

Intruder Detection and Recognition System

Divya Gurnani
Electronics Department
Vishwakarma Institute of
Technology Pune, India

Vijay Gaikwad
Electronics Department
Vishwakarma Institute of
Technology Pune, India

Pradip V Gurnani
Executive Director
Ador Powertron Limited,
Pune, India

ABSTRACT

Earlier intrusion detection systems included manpower, but in this paper an automatic system has been proposed which detects the intruder and recognizes it also. An ultrasonic sensor has been interfaced with an Arduino which detects an intrusion. For intruder recognition, object detection classification has been performed. This method includes both computer vision and machine learning for testing and training of data. Hence this proposed method uses technology which is new and growing.

General Terms

Automatic system, Arduino, Object detection, classification.

Keywords

Intrusion detection system, Ultrasonic sensor, intruder recognition, Computer Vision, Machine Learning.

1. INTRODUCTION

A Physical Protection System [1] is used to detect the entry of an intruder into an unauthorized area. Earlier versions of this system included using a human resource and relying on operator capabilities. But now it detects intrusions through many sensitive sensors. The method proposed uses an Ultrasonic Sensor to detect the intrusion instead of a man guarding the area. Along with this an algorithm for facial recognition of the intruder has also been implemented. Facial recognition algorithms are very difficult to execute because the amount of precision required, to match faces with the records present in the database is very high. Along with that, the requirement of quality of image also needs to reach a certain threshold for the face to be recognized. This poses a problem because the images captured might not be only from frontal view.

Combining these 2 systems, of intruder detection and intruder recognition, a machine has been proposed which detects an intruder and also recognizes the intruder. This overall system is very effective for security reasons, example for defense. It can be implemented on borders where it is not safe/permissible for humans to be present. It can also be used for criminal identification.

2. STUDY OF SIMILAR CONCEPTS

The use of received signal strength indicator (RSSI) and link quality indicator (LQI) value for intrusion detection has been implemented. The RSSI was used on its own to determine the presence of an intruder compared to LQI. Later on, it is found that if the transmitter receiver separation is increased, the RSSI level will decrease and the effect of the intruder on the signal falls below the one standard deviation limit making it harder to detect and increasing the number of generated false alarms. The detection of human presence based on principal components analysis (PCA) of the signal strength by means of several radio links has been introduced. PCA compresses raw RSSI inputs that were obtained from each radio link. The

appearance of a human subject brings about RSSI variation in the environment. The detection of a slight motion, such as typing, affects the reading of RSSI value and are detectable with few wireless sensor network (WSN) nodes. It proves that RSSI is a very sensitive to the subject being experimented on and with that sensitivity, a lot of things can be detected [2]. In tracking human body motion, a different technique has been discussed. Luttwak A. was the first to combine inertial sensors (INS) data and RSSI to track human bod motion so a good accuracy robust to environment changes and drift can be achieved. Quwaider and S Biswas discussed networked proximity sensing and Hidden Markov Model (HMM) based mechanism that can be applied for stochastic identification of body posture using a wearable sensor network [3].

The Robotics Institute of Carnegie Mellon University (Collins et al., 2000) developed autonomous video surveillance and monitoring systems. They proposed the ability to continuously monitor, detect, and track people and vehicles in complex environments. When a person is classified, the skeletal form is extracted, and then the behavior is analyzed and recognized using a Markov model [4].

The intelligent video surveillance (IVS) system proposed by Prakash and Thamaraielvi (2014) identifies security threats intelligently without user intervention. A Gaussian mixture model, which is an extended Kalman filter, is used for red-green-blue (RGB) background modeling and to detect moving objects using a background model. However, the IVS only includes a function for detecting an object, but does not include a function for classifying the detected object, and does not include a function for recognizing the behavior of the object. [5]

Dimou et al. (2016) proposed a pedestrian detection framework based on a faster region-based convolutional neural network (RCNN) and a recurrent neural network (RNN) to robustly cope with motion blur and scale deformation of objects generated by pan-tilt-zoom (PTZ) and low camera quality. In order to effectively detect pedestrians in motion-blurred images, blurred image data were used in training data sets. A transformation matrix was modeled to estimate object size and pose in successive input images[6].

Saran and Sreelekha (2015) suggested a surveillance system that uses CCTV to detect and classify vehicles. They applied real-time vehicle detection and classification algorithms. Object detection is performed with a background subtraction method, and the background is modeled by using a Gaussian mixture model. In order to classify the detected vehicles, a method combining histogram of oriented gradients (HOG) and artificial neural networks (ANN) was used. However, unlike CNN, which extracts features by itself, this method requires preliminary work to extract features of a vehicle using hand-crafted methods before classification using the ANN [7]. Seung Hyun Kim, Su Chang Lim and De Yeon Kim (2017) also designed an image based intelligent IDS with a

virtual fence, active intruder detection, classification, tracking and motion recognition to reduce manpower required for standby and as operators. They used a device which combined optical and thermal cameras [8].

The Active Appearance Models (AAMs) proposed by (Cootes, Edwards, & Taylor, 1998; 2001) have been used in face modeling. After fitting the model to a face image, either the model parameters, the location of the landmarks, or the local features extracted at the landmarks are used for face recognition [9] or facial expression analysis [10]. For face recognition, (Guillemaut, Kittler, Sadeghi, & Christmas, 2006) and (Heo&Savvides, 2008) proposed used warping the face images of normalized face images into the frontal pose [11,12].(Gao, Ekenel, &Stiefelhagen, 2009) improved the performance of this technique using affine warping modified piecewise. None of these methods, however, is fully automatic and they require a manual labeling or manual initialization [13].

In a recent publication, (Gao, Zhang, Jia, Lu, & Zhang, 2015) presented a face recognition approach based on deep learning using a single training sample per person. A deep neural network is an artificial neural network with multiple hidden layers between the input and output layers. In (Gao et al., 2015), the authors propose a supervised auto-encoder to build the deep neural network by training a nonlinear feature extractor at each layer. After the layer-wise training of each building block and building a deep architecture, the output of the network is used for face recognition. One of the shortcomings of this method is the manual cropping and alignment of the face images. It is also tested only on near frontal face images [14].

3. METHODOLOGY

In the proposed methodology, an ultrasonic sensor has been used to detect the entry of the intruder. On a signal returned by the ultrasonic sensor the camera switches on and captures the image of the intruder. This image is then run through the face recognition algorithm where the intruder is recognized.

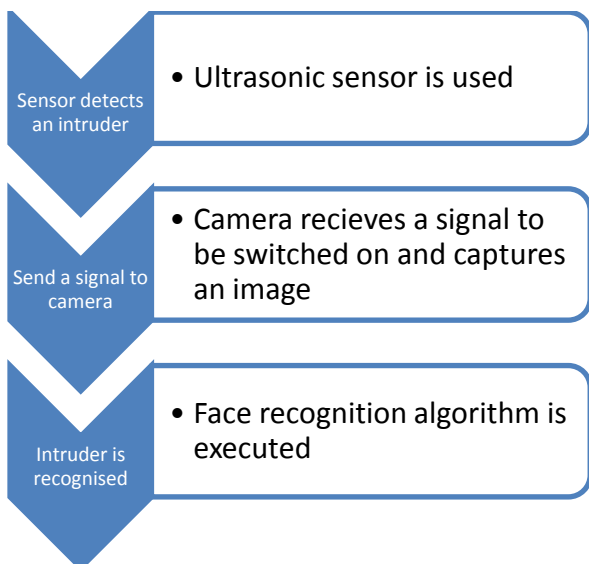


Fig 1: Data-flow diagram of proposed method

- 1) For intrusion detection, an ultrasonic sensor HC-SR04 has been interfaced with an Arduino Uno SMD R3. The sensor emits an ultrasound at 40000 Hz which travels through the air and if there is an object or obstacle on its path it will bounce back to the module. Considering the time the emission takes to travel and the speed of the

sound you calculate the distance between the sensor and the obstacle. The HC-SR04 Ultrasonic Module has 4 pins, Ground pin, VCC pin, Trig pin and Echopin . The Ground and the VCC pins of the module are connected to the Ground and the 5 volts pins on the Arduino Board respectively and the echo and Trig pins to any of the digital input/output pin on the Arduino Board. These pins need to be specified in the interfacing program It calculates the distance based on ECHO method.

$$d = (s * t)/2 \quad (1)$$

Where d is the distance, s is the speed and t is the time.

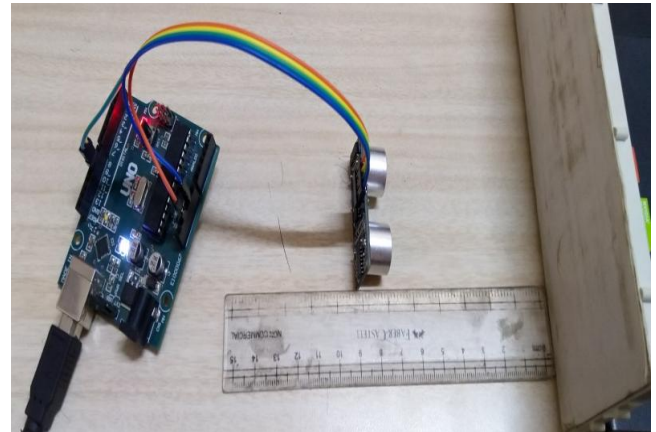


Fig 2 (a)

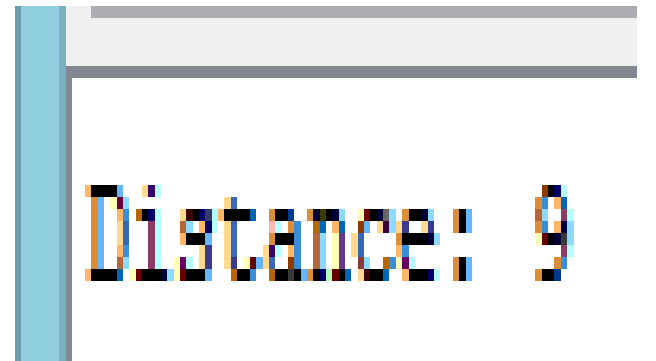


Fig 2 (b)

Fig 2: Calibration of sensor (a): Actual setup (b): Value displayed on arduino serial monitor

- 2) When fixed on the entry point, the reading of distance will have a constant value. In this case 70cm. As soon as someone enters, this value reduces and the intruder can be detected. As soon as this change in distance is detected, a signal is sent back to the Arduino via which a signal is sent to the camera. The camera then switches on and captures an image.
- 3) From the image, the face is detected. For Face Detection, there are 2 algorithms which can be followed :Haar-cascade and Local Binary Pattern(LBP) cascade. Both of them have approximately same efficiencies but LBP is faster. Hence LBP algorithm was used. The LBP labels the pixels by thresholding the 3x3 neighborhood in relation to the central pixel value

$$LBP_{(X_p, Y_p)} = \sum_{N=0}^7 s(i_n - i_p)2^n \quad (2)$$

where:(x_p, y_p) is the pixel of an image, n represents the neighboring pixel, i_n and i_p the respective gray level of

neighboring and central pixel, and $s(x)$ can be described by:

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (3)$$

4) Face Recognition mathematical model is designed as follows:

$$\varphi = \frac{1}{M} \sum_{i=1}^M r_i \quad (4)$$

where Ψ = average image, M = number of images, and r_i = image vector.

$$\Phi_i = r_i - \Psi$$

where, $i = 1, 2, 3, \dots, M$

It basically finds the face in the input image and subtracts it from the images present in the training database. For the recognition algorithm Tensorflow has been used for the Object detection classification. For this a graphics driver needs to be installed in your machine. For face recognition, a database was created consisting of 4 people, 40 images of each person. All images are used for training purpose and 70% of these images are used for testing purpose. Each image is labeled so that an xml file could be created using which the system recognizes an individual. This xml file stores the features detected in each image. From all these xml files, a single csv file is generated which contains the specific features of each individual in tabular form in an excel sheet. From this csv file, generate tfrecord files. These files are used to train the object detection classifier. Then a label map was created. This label map is created manually and consists of an id number and item name for each class, where a class is a label used while creating the database. Finally, the training is run. Lastly, the inference graph is exported.

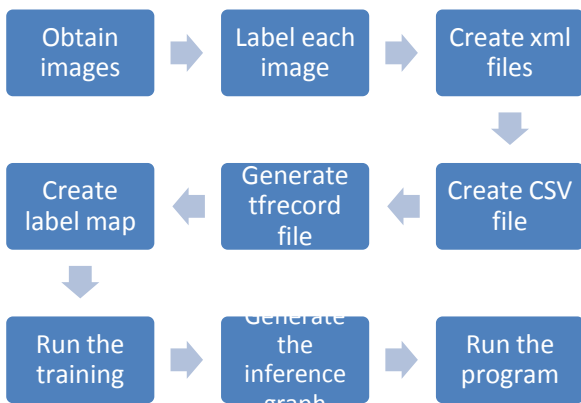


Fig 3: Data-flow diagram of Face Recognition Algorithm

4. EXPERIMENTATION AND RESULTS

4.1 Machine Properties

Table 1: Equipment properties

Machine Operating System	Windows 8.1
Processor	Intel® Core™ i5-4210U CPU @ 1.70GHz
RAM	4.00 GB
System Type	64-bit Operating System, x64-based processor
Graphics card	NVIDIA GeForce 820M , Verssion 333.11
Arduino	Arduino Uno SMD R3
Camera	OV7670 camera

Along with this, Tensorflow GPU installation is also necessary.

4.2 Face recognition testing

For face detection using LBP, the algorithm was tested on 800 images with an accuracy of 98.625%. While this training is going on, let the loss at each step go below 0.08 and stay below 0.08. This loss was also viewed on Tensorboard. Global step loss represents the sum of localization and classification losses.

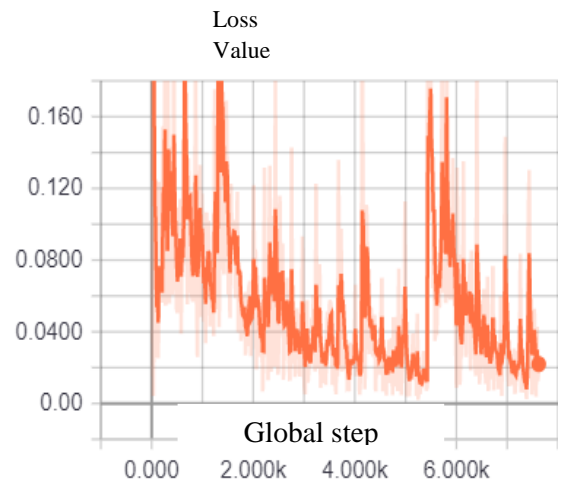


Fig 4: Loss graph obtained after a few hours of training.

4.3 Results of Ultrasonic Sensor and face recognition system

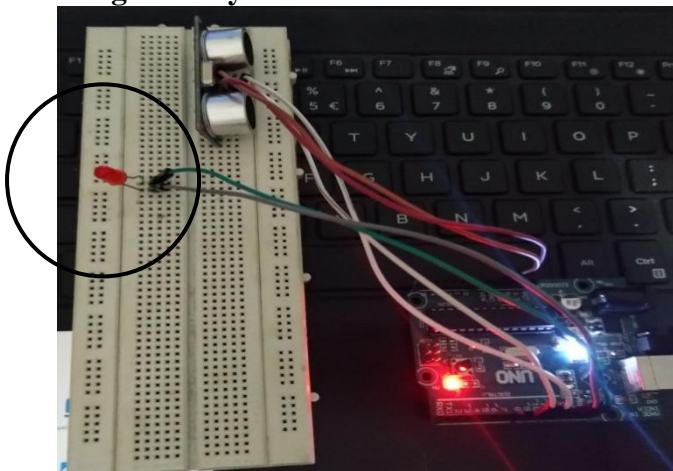


Fig 5(a)

Distance measured in cm: 3094

Fig 5(b)

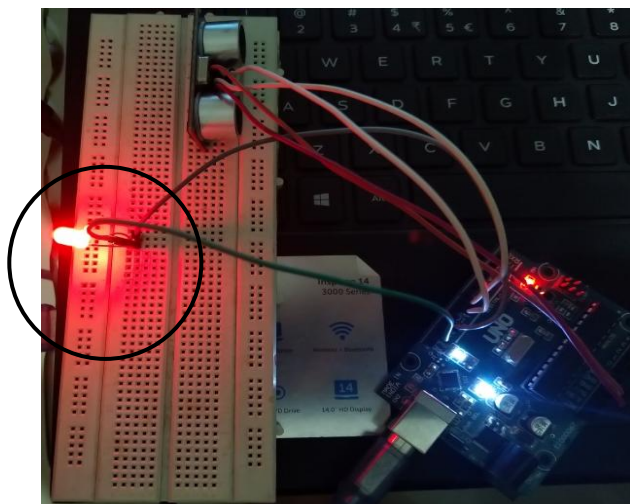


Fig 5(c)

Distance measured in cm: 36

Fig 5(d)

Fig 5: Results of Intruder Detection

- (a): Intruder absence demonstration using LED
- (b): When intruder is not detected
- (c): Intruder presence demonstration using LED
- (d): When intruder is detected

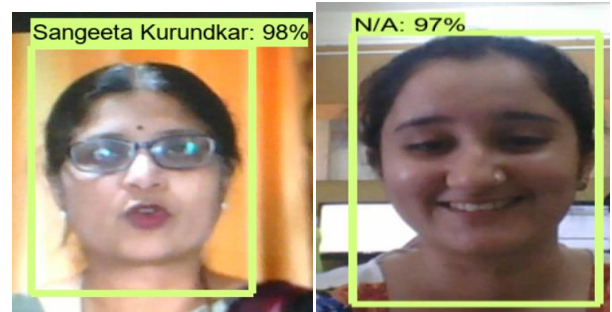
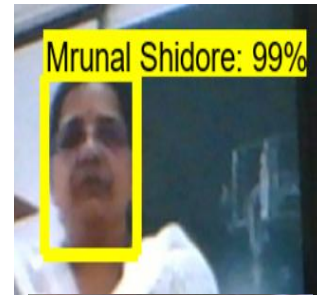


Fig 6: Results of Face Recognition.

5. CONCLUSION

In this paper, an intruder is detected and recognised. Integration of Ultrasonic sensor(model) as intruder detector is successfully implemented. Intruder recognition is done by object detection classification using Tensorflow.LBP had an accuracy of 98.625%, whereas Haar-cascade had an accuracy of 96.75%. Overall this system possesses an accuracy of 89.25%. This system has the potential of serving as a reliable and efficient security system for social and personal purposes. In the future, this system can be made more mobile so that it can send data to an individual on their mobile devices. The classifier can also be trained to distinguish between animals and humans to prevent wrongful detection by the sensor and hence prevent false alarms.

6. ACKNOWLEDGEMENTS

Sincere thanks to Dr Vijay Gaikwad and Pradip Gurnani for giving suggestions and sharing information to improve this research work.

7. REFERENCES

- [1] Mary Lynn Garcia,1 - Design and Evaluation of Physical Protection Systems,Editor(s): Mary Lynn Garcia,Design and Evaluation of Physical Protection Systems (Second Edition),Butterworth-Heinemann,2008,Pages1-11,ISBN 9780750683524,https://doi.org/10.1016/B978-0-08-055428-0.50005-1.
- [2] Hussain, Sajid & Peters, Richard & Silver, Daniel. (2008). Using received signal strength variation for surveillance in residential areas. Proceedings of SPIE - The International Society for Optical Engineering. 6973.

10.1117/12.778008.

- [3] Muhannad Quwaider and Subir Biswas. 2008. Body posture identification using hidden Markov model with a wearable sensor network. In *Proceedings of the ICST 3rd international conference on Body area networks (BodyNets '08)*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), ICST, Brussels, Belgium, Belgium, Article 19, 8 pages..
- [4] Robert Collins, Alan Lipton, Takeo Kanade, Hironobu Fujiiyoshi, David Duggins, Yanghai Tsin, David Tolliver, Nobuyoshi Enomoto and Osamu Hasegawa. A System for Video Surveillance and Monitoring. Tech. Report, CMU-RI-TR-00-12, Robotics Institute, Carnegie Mellon University, May, 2000
- [5] Prakash, U. M., Thamaraielvi, V. G., 2014. Detecting and tracking of multiple moving objects for intelligent video surveillance systems. In: *Proceedings of the 2nd International Conference on Current Trends in Engineering and Technology (ICCTET)*, July, 253-257.
- [6] Dimou, A., Medentzidou, P., Álvarez, F., et al., 2016. Multi-target detection in CCTV footage for tracking applications using deep learning techniques. In: *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, Sep, 928–932
- [7] Saran, K. B., Sreelekha, G., 2015. Traffic video surveillance: Vehicle detection and classification. In: *Proceedings of the IEEE International Conference on Control Communication & Computing India (ICCC)*, Nov, 516-521..
- [8] Seung Hyun Kim, Su Chang Lim, Do Yeon Kim, "Intelligent intrusion detection system featuring a virtual fence, active intruder detection, classification, tracking, and action recognition," Department of Computer Engineering, Sunchon National University, 255 Jungang-Ro, Sunchon, Jeonnam 57922, Republic of Korea
- [9] T. F. Cootes, G. J. Edwards and C. J. Taylor, "Active appearance models," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681-685, June, 2001. doi: 10.1109/34.927467
- [10] Ognjen Arandjelović, Abdelhakim Bendada ,Xavier Maldague. (2013). Infrared face recognition: A comprehensive review of methodologies and databases, Volume 47, Issue 9, September 2014, Pages 2807-2824
- [11] Guillemaut JY., Kittler J., Sadeghi M.T., Christmas W.J. (2006) General Pose Face Recognition Using Frontal Face Model. In: Martínez-Trinidad J.F., Carrasco Ochoa J.A., Kittler J. (eds) *Progress in Pattern Recognition, Image Analysis and Applications. CIARP 2006. Lecture Notes in Computer Science*, vol 4225. Springer, Berlin, Heidelberg
- [12] Heo J., Savvides M. (2008) Face Recognition Across Pose Using View Based Active Appearance Models (VBAAMs) on CMU Multi-PIE Dataset. In: Gasteratos A., Vincze M., Tsotsos J.K. (eds) *Computer Vision Systems. ICVS 2008. Lecture Notes in Computer Science*, vol 5008. Springer, Berlin, Heidelberg
- [13] Gao H., Ekenel H.K., Stiefelhagen R. (2009) Pose Normalization for Local Appearance-Based Face Recognition. In: Tistarelli M., Nixon M.S. (eds) *Advances in Biometrics. ICB 2009. Lecture Notes in Computer Science*, vol 5558. Springer, Berlin, Heidelberg
- [14] S. Gao, Y. Zhang, K. Jia, J. Lu and Y. Zhang, "Single Sample Face Recognition via Learning Deep Supervised Autoencoders," in *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 10, pp. 2108-2118, Oct. 2015. doi: 10.1109/TIFS.2015.2446438