An Intelligent Vision System for monitoring Security and Surveillance of ATM

Sambarta Ray
Dept. of Electronics & Communication
Heritage Institute of Technology,
Kolkata, India,
sambarta1202@outlook.com

Souvik Das
Dept. of Electronics & Communication
Heritage Institute of Technology,
Kolkata, India,
dsouvik.1993@outlook.com

Dr. Anindya Sen
Dept. of Electronics & Communication
Heritage Institute of Technology,
Kolkata, India,
anindya.sen@heritageit.edu

***Abstract*— This paper presents an automated system to increase the security and surveillance of ATM kiosks. Due to the increase of robbery in ATM kiosks, it is important to employ an automated surveillance system to protect and secure the ATM machine from threats. Currently, a camera attached with the ATM unit, records and transmits the video feed to the main server of the bank. Around the clock, this manual surveillance utilizes a lot of bandwidth for transmission. There is waste of memory and late response to emergency situation. Consequently, early detection of the situation is necessary to take preventive measures against an ongoing burglary. In this paper it is possible to detect whether a person is wearing a mask or not. The proposed system is also capable of counting the number of people present inside the ATM kiosk and generate a warning signal, thereby removes a constant human supervision, reducing the storage of unnecessary video feed and transmitting only an anomalous situation, a faster response to a threat by shutting down the ATM machine as soon the system detects the threat.**

Keywords—computer vision; image processing algorithm; surveillance; machine learning; Viola-Jones Algorithm

I. Introduction

Security is one of the primary concerns of the modern era. The biggest problem of using electronic devices and machinery is that the information can be hacked and leaked, privacy violated, theft and burglary. ATMs are global banking services which makes withdrawal of money quite convenient for the customers. Most ATMs are open 24 hours; their locations are spread all over a city/town. It might be impossible to steal money from a bank, as it is equipped with high manual security and is mostly located near populated locations of a city/town. However, ATMs due to the nature of its service, is under a greater threat of being burgled.

Since the entire purpose of an ATM service is to reduce manual interaction and provide service to the customer through automation, it is important that the security provided is also automatic and efficient. The frequent reports of burglary and theft in ATM are a growing concern for the banks. These security issues are becoming a hindrance to the expansion of ATM service and better reliability. The security system needs to be efficiently automated so that it can detect anomalous situations inside the ATM kiosk and report to the authorities or lock the machine from releasing money.

II. LITERATURE

Several research works has been done to improve the security of ATM kiosk. In last two decades security and surveillance of ATM Kiosk had became an active research area which attracts researchers from all around the globe. Initially security of ATM environment was limited to detection of individual objects like knife, helmet, pistols, daggers etc. In [1] Che-Yen-Wen et al. proposed a system which was able to detect safety helmets using modified Hough Transform [2]. In the year 2009[3], Yiping Tang et al. had developed an intelligent system which was based on Omni-Directional Vision sensor[4] and computer vision technology. In [5], Bayes Markov chains [6] were used to detect robberies at ATM machine in real time. The later was more robust and simple as it employs Markov chains to model a dynamic system.

Later research was more focused on using biometric information of a person to improve the overall security. Problems such as heavy facial occlusion, skin tone and blurred human shape were easily solved and improved over time. Several systems were modeled which employed face detection[7], facial recognition [8,9] and iris recognition technology [10] for easy detection of theft and burglary. But the major drawback in these systems is the requirement of huge database and storage facility for training. Secondly, these systems although being robust could detect the occurrence of robbery only when it is committed. Hence it lacked an early warning mechanism to notify the authorities. Also implementation of this system requires huge investments and high maintenance cost, which effects the socio-economic factors adversely. Therefore we need a system which is robust and fast, at the same time requires small database and bandwidth, can generate early warning signals as well as cheap and simple in terms of hardware and complexity

III. PROPOSED SYSTEM

This paper introduces a vision based automated surveillance system by which it is possible to monitor the ATM kiosk more efficiently. The proposed system considers two anomalous situations. First, it uses a detection algorithm to detect human faces and count the number of people present inside the ATM kiosk. Second, it can also detect whether a person is wearing mask or not, and depending on the situation it generates a warning signal to notify the authorities. The block diagram of the modeled system has been shown in Fig.1.

The detection system uses Viola-Jones algorithm[11] to detect the face and various facial aspects like the eye pair, nose and mouth. The system then counts the no. of person present inside the ATM and using the detected facial feature decide whether the person is wearing a mask. For the purpose of this paper, few selective mask models are considered as discussed in the following sections.

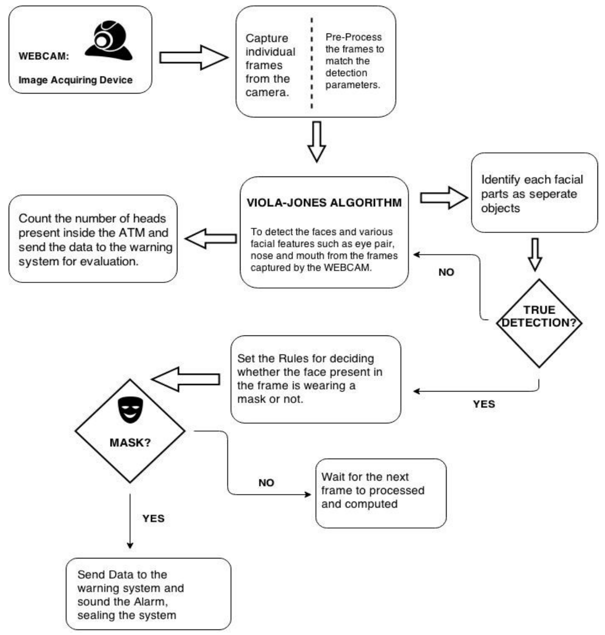

Fig. 1. Block diagram of proposed system

IV. VIOLA-JONES ALGORITHM

Viola-Jones Algorithm is one of the most robust algorithms present whose primary objective is to detect face very rapidly attaining very high true detection rate. The algorithm consists of mainly three parts:

- Integral Image
- Adaboost
- Cascading

This algorithm uses Haar Features which are a set of features given below:

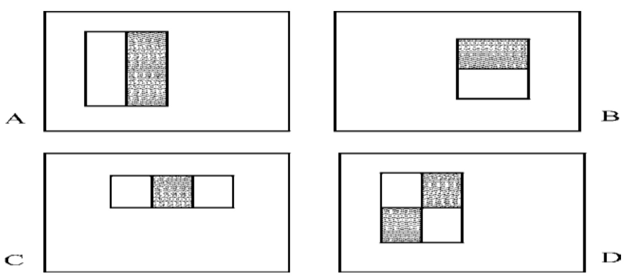

Fig. 2. Haar features used in Viola-Jones algorithm

V. METHODOLOGY

The proposed system finds its application inside the ATM kiosk. Hence the lighting condition and the background of the localized environment have been assumed to be constant while working on the project. Minor changes which are taken into account is the sunlight variation throughout the day and fluorescent tube lights (40 W & 2450 lumens; 2400-2800 K) during the night.

A. Parameters Used

The proposed system primarily uses the MATLAB function *vision.CascadeObjectDetector()* [12] to detect faces. The function uses Viola-Jones Algorithm, so certain parameters are used to increase the performance of the system and make the system more robust to external condition. The parameters used are as follow:

1) *MergeThreshold* : The value of this parameter ensures that a final detection is declared in an area where there are multiple detections around an object.
2) *MaxSize* : The value of this parameter specifies the largest object to detect.
3) *MinSize* : The value of this parameter specifies the smallest object to detect.

The sample pictures which had been used for research are all JPEG images and of dimensions *641 x 480* approximately.

B. Preprocessing Techniques[13]

Preprocessing techniques are used to eliminate the noise components and decrease the false detections.

1) *Rgb2gray :* The image is first converted to grayscale image as an preliminary treatment which eliminates the unwanted color components, consequently increasing the number of true detections.
2) *Intensity Adjustment :* The Function *imadjust()* is used for enhancing the intensity of the image. In this function a range of input intensity is mapped onto a range of output intensity.
3) *Histogram equalization :* The Function *histeq()* is used for histogram equalization which enhances the contrast of the image leading to a better detection. In this technique, the average of the overall intensity value of the image is taken and is used to smooth the intensity of the entire image.

C. Counting the number of people

The *Vision.CascadeObjectDetector* function is used to detect face-like objects. Once the face-like objects are detected it is isolated from the background. It is then treated as an individual object by forming a *Bounding Box* around the detected object which is a face in this case. The *Bounding Box* is an Nx4 dimensional matrix, where N denotes the number of faces detected in each frame. The matrix returned is named as *bboxes=[x,y,width,height]*. Effectively the numbers of rows are calculated from the *Bounding Box* matrix which is denoted by N and the value stored in N is used to determine the total number of faces in each frame.

D. Mask Detection

1) *Representation of Various Mask Models :* The proposed system is responsible for detecting whether the person present inside the ATM is wearing a mask or not. The covering of face can be done in multiple ways. For the purpose of this project, few sample models are taken for classifying both masked and non-masked images. Fig. 3, and Fig. 4, shows the various models that are considered.

Fig. 3(a), 3(b), 3(c). Models of masked images

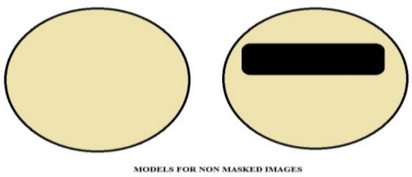

Fig. 4(a), 4(b). Models of nonmasked images

2) *State Diagram For Mask Detection:* The Mask Detection is based on the state diagram as described in Fig. 5. The Frame is collected from the webcam. Each of the final states P, Q and R is reached after the detection of the facial features from the frame is determined.

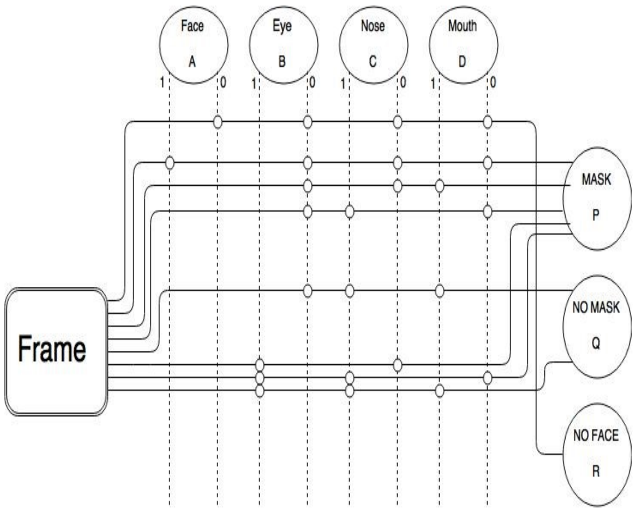

Fig. 5. State diagram for mask detection

*Rate of Detection of Mask is calculated as the ratio of the difference of positive detection and negative detection to that of positive detection. Similarly, the rate of detection of no mask is calculated in same way.

In the State Diagram, the detection is represented by 1 and the non-detection is represented by 0. The circles joining the horizontal and the vertical lines represent the fact whether the particular feature is detected or not i.e., the first line from the frame signifies that none of the facial parameter is detected in the frame, hence the final state is No face. Similarly when all the features such as the eye pair, nose and mouth is detected, the algorithm reaches the final state of No Mask.

VI . RESULTS

50 sample pictures were collected from our classmates and colleagues which were tested and ran on the system. The output of all the 50 samples were taken and recorded out of which only 12 positive results and 4 negative results are shown. The results are obtained considering an uniform environment the details of which are mentioned in Methodology section.

A. Positive Results for Mask Detection

Fig. 6, shows the samples where the proposed system has detected whether the person is wearing a mask or not correctly.
Different kinds of masks and clothes are used to cover the face but still the system is able to correctly identify the people wearing mask which verifies the robustness of the system.

B. Negative Results for Mask Detection

Fig. 7, shows some of the negative results that have occurred during evaluating the samples pictures. The negative results have occurred mainly due to false and no detection of the face objects. Although the external condition is assumed to be uniform theoretically it is not possible to attain *100%* uniformity. False detections have occurred due to following reasons:

i . Face orientation,
ii . Complexion of the person wearing the mask,
iii. Due to uneven lighting conditions.

C. Number of Head Counts

Fig. 8, illustrates the positive samples where the system can detect the number of people present inside the ATM kiosk. The system can even detect more than two heads in a frame provided the conditions are uniform.

Fig. 9, shows the positive occurrence of the three states(namely Mask, No Mask and No Face)of the modeled system. Out of the 50 samples taken, it can be seen that total positive detection of all the states is approximately 68%. From the experiments conducted, the rate of detection of mask and no mask* is found to be 75.00% and 100% respectively which depicts the efficiency as well as robustness of the modeled system as compared to any existing system.

Fig. 6. Positive results

Fig. 7. Negative results

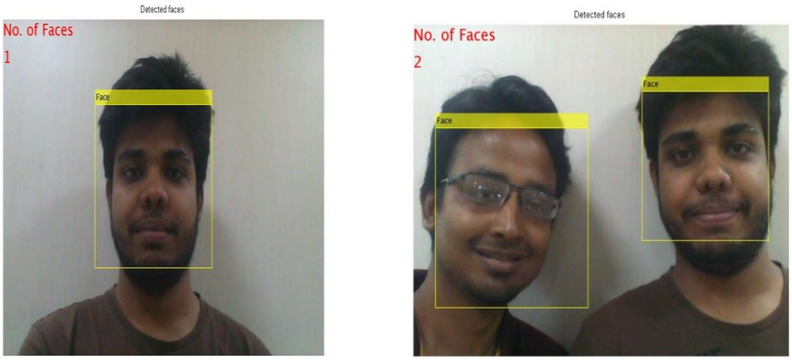

Fig. 8. Positive results illustrating the number of head counts

Fig.10 , shows the True positive, False negative, True negative and False negative values for the detection of eye-pair, nose, mouth and face for the samples collected

VII. CONCLUSION

A statistical analysis of negative detection has been accounted in this section. From the graph(Fig. 10) it can be seen that the False Negative value of face detection is very high which results in high negative detection. The reason being inefficient detection of face using Viola-Jones algorithm. Thus it can be concluded that because of the Viola-Jones algorithm's inability to detect face efficiently, the false negative value of face detection becomes significant thereby increasing the chances of getting erroneous results and limiting the overall performance of the system.

VII . FUTURE SCOPE

The modeled system can be upgraded and by taking live video feed, it can be used in real-time detection as well. By implementing advanced object detector algorithm the overall performance of the system can increased substantially.

Furthermore, in this paper certain environmental conditions have been assumed which can be further normalized by using more advanced filters and preprocessing techniques.. Also for the purpose of this paper few sample masks or face coverings have been taken, in the future, other types of masks can be included for detection which will make the system more robust in nature

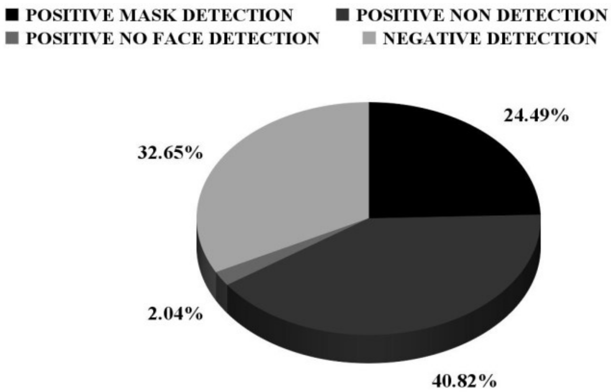

Fig. 9. Pie diagram representation of overall positive and negative detection

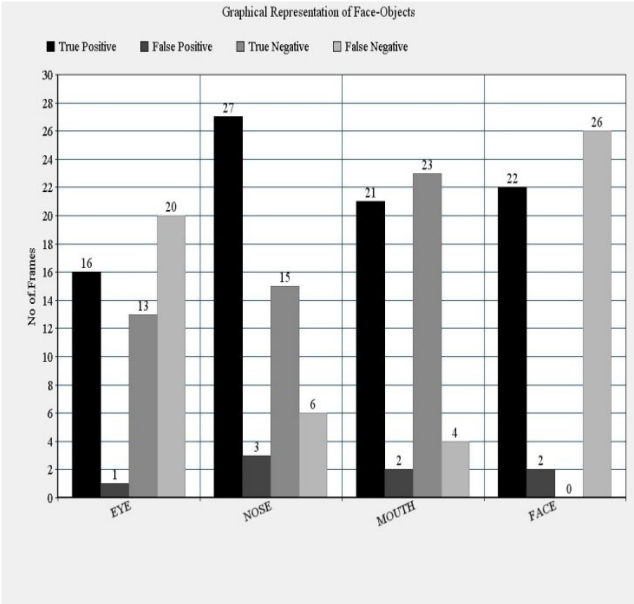

Fig. 10. Graph showing the TP, FP, TN, FN values of face object

ACKNOWLEDGEMENT

This paper would not have been completed without the constant support and guidance of Prof. Dinabandhu Bhandari. We would like to thank him for his valuable suggestions and encouragement. Finally we would also like to thank our colleagues and classmates for their precious time and faith.

REFERENCES

[1] Che-Yen Wen, Shih-Hsuan Chiu, Jiun-Jian Liaw, Chuan-Pin Lu, " The safety helmet detection for ATM's surveillance system via the modified Hough transform," Security Technology, 2003. Proceedings. IEEE 37th Annual 2003 International Carnahan Conference on, 2003, On page(s): 364-369.

[2] Jar-Ferr, Yang And Shu-Sheng Hao, "Modified Hough Transforms for object feature extraction," Journal Of Information Science And Engineering, vol. 17, 2001, pp. 133-144.

[3] Yiping Tang, Zuling He, Yaoyu Chen, Jinyi Wu, "ATM Intelligent Surveillance Based on Omni-directional Vision," Computer Science and Information Engineering, 2009 WRI World Congress on (Volume:4), 2009, On page(s): 660-664 .

[4] Yasushi Yagi, "Omnidirectional sensing and its application," IEICTE Transactions Information & System, Vol:E82-D, No.3, 1999, pp. 568-578.

[5] Arsić, D., Lyutskanov, A., Kaiser, M., Schuller, B., Rigoll, G., "Applying Bayes Markov chains for the detection of ATM related scenarios," Applications of Computer Vision (WACV), 2009 Workshop on, 2009, On page(s): 1-8.

[6] Wikimedia Foundation, Inc., [Online web]. https://en.wikipedia.org/wiki/Markov_chain/. [Last view: 22-06-2015].

[7] Zhang Jian, Song Wan-juan, " Face detection for security surveillance system," Computer Science and Education (ICCSE), 2010 5th International Conference on, 2010, On page(s): 1735-1738.

[8] Sungmin Eum, Jae Kyu Suhr, Jaihie Kim, "Face recognizability evaluation for ATM applications with exceptional occlusion handling," Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on, 2011, On page(s): 82-89.

[9] Peter, K.J., Nagarajan, G., Glory, G.G.S., Devi, V.V.S., Arguman, S., Kannan, K.S., "Improving ATM security via face recognition," Electronics

[10] Rajesh, V, Vishnupriya, S., "IBIO-A new approach for ATM banking system," Electronics and Communication Systems (ICECS), 2014 International Conference on, 2014, On page(s): 1-5.

[11] Paul Viola, Michael J.Jones, "Robust real-time face detection robust real-time face detection ," International Journal of Computer Vision 57(2), pp. 137–154, 2004.

[12] The MathWorks, Inc., [Online web]. http://in.mathworks.com/help/vision/ref/vision.cascadeobjectdetector-class.html/. [Last view: 14-06-2015].

[13] Rafael C.Gonzalez, Richard E.Woods, Steven L.Eddins, Digital Image Processing Using MATLAB, McGraw Hills, 2nd Edition.