Patrolling Robot with Facial Detection

Manisha Dhage* Mitrajeet Golsangi§, Vivek Ghuge†,
Divija Godse‡, Pravin Harne¶,
dept. of Computer Science
Vishwakarma Institute of Technology
Pune, India
Email : *manisha.dhage@vit.edu, §mitrajeetgolsangi@gmail.com, †vivekghuge2002@gmail.com,
‡divijagodse@gmail.com, ¶sunnyharne008@gmail.com,

***Abstract*—Security of any facility and the people inside it is a huge responsibility of the management of that facility. This, even though it can be achieved using humans, it leaves them at risk and introduces a factor of human error. Thus, The proposed system aims to improve the surveillance and security of a compound, without putting any human lives at risk and naturally removing the human error factor from the process. The paper discusses the implementation of such a system and all the work related to it. The proposed system is a fully functional robot which can follow the direction of a sound. It also has an on board camera which sends live footage of everything to a server. This feed is then processed for detection of a face and it is displayed to the user.**

***Index Terms*—security, IoT, facial detection, patrolling, arduino,**

I. Introduction

For billions of years mankind has faced threats which were potentially life threatening and has found ingenious ways to overcome them. Even so, one of the most pressing concerns in society today still remains the safety of its citizens. The threats in today's modern day world are not as simple as a lion attacking a human like the good old days, but have developed as the civilization has evolved. Sadly the measures to counteract these threats are not improving at the same speed. As it is rightly said by John Maxwell "Change is inevitable, but growth is optional". According to a research conducted in [1] the acid attacks on women in south Asian countries are still prevailing and cause a great deal of damage to the women. It was not until 2019 that Bangladesh could reduce the frequency of these attacks by a meere 20-30Another survey says that 38.7All the above is evidence that even in the 21st century the crimes are at an all time high and don't seem to be going down. Even though we have various measures to prevent or stop such threats all of them involve human lives being put in danger. This system has been put as a last resort and not as a foolproof solution. The bypass to this system has become so common that even in crime thrillers it is a common practice to introduce at least one scene in which the convict can slip through the "secure" facility when the guards are changing shifts. Thus, in order to reduce the risk of any further human being and remove the possibility of any human error the idea of the proposed system was put forth by the team. As this is a well researched field the team found lots of research papers which helped improve the overall concept and make the world a little safer.

II. Literature Review

One of the most important needs of modern times is security. The rise in population is directly proportional to the need for proper security. A night patrolling visionary robot would help achieve certainty, especially during the night. All the research papers have directed to the fact that the basic must-haves for the night patrolling systems are the logic for sound sensing, moving towards the area of target and back to the original location, image capturing and processing. The night vision patrolling robot has the primary objective to ensure the safety of the people without putting any human life at risk. There are certain essential features needed in the robot, such as a night vision, obstacle detection, and motor driver circuits for controlling them [2]. All of these are controlled using a microcontroller board such as an Arduino UNO or raspberry pi. In addition, a wireless IR transceiver would be helpful in the navigation of the robot [3]. According to [2] the proposed system has the best advantage over disadvantage ratio if it uses an IR Camera with a wireless controller and some form of an obstacle detection system. Furthermore, optional features such as a GPS Module, GSM radios [3], sound sensor [2] will greatly help increase the productivity of the project. Journals have shown that a sound sensor and a smart camera are primary necessities for the system. Along with this according to the paper by N.Hemavathi the robot is built with a DC motor and transistors. The movement of the robot is handled by multiple logics operated on the transistors. These transistors allow the motion of the DC motor in the desired direction [4]. Bluetooth technology has also been used in the process for serial communication with the robot [5]. In another paper, the authors have mentioned the use of a special GPS for location tracking. Object detection algorithms have been designed for the robot to understand the route [6]. Technology is advancing at a rapid pace nowadays. These shifts are also visible in the robotics industry. As a result, the system suggested in the given papers reflects the current state of the field. The main objective of this is to offer ladies safety at night. 2 Using a credit-card-sized Raspberry Pi and an open CV, the author of this paper, created a night patrolling robot (computer vision). [7] To detect the face, the suggested

system employs a Raspberry Pi camera. Anomaly detection is done using deep learning technologies [8]. As a result, the image is recorded using the pi camera and sent to the raspberry pi for face and human detection using OpenCV. The accuracy of this approach is around 83 percent [7]. Infrared (IR) and sound sensors are employed by the authors of this paper. According to this project, a night vision camera and a pre-programmed framework are used to monitor the entire terrain. When a sound is heard, the robot will travel along a predetermined path to the location, take a photo of the area, and transmit it to the user over IoT. A pre-programmed dazzling path for night vision seeing is behind this project. [9] The proposed system focuses on the security of the place where it is implemented. It is specially designed to carry out the security assessment function at the night time when it will be dark all over the place. It is equipped with a night vision camera which will be used to capture the pictures of the spot if any suspicious activity takes place. It moves on a predestined path, and it moves in the direction of sound. It is also equipped with human face recognition technology. This system uses the LAN protocol of IOT, and it sends the recorded images to the responsible authority so that they can take any actions if necessary. It system uses a microcontroller, night vision cameras, sound sensors, Infrared sensor and Motor drivers. [10] This system is built to ensure women safety in remote regions and public places. IOT gecko is used for receiving transmitted snapshots and displaying them to people with alert sounds. This system consists of a combination of 2 HD cameras to monitor the environment sharply and closely. The hardware used in this system are HD infrared camera with night vision, Sound sensor, DC motor, IR Sensor, Ultrasonic sensor, LCD display, and a motor driver. [11] Facial recognition is an important aspect in terms of the proposed system. This serves a greater importance in order to assess the threat and act accordingly. [12] Facial recognition is a subset of the highly expanding computer vision field. This field has been overly dominated by the areas of Machine Learning and Deep Learning from the very beginning [13]. Taking into consideration the scope of lighting in a night patrolling robot it is safe to assume that the best approach would be to perform IR recognition [12]. Thus according to [12] the best possible way of training the TIR based facial recognition is to use a Deep CNN Classifier with substantial training data. The losses given by equation 1

$$\sum_{i}^{N}[\| \int (x_i^a) - \int (x_i^p)\|_2^2 - \| \int (x_i^a) - \int (x_i^n)\|_2^2 + \alpha] \quad (1)$$

Are the minimum and give the best possible output with a method called triple loss learning. [13] There are three main steps for facial recognition and tracking in real time [15]. The First Step is Detection of Faces to track them, the next being recognition of facial features, which is determined using equation 2

$$\begin{aligned}\Omega_{[w,h]}^{m} = \big([\tfrac{m}{w}] + [\tfrac{m-1}{w}] + \ldots + [\tfrac{w+1}{w}] + 1\big) \cdot \\ \big([\tfrac{m}{h}] + [\tfrac{m-1}{h}] + \ldots + [\tfrac{h+1}{h} + 1]\big)\end{aligned} \quad (2)$$

The last step is to begin tracking the face in real time. Object detection bases the image processing. The image is seen in a digital form of the same. Process of Image processing, object detection majorly involves object recognition, object localisation, image classification. According to a survey for different processes for image processing, multiple techniques have been identified which present different accuracies for the same. The survey paper implies that numerous techniques have separate sets of specifications. For example, Deep CNN shows 85% sensitivity especially for medical cases, face recognition vendor test (FRVT) makes identifying a male easier than female. According to this paper, Deep neural networks and CNN which are AI based techniques are recommended for object detection [14]. Furthermore, another paper by Sandeep Kumar, has emphasized on the kind of image processed. It mentions the use of Convoluted neural networks, for static images with static background. After preprocessing, and feature extraction, single neural networks are integrated for image recognition and processing [1] . For each stage of a face recognition system, the author of this article offers a number of cutting-edge models. The authors propose a hybrid model incorporating AdaBoost and Artificial Neural Networks to carry out the process of face detection effectively (ABANN). The final step involves aligning the labelled faces discovered by ABANN using the Active Shape Model and Multi-Layer Perceptron. In this alignment stage, the author suggests a brand-new Multi-Layer Perceptron-based 2D local texture model. The classifier in the model improves the reliability and accuracy of local searching on faces with complex expressions and ambiguous contours. The authors of the study suggest combining the geometric feature-based method and the Independent Component Analysis method to increase the effectiveness of the feature extraction stage. [15]

III. Related Work

A. Haar Cascade Classifier

One of the most popular study topics is the facial detection system. This section discusses the work defined by prior scholars on facial detection. Face detection algorithms and approaches have been the subject of numerous investigations. For the algorithm, Harleen Kaur and Arisha Mirza used OpenCV and NumPy, and they divided facial recognition into two parts. The first step in this technique is the classification task, which takes any random image or media as input and outputs a binary value of yes or no, indicating whether or not there are any faces in the image. The second phase is a face localization job, which takes an image as input and outputs the location of any face within that picture in a bounding box with width and height parameters [16]. Di Lu and Limin Yan describe that In the field of face detection, the OpenCV approach is widely used. It begins by extracting the face from a big sample set and then extracting the feature photos. The face detector is Haar characteristics in the image, followed by the AdaBoost algorithm. The algorithm detects faces. can easily adapt to challenging conditions, su 3 as insufficient illumination and backdrop blur, resulting in a significant improvement in performance [17]. The author

of the last paper describes various algorithms which are available for face detection. The author of this paper says that their paper is for anyone interested in learning about the various types of face detection algorithms. The following is how the rest of the paper is structured: first, they go over the feature-based techniques in more detail. After that the image-based techniques are discussed. Also., the face detection techniques are compared in a thorough manner [18]. An Object Detection Algorithm called Haar Cascade can be used to identify faces in still photos or moving films. In their research work, Viola and Jones suggested edge or line detecting features [19]. Features are one of the most significant aspects of the haar cascade approach. This trait is depicted in Fig. 1. These elements on the image make it simple to locate the image's boundaries or lines, as well as locations where the brightness of the pixels abruptly shift.

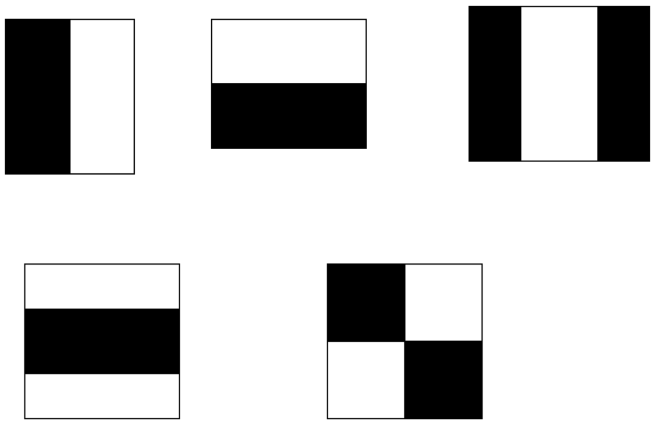

Fig. 1. A sample of Haar feature [19]

The haar feature makes pixels with values of 1 and 0 black and whiter, respectively. Each of these is in charge of identifying a certain feature in the image. any structure in the image with an abrupt shift in intensity, such as an edge, a line, or another structure. The goal is to calculate the combined number of image pixels in the haar feature's darker and brighter regions, respectively. and then identify the differences. The haar value will be closer to 1 if an edge separates the black pixels on the right from the white pixels on the left. In other words, if the haar value is closer to 1, we claim that an edge has been found. There is no edge because the haar value in the aforementioned example is quite far from 1. The haar feature needs to go through the entire picture in order to locate an edge anyplace in it.

B. MediaPipe by Google

Another popular framework for facial detection is MediaPipe developed by Google. It has many features like object detection, facial detection, posture recognition, facial recognition and more. Fig.2 shows the internal workings of the mediapipe framework [20]

MediaPipe uses a unique architecture which allows the user to incrementally prototype a pipeline. Here each component is a calculator. Graphs are specified in the Graph object using the

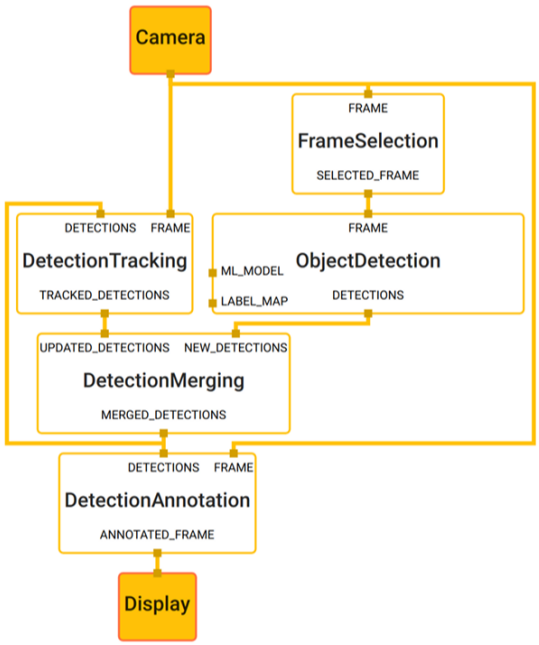

Fig. 2. Media Pipe Internal Workings

GraphConfig Protocol. Data streams connect the Graphs and Calculators where each stream is a time-series of data packets [20]

IV. Methodology

A. Hardware

1) Arduino UNO: Arduino UNO is a microcontroller board based on ATmega328p. The board has a set of input and output digital pins as well as analog pins. The board can be interfaced with several other expansion boards. The Arduino board presents a facility for communicating with computers.

2) DC motors: DC motors belong to a class of rotatory motors, which convert DC current into Mechanical energy. Using this principle, the DC motors would act as an agency for mobility.

3) Gyroscope (GY521): This is a breakout board which consists of a 3 axis gyroscope, 3 axis accelerometer, and a temperature sensor. Gyroscope helps in calculating the orientation and the angular velocity. The digital motion processor can be used to process complex algorithms on board.

4) Ultrasonic sensors: Ultrasonic sensors are used to measure the distance of any object in the vicinity of the sensor. It has a wide range and uses the emission and detection of ultrasonic waves as the principle. Primarily ultrasound sensors are used for the purpose of object detection in the proposed system.

5) ArduCam mini: ArduCam mini is a high definition 2MP SPI camera, which reduces the complexity of camera control interfaces. It integrates 2MP CMOS image sensor OV2640, and provides miniature size, as well as the easy to use hardware interface and open source code library.

B. Software

The software part of the proposed methodology is further subdivided into 2 parts the first being the Flutter application and the second being the facial detection. The Flutter application serves as a Frontend / User Interface for the system, it notifies the user about the status of the robot using the live feed of the onboard camera from the actual robot. The workflow of this application is explained in Fig. 3 Now

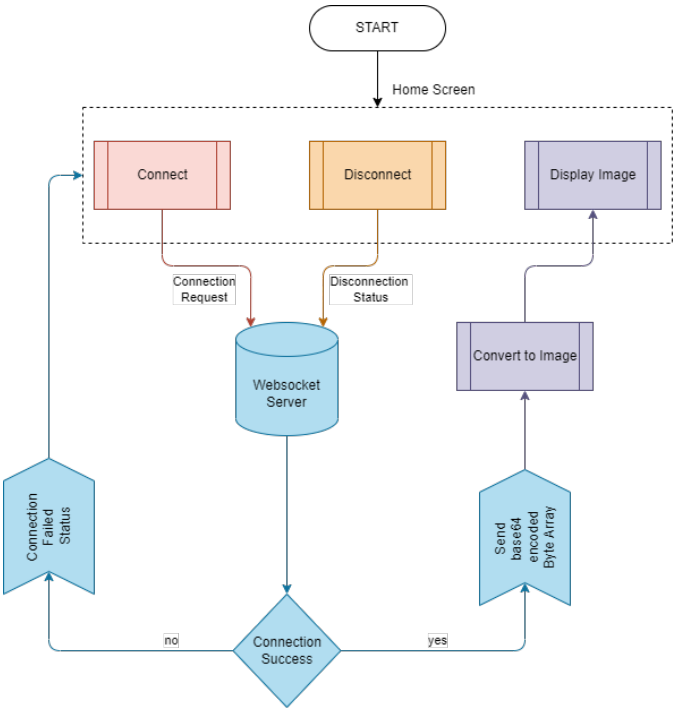

Fig. 3. Application workflow

in Fig. 3 as you can see the Flutter application only makes a connection request to the websocket. This is because as mentioned earlier the websocket protocol can directly dump data on the client without it specifically requesting for any data. The websocket server serves a simple purpose in this application. It first reads the data from the camera (which is the live video feed), frame by frame and then it performs facial detection tasks on that frame. The program then writes the confidence percentage on the image and then reads it as a byte array. This byte array is encoded in a base64 string and packaged and sent from the server to the client side. The Flutter client then reads this string and decodes it which in turn leaves it with a byte array, it then renders that byte array onto the application thus giving the user a face detected live video feed.

As illustrated in Fig. 4, the camera should be started first; if there is a problem with the camera opening, the code will be terminated. Then it will read the frame/image identified by the camera; Haar cascade is a grayscale image-only algorithm, hence the acquired image must first be transformed to grayscale before the algorithm can be applied. The frame is now ready for the haar cascade algorithm to run. The algorithm will scan the facial feature after passing through the haar cascade classifier to determine the face in that frame or image.

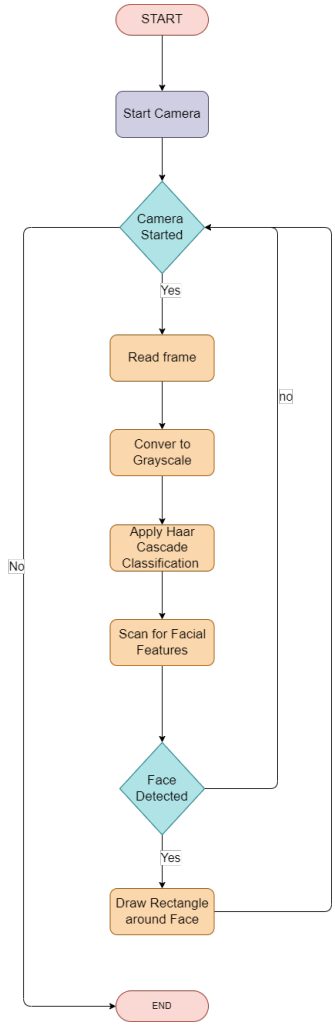

Fig. 4. Flowchart for implemented model

If a face is detected, a rectangle will be drawn around it. Otherwise, the code will be terminated.

V. Results and Discussion

The patrolling robot is tested in the different conditions. The sound sensors mounted at the different positions give us the sound decibel. Conditions of the movement provided satisfy all the directions from which the sound is perceived. The bot functions in the supportive way where the robot also performs phases like obstacle detection and image processing. It can be seen that certain logics are implemented efficiently of the different cases of rotations of the bot. The limitation of this is seen to be the size and scalability. As arduino Uno has been used, the a limited amount of features can be physically implemented. The sound sensors do not have the efficient space to be mounted, hence there are sound reception discrepancies owing owing some margin of error in the bot working. However the results obtained from the created bot have been efficient in studying it's activities. Following graphs have been plotted according to the values obtained.

Fig. 5 represents the relationship between distance and sound sensitivity. As we move away fro the sound source

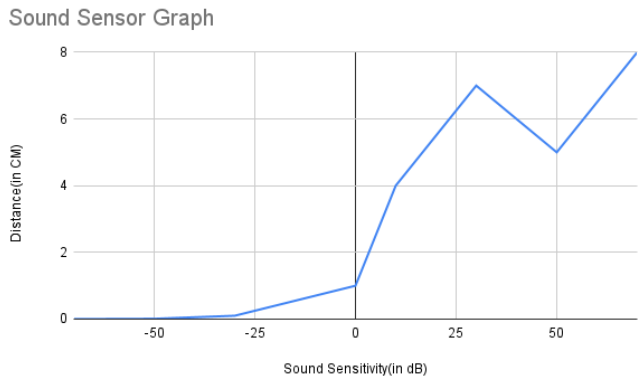

Fig. 5. Sound Sensor Readings

the intensity of the sound sensor goes on decreasing, as shown above Fig. 6 represents the relationship between the

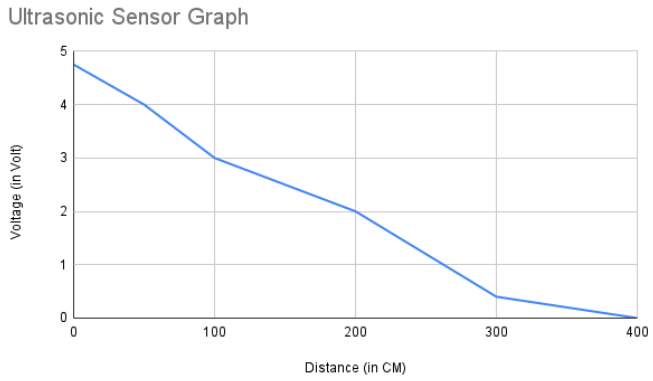

Fig. 6. Ultrasound Sensot Readings

output voltage given by the ultrasound sensor and the distance detected by the same.

The ultrasound sensor used for the obstacle detection in the path of the robot. This makes robot more self dependent and reduces human efforts of controlling it. As soon as any obstacle is detected and the robot is moving towards it, the output voltage of the ultrasonic sensor goes on increasing so when it crosses the threshold it changes its path and move right backward. In this way we can prevent wear and tear of the robot.

Fig. 7 is the measure of accuracy (determined in facial detection accuracy) of the camera vs the distance of the camera from the required object.

The arducam used in the research work is trained using facial detection model using OpenCV. The model's accuracy is tested at the different distance. So as soon as distance changes model's accuracy decreases. The above figure is the graphical representation for the same.

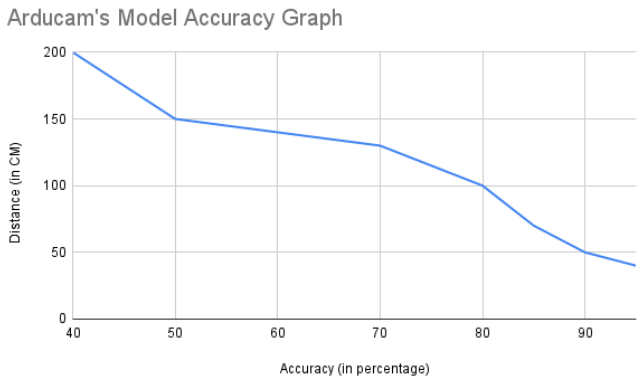

Fig. 7. ArduCAM Readings

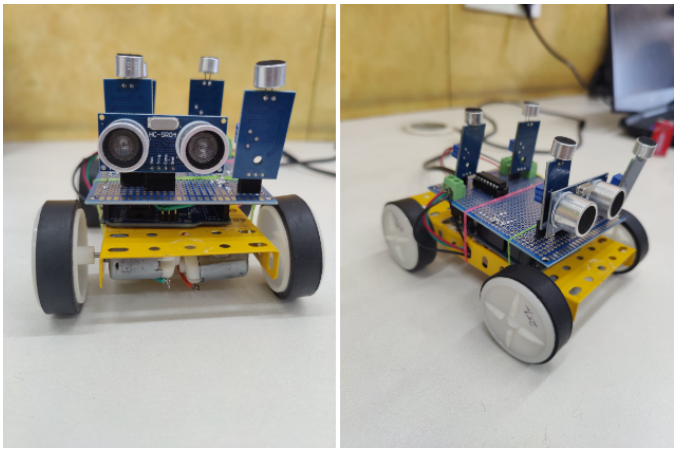

Fig. 8. Actual Implementation of the System

Fig. 8 is the actual implementation of the proposed system mounted on a small vehicle chasis, with the circuit encompassed into a PCB.

Fig. 9. Realtime Face Detection by Proposed System

The camera in the public space is able to identify faces, as seen in the Fig .9. The facial detection models range is also impressive, and it can recognise numerous faces simultaneously. This makes patrolling of the area much more convenient and easy. After a face is detected a notification is sent to the user alerting him/her about an intruder

VI. Conclusion

There are numerous conclusions that can be drawn based on the research conducted. The project concludes with the

creation of a patrolling safety robot that uses an ultrasonic sensor and a vision camera to transmit video of the entire premises.The proposed approach also addresses the flaws in the current system, which has been proposed by a number of scholars. A websocket is constructed and connected to the app interface, and with this method, a live feed captured by the robot can immediately deliver the application generated. Also, robots can patrol both during the day and at night. When the robot's sound sensors detect a sound, the motor driver attached to the robot rotates the robot towards the direction of the sound. When the ultrasonic sensor in front of the robot finds an obstruction, it moves to the right backward and detects the sound again. It then uses the vision camera to recognize the user's face and sends it to them. As a result, the robot will not get stuck at any obstruction and hence the user's safety is ensured. The robot is compact in size and can travel into areas where humans are unable to access. The Robot is hard to identify because it mixes in with its environment. The patrolling robot is a highly efficient and functional robot that saves time and effort while still completing convincing checking chores.

REFERENCES

[1] S. Hameed and B. Bhattacharya, "Scarred for life: Thoughts on legal perspectives of acid attacks in selected countries with a focus on india," *Journal of International Women's Studies*, vol. 23, no. 1, p. 12, 2022.

[2] C. Kulkarni, S. Grama, P. G. Suresh, C. Krishna, and J. Antony, "Surveillance robot using arduino microcontroller, android apis and the internet," *International Conference on Systems Informatics, Modeling and Simulation*, 2014.

[3] F. Jalled and I. Voronkov, "Object detection using image processing," 2016.

[4] S. Soo, "Object detection using haar-cascade classifier," *International Journal of Computer Science & Information Technology (IJCSIT)*, 2019.

[5] N.Hemavathy, Arun.K, Karthick.R, Srikanth.A.P, and Venkatesh.S, "Night vision patrolling robot with sound sensor using computer technology," *International Research Journal of Engineering and Technology (IRJET)*, vol. 07, 2020.

[6] J. Patoliya, H. Mehta, and H. Patel, "Arduino controlled war field spy robot using night vision wireless camera and android application," *Nirma University International Conference on Engineering (NUiCONE)*, 2015.

[7] F. Khalid, I. H. Albab, D. Roy, A. P. Asif, and K. Shikder, "Night patrolling robot," *International Conference on Robotics and Signal Processing Techniques (ICREST)*, 2021.

[8] M. Vashisht and B. Kumar, "A survey paper on object detection methods in image processing," *Institute of Electrical and Electronics Engineers*, 2020.

[9] S. Kumar, A. Balyan, and M. Chawla, "Object detection and recognition in images," *IJEDR*, vol. 05, 2017.

[10] B. N. Divya, B. Hegde, B. R. Chaithanya, M. T. M. Raju, and S. Shambhavi, "Smart motion detection surveillance rover with night patrolling for women's safety and monitoring," *International Journal of Research in Engineering, Science and Management*, vol. 4, 2021.

[11] P. Kitchagiri, S. Vaddi, S. R. Rajanala, B. Manthena, S. Sanepalli, and I. A. Pindoo, "Night vision patrolling robot for security patrolling using raspberry pi," *International Journal of Research and application*, vol. 11, 2021.

[12] V. K. Bhanase and Dr.M.D.Jaybhaye, "Face detection and tracking using image processing on raspberry pi," *International Conference on Inventive in computing applications*, 2018.

[13] T. H. Le, "Applying artificial neural networks for face recognition," *Advances in Artificial Neural Systems*, 2011.

[14] D. R. S, M. R. M. Taj, M. N. Reddy, and M. Greeshma, "Night vision patrolling navigation system for women's safety," *International Journal of Creative Research Thoughts (IJCRT)*, 2021.

[15] K. Gero, F. Noubary, I. Kawachi, C. F. Baum, R. B. Wallace, B. A. Briesacher, and D. Kim, "Associations of state-level and county-level hate crimes with individual-level cardiovascular risk factors in a prospective cohort study of middle-aged americans: the national longitudinal survey of youths 1979," *BMJ open*, vol. 12, no. 1, p. e054360, 2022.

[16] H. Kaur and A. Mirza, "Face detection using haar cascades classifier," *ICIDSSD 2020*, p. 158, 2021.

[17] D. Lu and L. Yan, "Face detection and recognition algorithm in digital image based on computer vision sensor," *Journal of Sensors*, vol. 2021, 2021.

[18] M. K. Hasan, M. Ahsan, S. Newaz, G. M. Lee *et al.*, "Human face detection techniques: A comprehensive review and future research directions," *Electronics*, vol. 10, no. 19, p. 2354, 2021.

[19] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, vol. 1. Ieee, 2001, pp. I–I.

[20] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. G. Yong, J. Lee *et al.*, "Mediapipe: A framework for building perception pipelines," *arXiv preprint arXiv:1906.08172*, 2019.