil.pandya84@gmail.com
2 Center for Intelligent Medical Electronics, Fudan University, Shanghai 200433, China; 17110720061@fudan.edu.cn (M.A.); sd.akbarzadeh@gmail.com (S.A.); w_chen@fudan.edu.cn (W.C.)
3 Director, SIT, Symbiosis International University, Pune,411004, India; director@sitpune.edu.in
4 School of Engineering and Computer Science, University of Hull, Hull, HU1 1DB, UK; p.gope@hull.ac.uk
5 Mechanical/Electronics Engineering, Macquarie University, Sydney, NSW 2109, Australia; subhas.mukhopadhyay@mq.edu.au
* Correspondence: ghayvat@gmail.com; Tel.: +86-132-6285-4329

Received: 19 September 2018; Accepted: 16 October 2018; Published: 23 October 2018


**Abstract:** The proposed research methodology aims to design a generally implementable framework for providing a house owner/member with the immediate notification of an ongoing 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) and Cognitive Internet of Things are expanding smart home concepts and solutions, and their applications. The present research proposes 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 capability. 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. The investigation results validate the success of the proposed system. The system accuracy has been enhanced to 97.01%, 84.13, 78.19%, and 66.5%, in scenarios where a detected intruder had not hidden his/her face, hidden his/her face partially, fully, and was detected in the dark from 85%, 64.13%, 56.70%, and 44.01%.

**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 [3–5]. In addition, it is a tedious task to go through all the recorded video clips after a possible theft has become known. It might be 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 Figures 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.

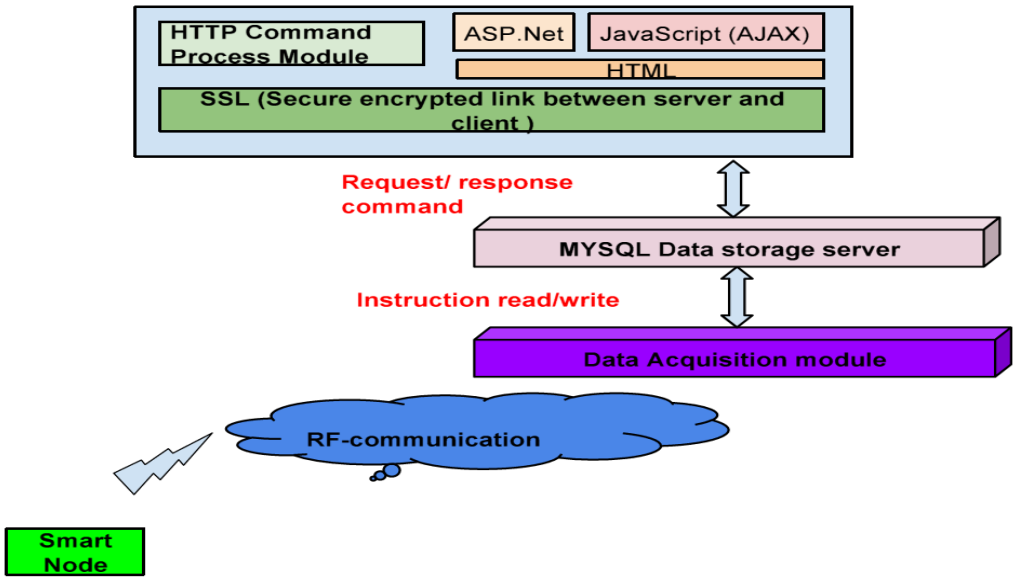
The work is presented as follows. The initial section provides an introduction on legacy systems, potential issues, and their impact on society. Section 2 describes the necessity of the present research. It also describes the intelligent features of the system that can detect obscured faces using graphical 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.

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


(**a**)

(**b**)

**Figure 2.** *Cont.*




(**c**)

**Figure 2.** Detailed workflow of the system. (**a**) CP Plus 850 tvl analog camera; (**b**) customized hardware in which the proposed research methodology is coded; and (**c**) the layered architecture of an anti-theft system for an Internet of Things (IoT)-based smart home solution.

## 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-cutters, smart anti-lock systems, and many more. For such intruders, it is straightforward to disconnect 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 [5–27] and propose an intelligent approach that can not only provide unsupervised human activity monitoring, but can also stop an ongoing 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.

**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 |

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].

## 3. Design and Experimental Setup

As shown in Figure 2, a CP-PLUS analog camera model 850 tvl, a Wi-Fi- and Bluetooth-enabled customized hardware with 1 GB RAM and a 256 GB memory card, and a Samsung Grand mobile device were used. The camera was placed in each room of the smart home to protect the house from 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 Bluetooth and 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 3, 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 4. 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 second for up to 40 seconds and stores them in mp4 video format. The captured 40 s mp4 video is then compressed to approximately 10 seconds 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 a 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.

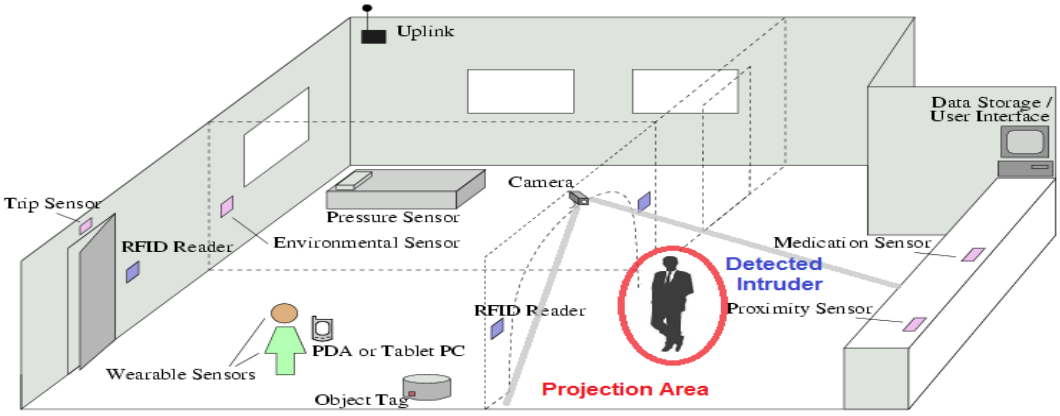



**Figure 3.** Projection area of the smart home anti-theft system.

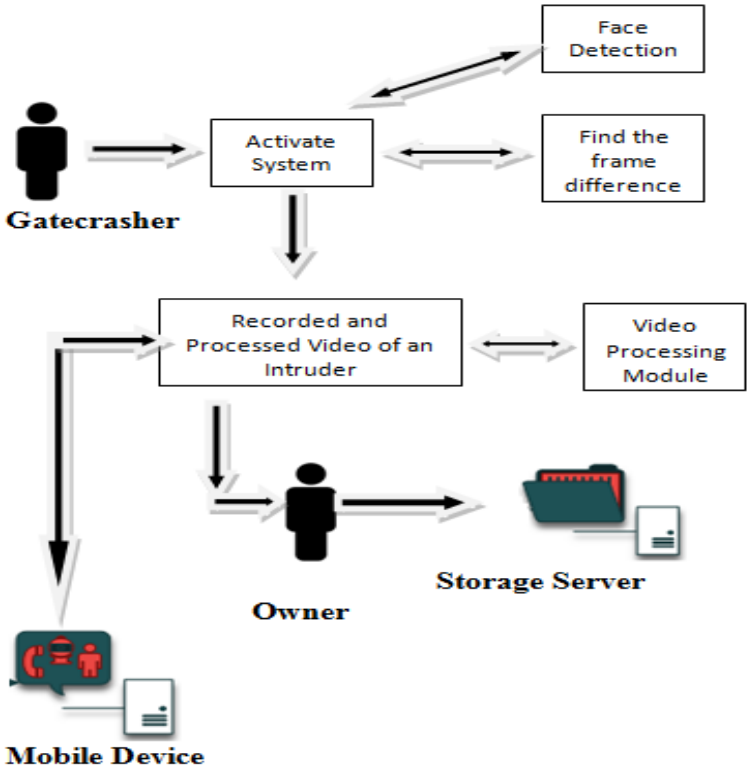



**Figure 4.** Detailed workflow of the system.

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 3 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 submodules, as shown in Figure 5.

   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 [28,29]. 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 main reason behind increasing the speed of the captured frames from 15 to 45 is to allow the house owner to access the captured video sequences in slow internet access scenarios. 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 [13,14] 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.

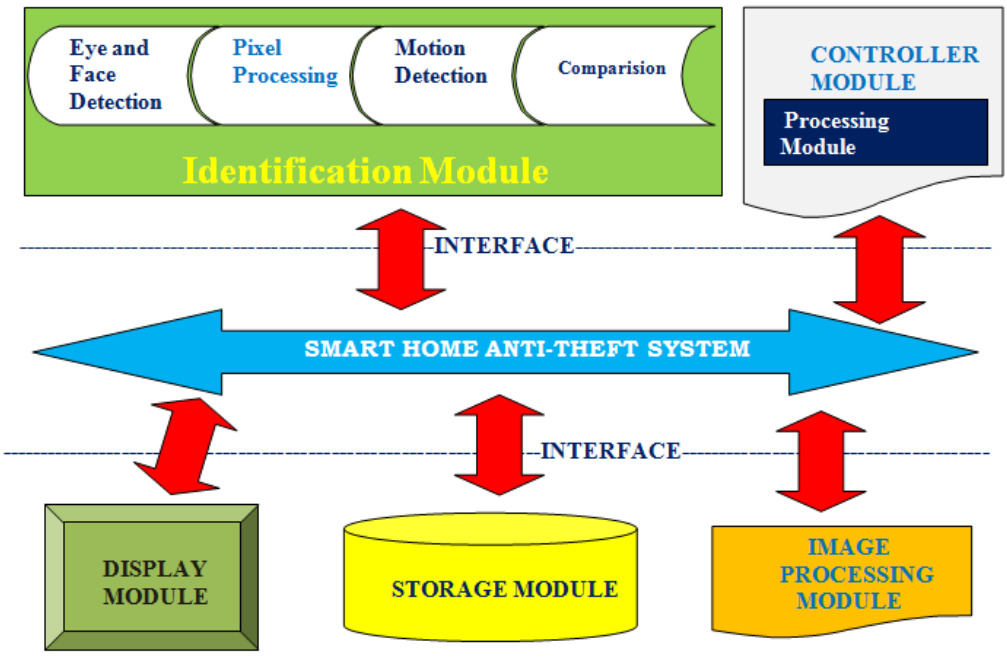



**Figure 5.** Communication block diagram of the smart home anti-theft system.

## 4. Methodologies and Results Analysis

According to our rigorous and detailed literature survey [3–40], we have identified that to date, no feasible solutions are available which can detect intruders with partially or fully covered faces and provide immediate notification to the house owner to stop the ongoing theft. Furthermore, it was

also identified that the main reason why such methodologies have not been introduced in the market is due to the complexity, efficiency, accuracy, and the time-duration required to develop specialized methodologies for the intrusion detection with partially or fully covered faces. To overcome the above challenges we have proposed a research methodology that aims to design a generally implementable framework for providing a house owner/member with the immediate notification of an ongoing theft (unauthorized access to their premises). The research methodology was classified into three stages with corresponding results: (a) primary phase; (b) secondary phase; and (c) final phase.

### *4.1. Primary Phase*

The eye and face detection module was programmed with the modified feature-based Haar cascade face detection classifier algorithm [13,22,26,28,31,32]. 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 6. As shown in Figures 6 and 7, the rectangle features can be captured and computed based on the intermediate representation of the captured intruder's image. The results of the face detection rate analysis of Haar Cascade classifier with existing methodologies is represented in Table 2, which shows that the accurate face detection rate of the modified Haar cascade classifier was very high when compared to previous methodologies for approximately 5000 face samples in ordinary scenarios where intruders have not obscured their faces. This is the primary reason why the modified Haar cascade classifier was used in the present research experiments. 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 in the following equation.

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

where $ii(a, b)$ is the detected image of an intruder, $i(a', b')$ is the original image, $(a, b)$ represents $x$ and $y$ coordinates of the detected image, $(a', b')$ represents $x$ and $y$ coordinates of the original image, and S represents sum of the pixels. The initial condition for $S$ is set to zero when the proposed system is in the idle state or has not captured the trespasser in the smart home using the following pair of recurrences.

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

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

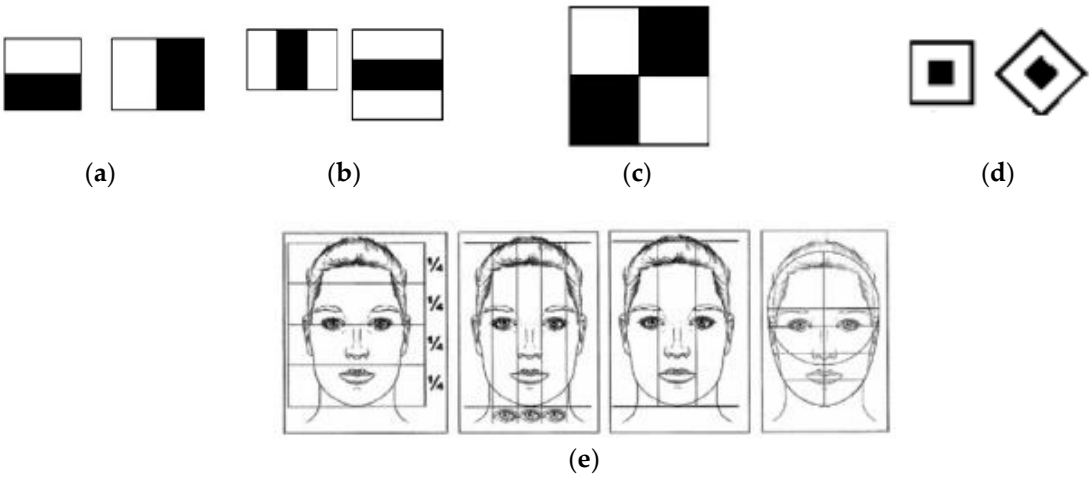



**Figure 6.** Haar cascade classifiers, (**a**) edges features; (**b**) lines features; (**c**) four rectangle features; (**d**) center-surround features; and (**e**) face classifiers.

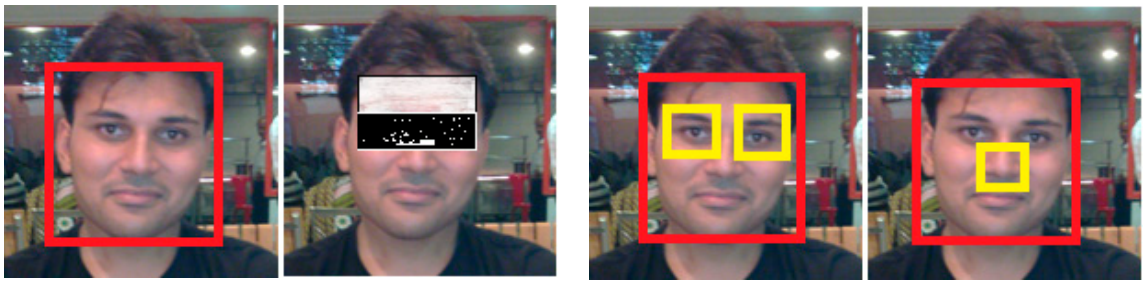


**Figure 7.** Haar cascade image classifier samples (all of the shown human faces in the conducted experiments required copyright permission obtained on 19 March 2017).

**Table 2.** Face Detection Rate Accuracy Analysis of the modified Haar Cascade Classifier with existing methodologies in a scenario, where an intruder has not occluded their face.

| Types of Face Detection Methodologies | Face Detection Rate Accuracy in an Ordinary Scenario (%) [Total Sample Size = 5000] |
|---|---|
| Modified Haar Cascade | 97.01 |
| Neural Networks | 75.4 |
| Skin Segmentation Template | 52.3 |
| Matching Modified Haar Cascade | 65.87 |
| Higher Order Statistics | 72 |
| Two Step R-CNN Method | 94 |

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 [27], 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

$$x = \left(cx - \left(\frac{2r}{5}\right)\right) \text{ and } y = \left(cy + \left(\frac{r}{3}\right)\right) \quad (4)$$

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

The main reason of the usage of cascade classification represented in Equation (4) in the conducted experiments is to allow the proposed system to detect multiple faces, which includes detection of overlapping face patterns by distinguishing strong and weak classifiers as shown in Figures 8 and 9a. Here, the feature value is defined as the total difference between sum of the pixels S in the white rectangle and black rectangle as shown in Figure 6. The algorithm defines weak classifiers by thresholding captured feature values. The modified Haar Cascade classification algorithm trains strong classifiers to achieve very low false acceptance and false recognition rates. As shown in Figure 8, 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 subwindow. In this phase, numerous classifier stages are built using the modified Haar cascade to minimize false negatives.

In comparison to the existing methodologies, this framework was 15 times faster than the other recognition frameworks [31]. As per the face location framework portrayed by Rowley et al., a Multiclassifier-based Near-Real-Time Face Detection System is the fastest algorithm in detecting human faces as compared to existing methodologies. 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 9a. Figure 9b represents non-human object samples which were not captured by the smart home anti-theft system during the conducted experiments. To conduct non-human object detection experiments, samples of dogs, cats, sparrows, and parrots were used as described in Figure 9b.

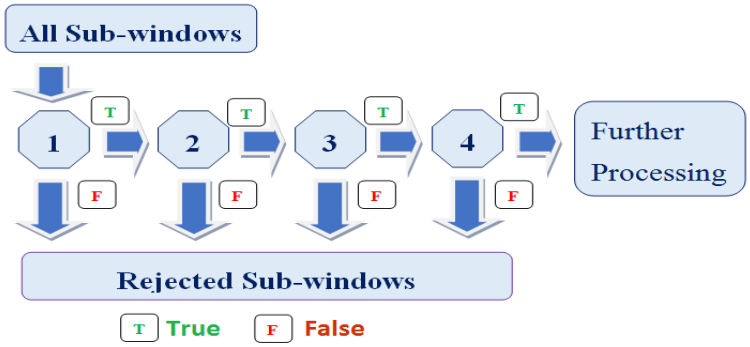
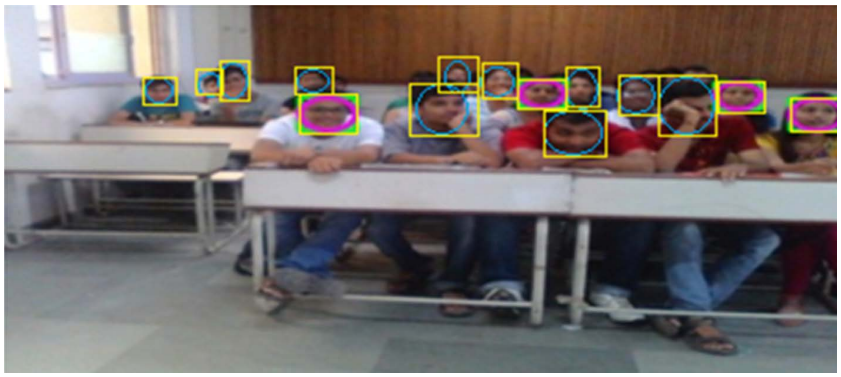
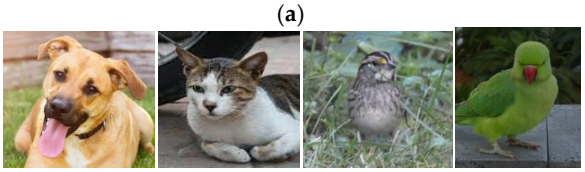



**Figure 8.** Schematic representation of cascade classifiers.



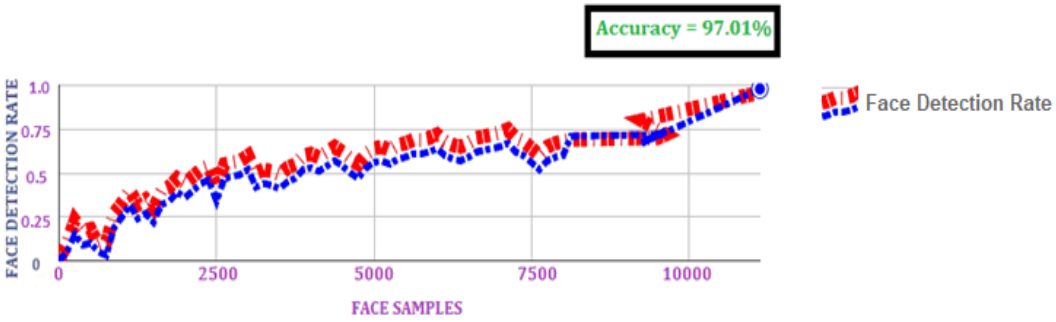
(**a**)



(**b**)

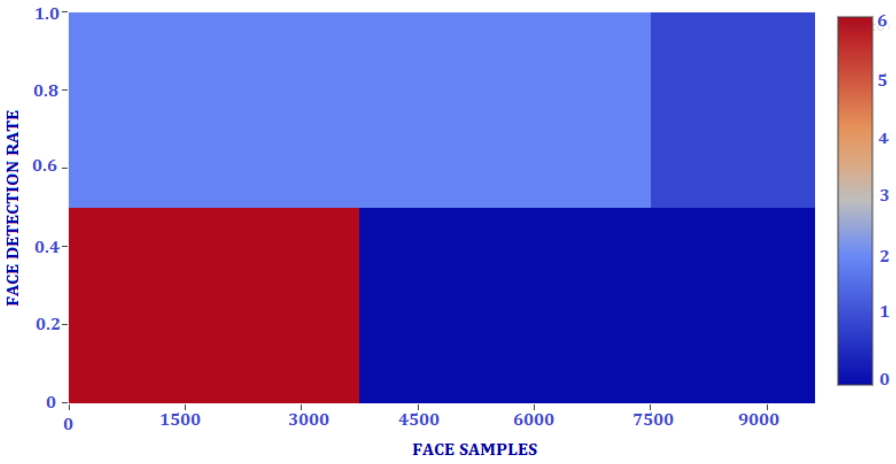
**Figure 9.** (**a**) Crowd Cascade Image Classifier Sample (copyright permission for photograph obtained on 1 April 2017). (**b**) Non-human objects samples (copyright permission for photograph obtained on 9 July 2017).

As shown in Figure 10a, rigorous research experiments were conducted with approximately 10,000 diverse samples in scenarios where an intruder was detected in ordinary scenarios and where an intruder had not fully/partially covered his/her face. The ratio of accurate face detection was 100% in more than 97% of cases for more than 10,000 face samples. Figure 10b depicts the face detection rate analysis of 97% accuracy cases in the form of heat map representation.




(**a**)

**Figure 10.** *Cont.*




(**b**)

**Figure 10.** (**a**) Face detection rate analysis of the proposed system with 10,000 face samples in an ordinary scenario where an intruder has not occluded their faces (**b**) Heat map representation of 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.

### *4.2. Secondary Phase*

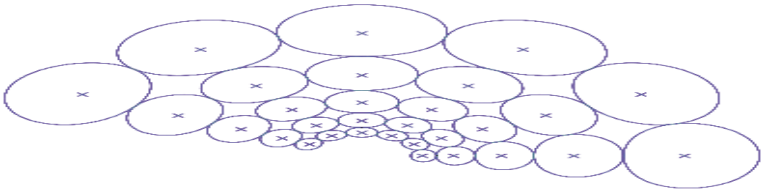
After initial face detection is complete, an intelligent secondary stage algorithm was designed to eliminate false positives which capture half-hidden human faces, based on the brightest features of the face [26,27,33–38]. For each positive sample, a dynamic threshold (range from 0.5 to 0.7) was set to capture the brightest part of the intruders partially or fully covered faces. After conducting numerous experiments, it has been identified that the proposed system does not require different dynamic threshold values for day-time or night-time intrusions. For detecting the retinal point and other bright facial features, a modified Gabor filter was used [38]. The Gabor filter is a powerful tool for heterogeneous image feature extraction. The Gabor filter can extract common features from diverse face images. Gabor filters make the use of Gabor wavelets which are very powerful in detecting and discriminating image feature extraction. Gabor filters also possess similar characteristics as those of a human visual system [19].

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 [20]. Figure 11a,b 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

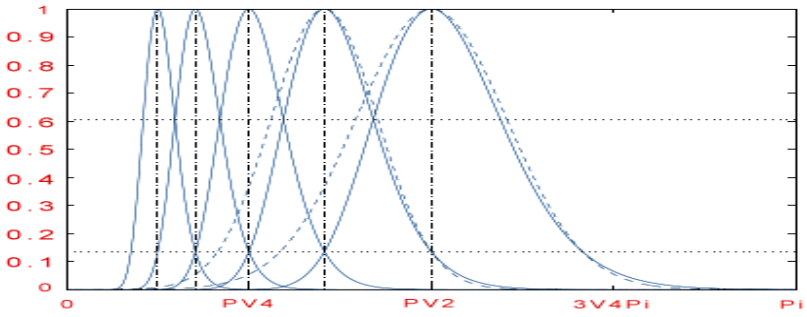
$$\omega_0 = \exp(\delta_0)$$

$$G(\omega, \delta) = \left( \frac{A \exp\left\{ (\delta - \delta_0)^2 \right\}}{2{\sigma_\delta}^2} \right) \times \exp\left( \frac{-\left\{ (\varphi - \varphi_0)^2 \right\}}{2{\sigma_\varphi}^2} \right) \quad (5)$$

where $A$ is a normalization constant and $(\delta, \phi)$ are polar frequency coordinates.



(**a**)




(**b**)

**Figure 11.** (**a**) Gabor filters in the Fourier domain, level curves, and (**b**) cross-sectional plot.

Figure 12 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 13a 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 13a reports the correct face detection rate analysis for 10,000 humans under these conditions. The correct face detection rate accuracy achieved was more than 84%, with similar rates achieved when more than 10,000 images were used. Figure 13b represents the face detection rate analysis of 84% accuracy cases in the form of a heat map representation.

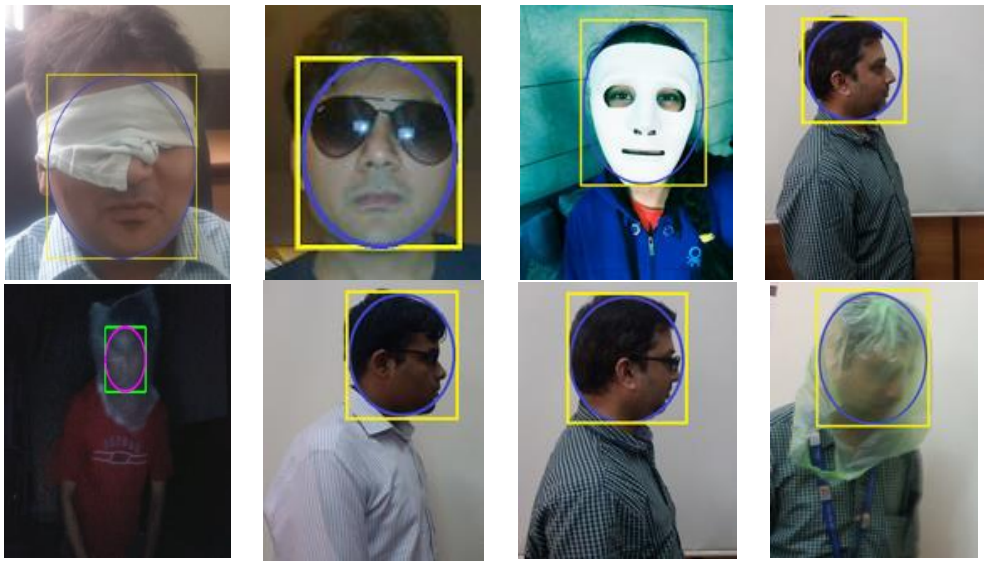
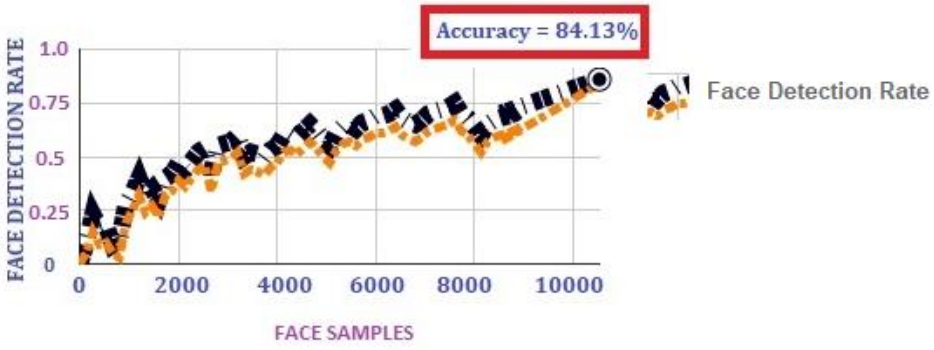
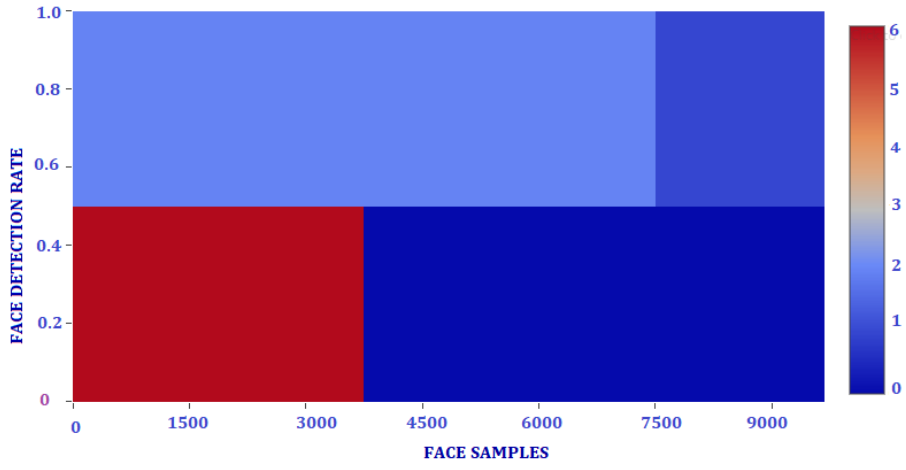


**Figure 12.** *Cont.*

**Figure 12.** 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, and face covered in a transparent cloth (copyright permission for images obtained on 21 May 2017).




(**a**)




(**b**)

**Figure 13.** (**a**) 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 material. (**b**) Heat map representation of 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.

### *4.3. 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

$$M(a,b) = \begin{cases} 1, if|I_t(a,b) - B(a,b)| > T \\ 0, if|I_t(a,b) - B(a,b)| \leq T \end{cases} \quad (6)$$

Here, $T$ is the predefined 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 that it contains background. Generally, dynamic pixels are named "1", which is assigned for moving objects. 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 $\mu_{xy}$ 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 given by a standard deviation $\sigma_{ab}$ as follows.

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

$$\sigma_{ab} = \left(\frac{1}{k}\sum_{k=0}^{k-1}\left\{(I_k(a,b) - \mu_{ab})^2\right\}^{1/2}\right) \quad (7)$$

In order to achieve object motion detection, the difference between the background of the image and captured video frame was calculated. Here, the predefined parameter is $\lambda$. 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.

$$M(a,b) = \begin{cases} 1, if\left|I_t(a,b) - B(a,b)\right| > \lambda\sigma_{xy} \\ 0, if\left|I_t(a,b) - B(a,b)\right| \leq \lambda\sigma_{xy} \end{cases} \quad (8)$$

Figure 14 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 14 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.

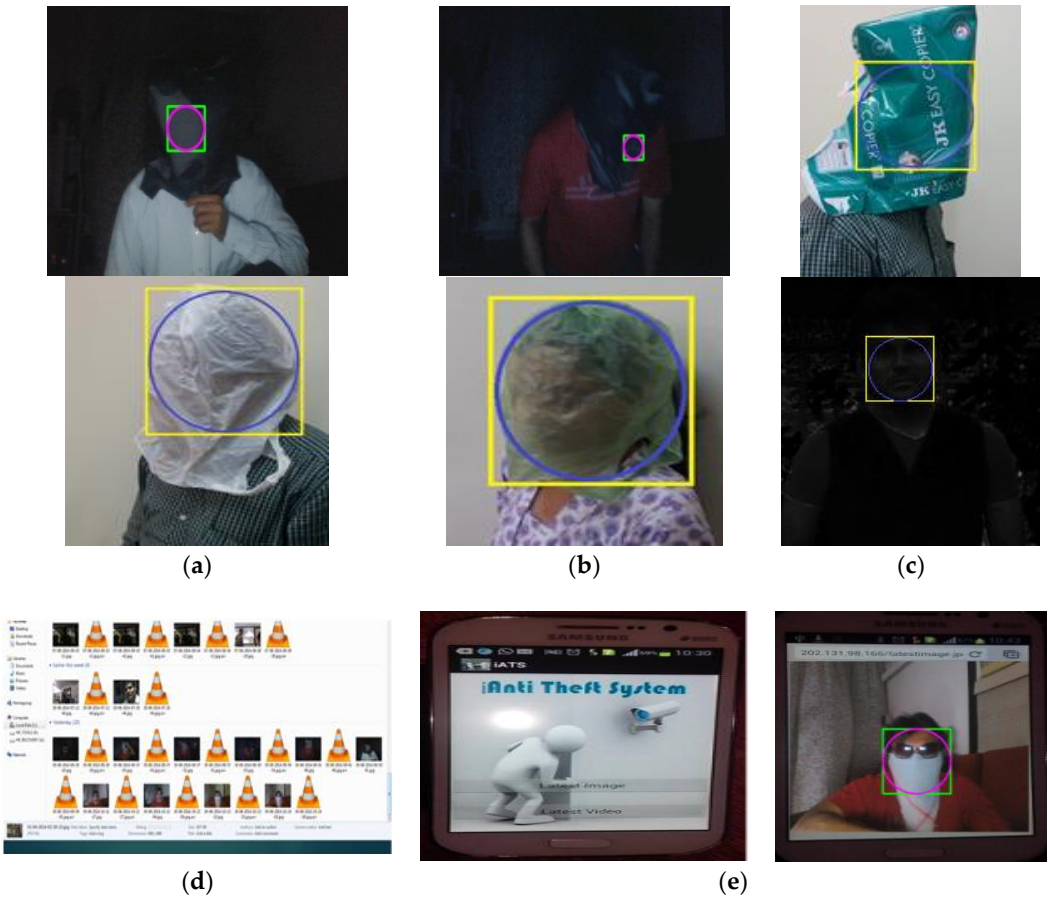
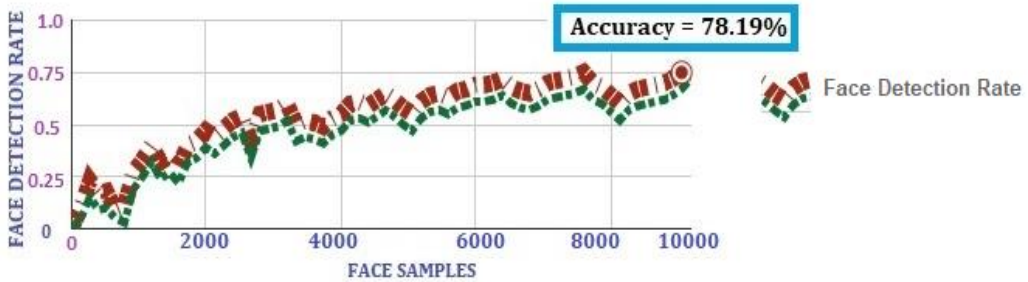


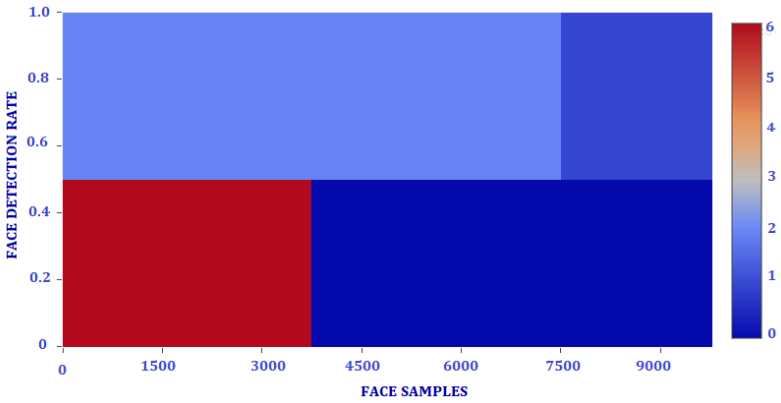

**Figure 14.** Final Stage Classifier Samples where an intruder has a fully covered face with some type of transparent, solid, plastic, leather material and detection in the dark night. Face hidden with (**a**) black plastic, (**b**) black plastic and in motion, (**c**) face captured on a dark night, (**d**) storage samples of captured intruder's image and a compressed video with a timestamp, and (**e**) a customized mobile app with samples of ongoing theft notification (copyright permission was obtained for all of the human faces shown on 11 May 2017).

As shown in Figure 15a, 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 with accuracy of approximately 78% as shown in Figure 15a. Figure 15b represents face detection rate analysis of 78% accuracy cases in the form of the heat map representation.




(**a**)

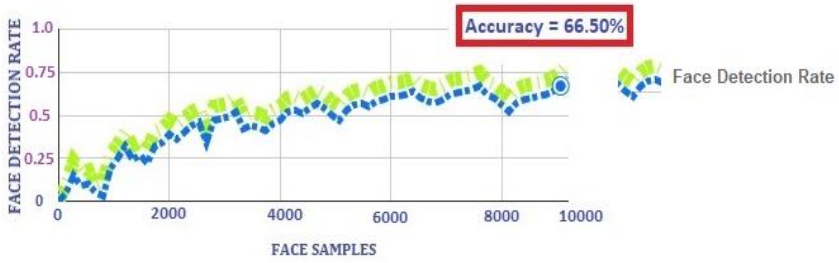
**Figure 15.** *Cont.*



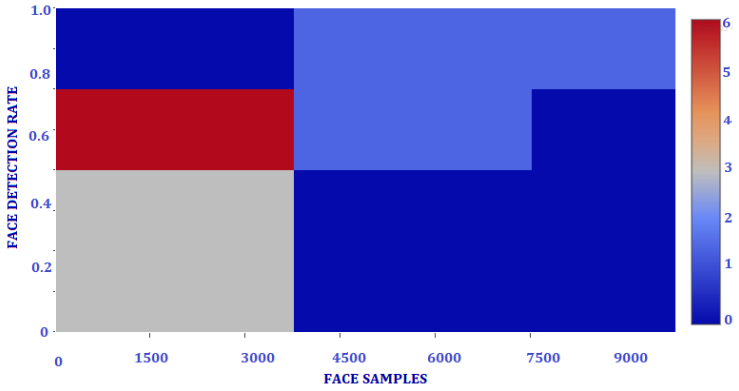

(**b**)

**Figure 15.** (**a**) 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. (**b**) Heat map representation of face detection rate analysis of the proposed system with 10,000 face samples in scenarios where an intruder has a fully covered face.

Figure 16a depicts the correct face detection rate analysis against 10,000 images of humans detected at night by an ordinary analog camera without night vision capability. Faces were accurately detected in approximately 67% of cases, with similar accuracy achieved for high numbers of images. Figure 16b represents face detection rate analysis of detected 67% cases in form of the heat map representation.



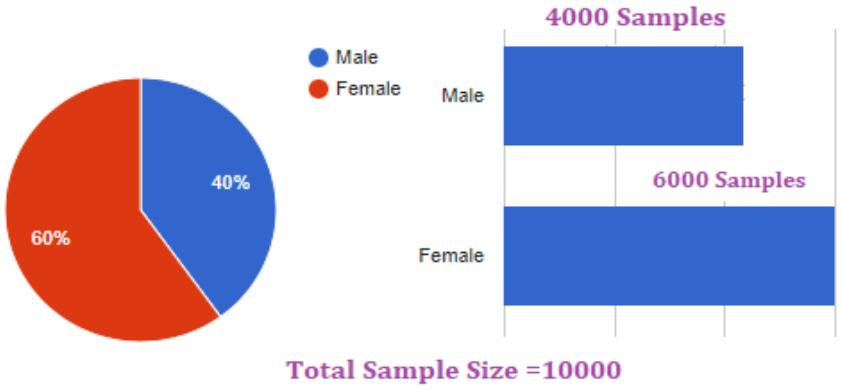

(**a**)



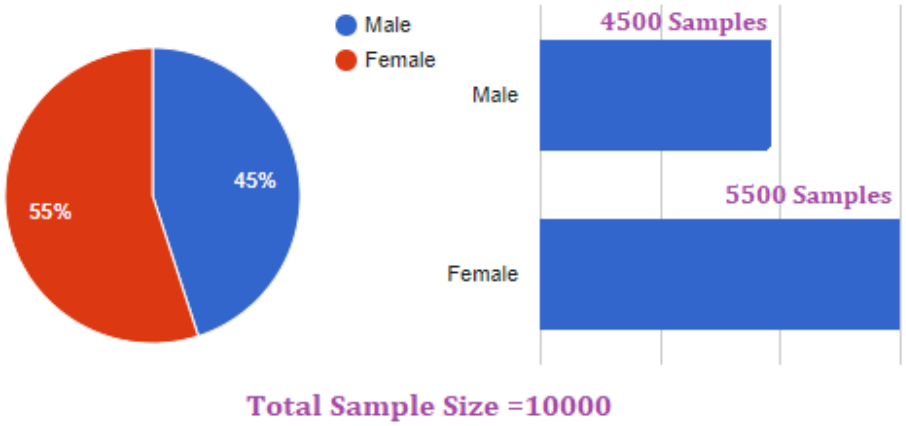

(**b**)

**Figure 16.** (**a**) 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. (**b**) Heat map representation of 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.

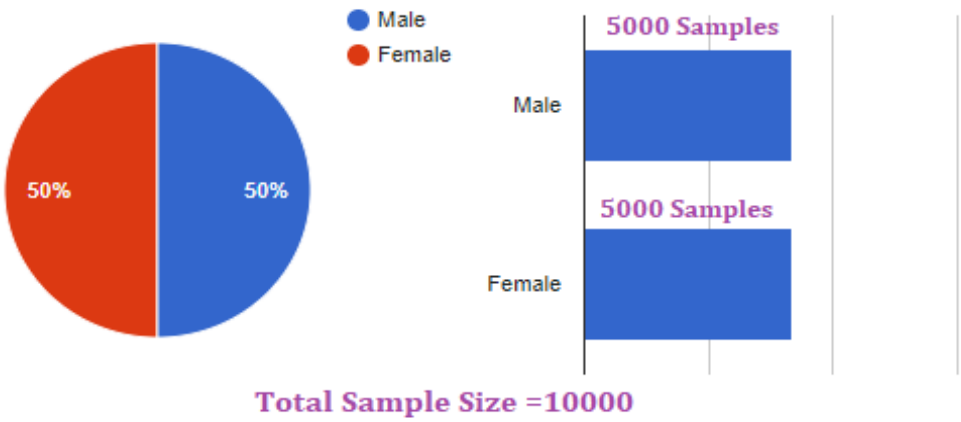
Analysis of the male and female ratio used in the conducted rigorous research experiments in all the above-mentioned scenarios is presented in Figure 17, which depicts a variety of male and female samples in the following scenarios: (a) face not obscured, (b) face partially covered, (c) face fully covered, and (d) captured in the dark by an analog camera.The total sample size used for all the research experiments was 10,000 images.




(**a**)

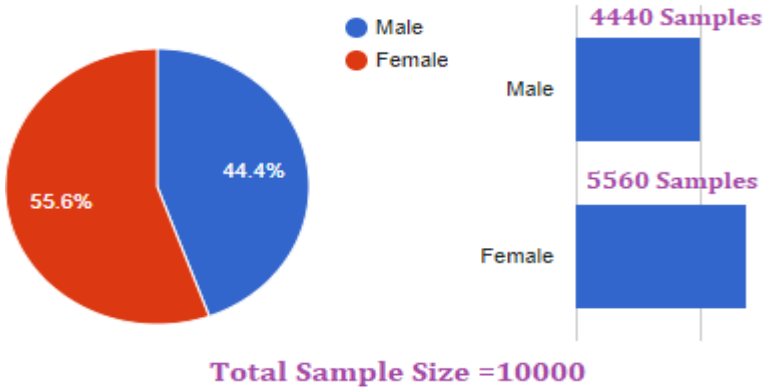



(**b**)




(**c**)

**Figure 17.** *Cont.*




(**d**)

**Figure 17.** Analysis of male–female ratio where the face is (**a**) face not obscured, (**b**) face partially covered, (**c**) face fully covered, and (**d**) captured in the dark by an analog camera.

An analysis of various skin tones: dark, whitish-brown, whitish, fair, very fair, and white is given in Figure 18. The proposed system achieved high accuracy for all varieties of skin tones and for all the above-mentioned scenarios, as shown in Table 3. It clearly depicts that the conducted experiments on a variety of skin tones such as white, very fair, fair, wheatish-brown, wheatish, and dark achieved above 80% accuracy when an intruder has partially hidden their face, above 70% accuracy when an intruder has fully hidden their 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 analog camera without night vision capability.

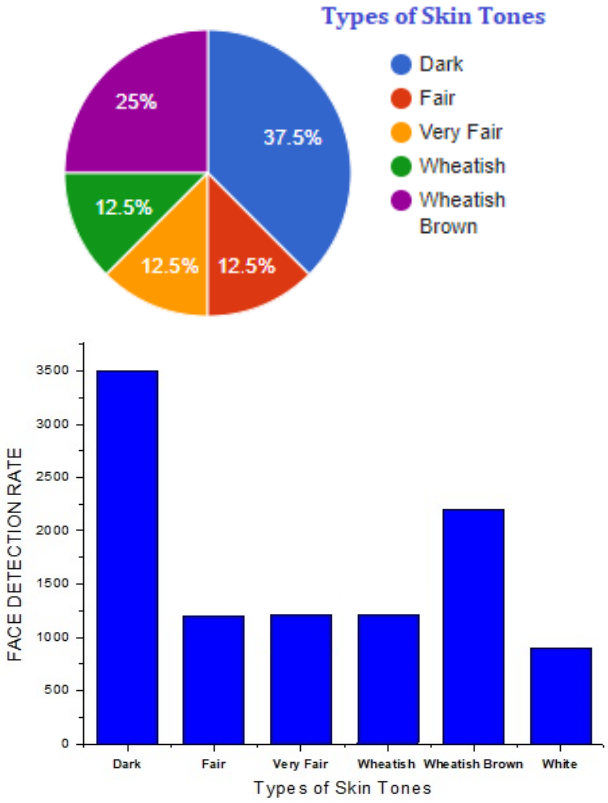



**Figure 18.** Detailed Analysis of various skin tones samples used in scenarios: face not obscured, face partially covered, face fully covered, and captured in the dark by an analog camera.

**Table 3.** 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 19, a comparison was carried out between all the scenarios mentioned in Figure 12, Figure 13, Figure 15, and Figure 16 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 20 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 analog camera.

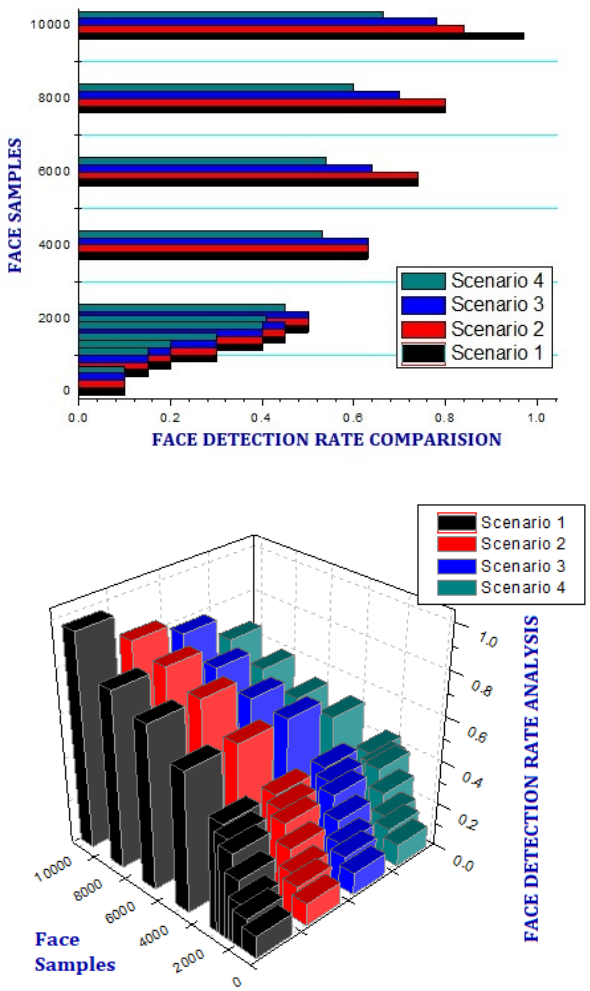



**Figure 19.** Comparison of face detection rate in various scenarios with 2000 face samples: face not obscured; face partially covered; face fully covered; and captured in the dark by an analog camera.

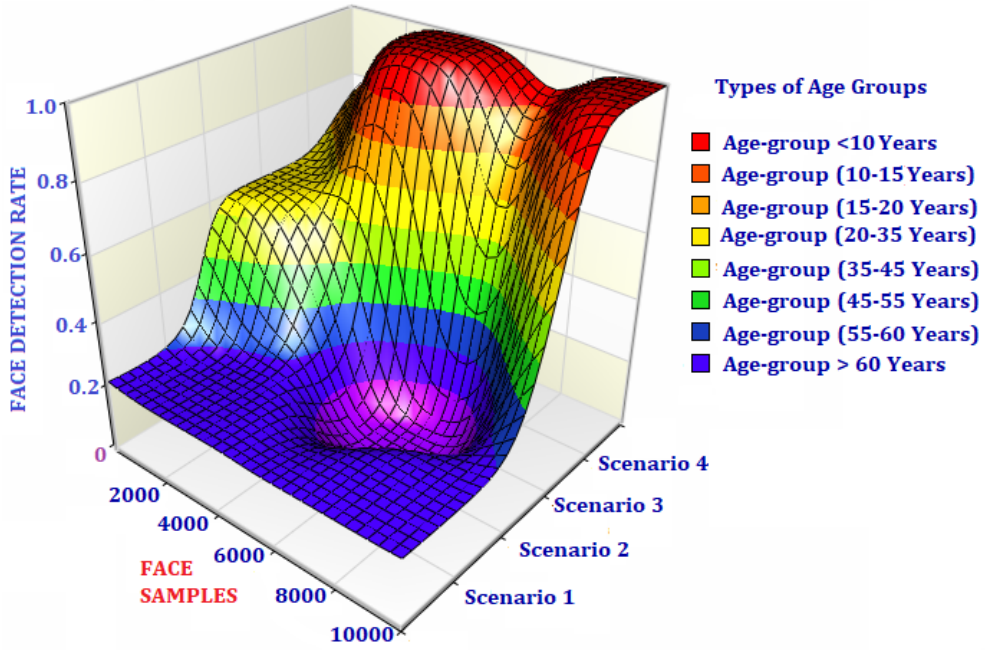



**Figure 20.** Comparision Analysis of face detection for various age groups in four scenarios: face not obscured; face partially covered; face fully covered; and captured in the dark by an analog camera.

Analysis of the various age groups used in the conducted rigorous research experiments in all the above-mentioned scenarios is presented in Figure 20, 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.

Table 4 represents a face detection rate accuracy analysis of the modified Haar Cascade Classifier with existing methodologies in four scenarios: (a) face not obscured, (b) face partially covered, (c) face fully covered, and (d) captured in the dark by an analog camera. In the case of the first two scenarios, where the intruder has not obscured or partially obscured their face, the proposed system has enhanced system accuracy to 97.01% and 84.13% from 95% and 80.8%, to 75.6% and 65.5% from 69% and 34.5%, and to 72% and 50% from 94% and 75% resulting from the use of neural networks, a skin segmentation template, matching modified haar cascade methodologies, a two step R-CNN, higher order statistics, and faceness-net methodologies. However, in scenarios 3 and 4, where the intruder has fully obscured their face using some type of transparent, solid, paper, plastic, or leather material and the intruder has been captured in the dark using an analog camera without night vsion facility, all the existing methodologies have failed in detecting the intruder except for the neural networks and two step R-CNN method. Howver, the two step R-CNN method requires the usage of a thermal camera to detect an intruder in the dark night [20]. The proposed system can detect an intruder using an analog camera without night vision capability. In scenarios 3 and 4, the proposed system has enhanced accuracy to 78.19% and 66.5% from 56.70%, 44.01% and 69%, 55% as represented in Table 4.

**Table 4.** Face Detection Rate Accuracy Analysis of the modified Haar Cascade Classifier with existing methodologies in four scenarios: (**a**) face not obscured; (**b**) face partially covered; (**c**) face fully covered; and (**d**) captured in the dark by an analog camera.

| Methodologies | 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 Material (%)** | **Scenario 3: Intruder Has Fully Covered Face with Some Type of Transparent, Solid, Plastic, Leather Material (%)** | **Scenario 4: Intruder Is Captured in the Dark Night or in Bad Light Conditions (%)** |
| Modified Haar Cascade | 97.01 | 84.13 | 78.19 | 66.5 |
| Neural Networks | 95 | 80.8 | 56.70 | 44.01 |
| Skin Segmentation Template | 75.6 | 65.5 | CD3 | CD4 |
| Matching Modified Haar Cascade | 69 | 34.5 | CD3 | CD4 |
| Two Step R-CNN | 95 | 75 | 69 | 55 |
| Higher Order Statistics | 72 | 50 | CD3 | CD4 |
| Faceness-Net | 94 | 75 | CD3 | CD4 |

CD3—Can not detect intruder in Scenario 3, CD4—Can not detect intruder in Scenario 4.

## 5. Conclusions

This research paper presents an innovative method to prevent smart home theft by providing spontaneous notification of ongoing 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 performance evaluation parameters of the Smart Home Antitheft System for intruder detection are recorded for the four different scenarios: (a) face not obscured, (b) face partially covered, (c) face fully covered, and (d) captured in the dark by an analog camera. 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. The main advantage of the proposed system is that it requires only 50% cost of the DVR and other surveillance-based solutions available on the market.

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. If an intruder disables WiFi connection using DoS attack then the proposed system will not be able to notify the house members about the ongoing theft. However, the proposed system is equipped with Bluetooth network, which can still record the ongoing theft but cannot send the notification to the house owner due to the lack of WiFi/Internet connections. 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.P., H.G., K.K. and W.C. conceived, designed, and performed the experiments; the rest contributed to data analysis; all authors reviewed the manuscript.

**Funding:** This work is supported by Shanghai Municipal Science and Technology International R&D Collaboration Project (Grant No. 17510740500) and National Key R&D Program of China (Grant No. 2017YFE0112000).

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

## References


1. Zhang, Z.; Yi, D.; Lei, Z.; Li, S.Z. Regularized Transfer Boosting for Face Detection Across Spectrum. *IEEE Signal Process. Lett.* **2012**, *19*, 131–134. [CrossRef]
2. Alobaidi, W.H.; Aziz, I.T.; Jawad, T.; Flaih, F.M.F.; Azeez, A.T. Face detection based on probability of amplitude distribution of local binary patterns algorithm. In Proceedings of the 2018 6th International Symposium on Digital Forensic and Security (ISDFS), Antalya, Turkey, 22–25 March 2018; pp. 1–5.
3. Jian, Z.; Chao, Z.; Shunli, Z.; Tingting, L.; Weiwen, S.; Jian, J. Pre-detection and dual-dictionary sparse representation based face recognition algorithm in non-sufficient training samples. *J. Syst. Eng. Electron.* **2018**, *29*, 196–202.
4. Ahmed, T.; Ahmed, S.; Ahmed, S.; Motiwala, M. Real-Time Intruder Detection in Surveillance Networks Using Adaptive Kernel Methods. In Proceedings of the 2010 IEEE International Conference on Communications, Cape Town, South Africa, 23–27 May 2010.
5. Yang, S.; Luo, P.; Loy, C.C.; Tang, X. Faceness-Net: Face Detection through Deep Facial Part Responses. *IEEE Trans. Pat. Anal. Mach. Intell.* **2018**, *40*, 1845–1859. [CrossRef] [PubMed]
6. Sepas-Moghaddam, A.; Pereira, F.; Correia, P.L. Light Field-Based Face Presentation Attack Detection: Reviewing, Benchmarking and One Step Further. *IEEE Trans. Inf. Forensics Secur.* **2018**, *13*, 1696–1709. [CrossRef]
7. Zhang, H.; Li, Q.; Sun, Z.; Liu, Y. Combining Data-Driven and Model-Driven Methods for Robust Facial Landmark Detection. *IEEE Trans. Inf. Forensics Secur.* **2018**, *13*, 2409–2422. [CrossRef]
8. Biegelman, M.T. *Identity Theft Handbook: Detection, Prevention, and Security*; John Wiley & Sons: Hoboken, NJ, USA, 2015; pp. 263–276. ISBN 978-1-119-20316-2.
9. Kim, J.S.; Yeom, D.H.; Joo, Y.H.; Park, J.B. Intelligent Unmanned anti-theft system using network camera. *Syst. Int. J. Control Autom. Syst.* **2010**, *8*, 967–974. [CrossRef]
10. Jog, V.V.; Jain, D.; Arora, R.; Bhat, B. Theft prevention ATM model using dormant monitoring for transactions. In Proceedings of the 2013 IEEE Conference on Information And Communication Technologies, Azerbaijan, Baku, 23–25 October 2013.
11. Li, H. Design forhome security video monitoring system based on SOPC. *J. Electron. Meas. Instrum.* **2010**, *24*, 294–300. [CrossRef]
12. Ghayvat, H.; Mukhopadhyay, S.; Gui, X.; Suryadevara, N. WSN-and IoT-Based Smart Homes and Their Extension to Smart Buildings. *Sensors* **2015**, *15*, 10350–10379. [CrossRef] [PubMed]
13. Xin, Y.; Kong, L.; Liu, Z.; Wang, C.; Zhu, H.; Gao, M.; Zhao, C.; Xu, X. Multimodal Feature-Level Fusion for Biometrics Identification System on IoMT Platform. *IEEE Access* **2018**, *6*, 21418–21426. [CrossRef]
14. Pengfei, H.; Huansheng, T.; Tie, Q.; Yue, X.; Xiong, L.; Sangaiah, A.K. A unified face identification and resolution scheme using cloud computing. *Internet Things Future Gener. Comput. Syst.* **2018**, *81*, 582–592.
15. Wang, N.; Gao, X.; Tao, D.; Yang, H.; Li, X. Facial feature point detection: A comprehensive survey. *Neurocomputing* **2018**, *275*, 50–65. [CrossRef]
16. Sharifi, O.; Eskandari, M. Cosmetic Detection Framework for Face and Iris Biometrics. *Sensors* **2018**, *10*, 122. [CrossRef]
17. Savas, B.K.; Ilkin, S.; Becerikli, Y. The realization of face detection and fullness detection in medium by using Haar Cascade Classifiers. In Proceedings of the 2016 24th Signal Processing and Communication Application Conference (SIU), Zonguldak, Turkey, 16–19 May 2016.
18. Nguyen, D.; Pham, T.D.; Baek, N.R.; Park, K.R. Combining Deep and Handcrafted Image Features for Presentation Attack Detection in Face Recognition Systems Using Visible-Light Camera Sensors. *Sensors* **2018**, *18*, 699. [CrossRef] [PubMed]
19. Cho, S.; Baek, N.; Kim, M.; Koo, J.; Kim, J.; Park, K. Face Detection in Nighttime Images Using Visible-Light Camera Sensors with Two-Step Faster Region-Based Convolutional Neural Network. *Sensors* **2018**, *18*, 2995. [CrossRef] [PubMed]
20. Xu, Y.; Liang, F.; Zhang, G.; Xu, H. Image Intelligent Detection Based on the Gabor Wavelet and the Neural Netowork. *Symmetry* **2016**, *8*, 130. [CrossRef]
21. Ma, C.; Trung, N.; Uchiyama, H.; Nagahara, H.; Shimada, A.; Taniguchi, R. Adapting Local Features for Face Detection in Thermal Image. *Sensors* **2017**, *17*, 2741. [CrossRef] [PubMed]

22. Da Silva, J.; Caldeira, J.; Ferreira, D. Face Recognition Based on Higher-Order Statistics. *IEEE Latin Am. Trans.* **2018**, *16*, 1508–1515. [CrossRef]
23. Li, J.; Zhao, F.; Feng, J.; Roy, S.; Yan, S.; Sim, T. Landmark Free Face Attribute Prediction. *IEEE Trans. Image Process.* **2018**, *27*, 4651–4662. [CrossRef] [PubMed]
24. Shen, J.; Yan, G.L.W.; Tao, W.; Xu, G.; Diao, D.; Green, P. Nighttime Driving Safety Improvement via Image Enhancement for Driver Face Detection. *IEEE Access* **2018**, *6*, 45625–45634. [CrossRef]
25. Li, C.; Tai, C. A Smart Spoofing Face Detector by Display Feature Analysis. *Sensors* **2016**, *16*, 1136. [CrossRef] [PubMed]
26. Saeed, A.; Al-Hamadi, A.; Ghoneim, A. Head Pose Estimation on Top of Haar-Like Face Detection: A Study using Kinect Sensor. *Sensors* **2015**, *15*, 20945–20966. [CrossRef] [PubMed]
27. Yang, H.; Wang, X.A. Cascade classifier for face detection. *J. Algorithms Comput. Technol.* **2016**, *10*, 187–197. [CrossRef]
28. Yuan, C.; Xu, W. Multi-object events recognition from video sequences using an extended finite state machine. In Proceedings of the 4th International Congress on Image and Signal Processing, Shanghai, China, 15–17 October 2011.
29. Venkatesan, R.; Raja, P.D.A.; Ganesh, A.B. Unsupervised Learning Based Video Surveillance System Established with Networked Cameras. In *Advancs in Signal Processing and Intelligent Recognition Systems*; Springer International Publishing: Cham, Switzerland, 2016; pp. 603–614. ISBN 978-3-319-28656-3.
30. Kankanhalli, M. Multimedia Surveillance and Monitoring. In Proceedings of the 2006 IEEE International Conference on Video and Signal Based Surveillance, Sydney, Australia, 22–24 November 2006.
31. Yang, B.; Cao, J.; Ni, R.; Zhang, Y. Facial Expression Recognition Using Weighted Mixture Deep Neural Network Based on Double-Channel Facial Images. *IEEE Access* **2018**, *6*, 4630–4640. [CrossRef]
32. González-Briones, A.; Villarrubia, G.; de Paz, J.F.; Corchado, J.M. A multi-agent system for the classification of gender and age from images. *Comput. Vis. Image Underst.* **2018**. [CrossRef]
33. Ding, C.; Tao, D. Pose-invariant face recognition with homography-based normalization. *J. Pat. Recognit.* **2017**, *66*, 144–152. [CrossRef]
34. Zhou, W.; Dai, L.; Zou, Y.; Zeng, X.; Han, J. A High Speed Reconfigurable Face Detection Architecture Based on AdaBoost Cascade Algorithm. *IEICE Trans. Inf. Syst.* **2012**, *E95-D*, 383–391. [CrossRef]
35. Perera, P.; Patel, V.M. Efficient and Low Latency Detection of Intruders in Mobile Active Authentication. *IEEE Trans. Inf. Forensics Secur.* **2018**, *13*, 1392–1405. [CrossRef]
36. Park, J.-G. Face Detection Algorithm Using the Skin Color and Region of Interest. *J. Korean Inst. Inf. Technol.* **2015**, *13*, 53. [CrossRef]
37. Neugebauer, J.; Kramer, O.; Sonnenschein, M. Improving Cascade Classifier Precision by Instance Selection and Outlier Generation. In Proceedings of the 8th International Conference on Agents and Artificial Intelligence (ICAART2016), Rome, Italy, 24–26 February 2016.
38. Wang, K.; Yu, T.; Meng, Q.Y.; Wang, G.K.; Li, S.P.; Liu, S.H. Edge Detection from High Resolution Remote Sensing Images using Two-Dimensional log Gabor Filter in Frequency Domain. *IOP Conf. Ser. Earth Environ. Sci.* **2014**, *17*, 012191. [CrossRef]
39. Chiu, W.-Y.; Tsai, D.-M. Moving/motionless foreground object detection using fast statistical background updating. *Imaging Sci. J.* **2013**, *61*, 252–267. [CrossRef]
40. Shone, N.; Ngoc, T.N.; Phai, V.D.; Shi, Q. A Deep Learning Approach to Network Intrusion Detection. *IEEE Trans. Emerg. Top. Comput. Intell.* **2018**, *2*, 41–50. [CrossRef]