International Journal of Engineering & Technology, 7 (2.22) (2018) 35-38

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

State model based face mask detection

Kavitha M[1]*, Mohamed Mansoor Roomi S[2], Priya K[3] , Bavithra Devi K[4]

[1]PG Student, Thiagarajar College of Engineering, Madurai, TamilNadu.
[2]Assistant Professor, Thiagarajar College of Engineering, Madurai, TamilNadu.
[3] PhD Scholar, Thiagarajar College of Engineering, Madurai, TamilNadu.
[4]PG Student,Thiagarajar College of Engineering, Madurai, TamilNadu.
**Corresponding author E-mail:mkavithamuruganece@gmail.com*

Abstract

The Automatic Teller Machine plays an important role in the modern economic society. ATM centers are located in remote central which are at high risk due to the increasing crime rate and robbery.These ATM centers assist with surveillance techniques to provide protection. Even after installing the surveillance mechanism, the robbers fool the security system by hiding their face using mask/helmet. Henceforth, an automatic mask detection algorithm is required to, alert when the ATM is at risk. In this work, the Gaussian Mixture Model (GMM) is applied for foreground detection to extract the regions of interest (ROI) i.e. Human being. Face region is acquired from the foreground region through the torso partitioning and applying Viola-Jones algorithm in this search space. Parts of the face such as Eye pair, Nose, and Mouth are extracted and a state model is developed to detect mask.

Keywords: ATM; Masked face Images; Surveillance; Viola-Jones Algorithm.

1. Introduction

Security is one of the primary concerns of the modern era. The biggest problem of using electronic devices is the information can be hacked, privacy violated, theft and burglary. The Automatic Teller Machine (ATM) plays an important role in the modern economic society. The 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 the money. Sambarta Ray et al [1] proposes a vision based automated surveillance system by which it is possible to monitor the ATM kiosk more efficiently. First, it uses a detection algorithm to detect human faces and count the number of people present inside the ATM kiosk. The problem with this algorithm is that the False Negative value of face detection is very high which results in high negative detection. 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. Y. The Viola-Jones algorithm used as a method of face detection is discussed with ref [3,4].Gabor filter is used to detect the Face as a reference of [5]. To extract areas using head and shoulder detection method even in cases of Facial occlusions and pose variations and the global and local skin color area ratios-based Occlusion verifier can effectively determine whether an area is occluded and which part is occluded even under slight pose variations. The problem, it is not robust in mask verification performance by using a head pose estimator and face wearing detector [6].

2. Database creation

Image databases are highly essential for the vision-based approaches. The database is used to analyze the features of an object and evaluate the performance of detection algorithms. Only a few datasets are available for face mask approaches. The AR face and MAFA datasets are available for face mask approaches. AR face has 4,000 images with the subjects of 126 people. MAFA is designed for the evaluation of masked face detection, which contains 35806 face annotations with a minimum size of 32×32. Only a few Facial masks images available on the Internet and do not cover a wide variety of different combinations. The datasets of AR Face and MAFA datasets do not have different combinations such as

- Nose and Mouth are visible
- Forehead and Mouth are visible
- Forehead and Nose are visible
- Eyes and Mouth are visible
- Eyes and Nose are visible
- Forehead and Eyes are visible
- Forehead, Eyes, and Nose are Visible.

Therefore, the Facial mask database is essential to detect whether a person is wearing a Mask or not. To attain these criteria the new database is created. It contains over 540 color images relevant to 5 people's faces with different combinations(shown in Fig.1).

Fig.1: Datasets of different combinations

The Images cover frontal view faces with masks (scarf). This arrangement gives the shadow-free environment for the object with 1226 Lux of light. The images are captured using a NikonD3100 camera with the auto white balance mode and resolution of 3.5MP.The videos are taken from uneven lighting conditions such as high, medium and low illumination and it is also taken as the Facial color and Mask color are more or less same is shown in Fig.2.

Fig.2: Datasets of different illuminations

3. Proposed method

This work proposes a vision based automated surveillance system. The proposed method of the system is shown in fig.3.

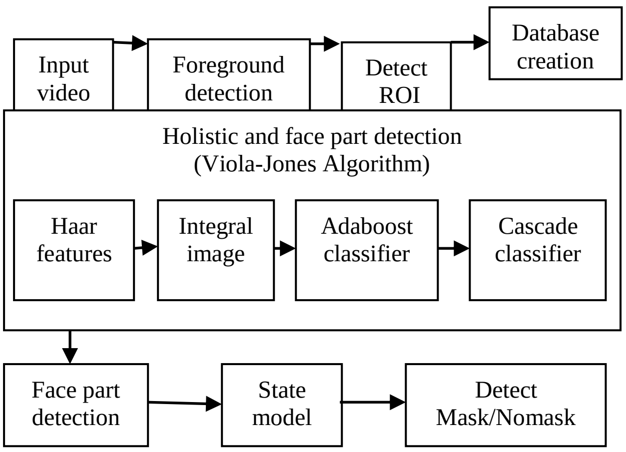

Fig.3: Proposed methodology

3.1. Gaussian mixture model

The detection of moving objects such as Foreground detection is the first process of Mask Detection in ATM scenario. The detection of Foreground objects in an image taken from a stationary camera. There are many ways to detect moving object such as a Gaussian Mixture Model (GMM), Background subtraction, etc. The Background subtraction algorithm was affected in

- Gradual Illumination Changes
- Sudden illumination changes
- Camouflage
- Shadows
- Relocation of the Background object

The Foreground detection algorithm of GMM was used to detect a Foreground region. It is one of the powerful algorithms to detect Foreground region. The moving object is extracted from the input video. The Histogram of probability density function of a large number of samples always belongs to Gaussian. Mixture means it has multiple surfaces appear in a pixel.GMM is a type of clustering algorithm used for Foreground detection. In mixture model, the clusters are classified as Hard clustering and soft clustering. In GMM the clusters are found using a technique called "Expectation Maximization". It is similar to K-means clustering. Because the probability density function of every sample in each class is calculated using Expectation-Maximization algorithm. The probability density function is calculated using mean vectors and covariance matrices. A Gaussian mixture model is classified as an Unicomponent and Multicomponent model. It further classified as univariate and multivariate Gaussian. For example, measured the height of the single person is classified as a Unicomponent-Univariate model. If measure more than one variable of a single person such as height, weight it classified as a Unicomponent-Multivariate model and More than single variable and component is called as Multicomponent-Multivariate Gaussian. Similarly, The measurement of more than one component, but a single variable is known as a Multicomponent-Univariate model. In the ATM scenario, Foreground detection mostly classified as Unicomponent model is represented as equation.1.Extract the moving object using Gaussian Mixture Models.

$$N(x:\mu,\sigma^2)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad (1)$$

Where x is a random variable, Mean μ ,Variance σ .

3.2. Region of interest

In Region of Interest, the face is an important part of Mask Detection Approach. A human body always contains the facial region in the top of the body. Divide the human body into two portions. The upper portion always contains the face region. So we extract the upper portion of the body which is called as Region of Interest of a person.

3.3. Holistic and face part detection

The Viola-Jones algorithm is a powerful algorithm for detecting faces in both holistic and by part method. The Viola-Jones algorithm consists of different techniques such as Haar Features, Integral image, Adaboost classifier, Cascade classifier. Haar features are used to extract the features of a person from the input image using convolution kernels of the matrix. It extracts 1,60,000 features and it evaluated as a window size of 24*24. An integral image is calculated from the input image. The pixel of (x,y) is calculated as the sum of all the above pixel and left of the pixel of (x,y).where x and y are the coordinates of the input image. The AdaBoost classifier is used to find the related features of a face and face parts. It eliminates the unrelated features of the extracted features of 1,60,000 features. The related features also called as a weak classifier. The several weak classifier features multiply by the weight of the future. The features are evaluated as any size of the window. The Combination of all the weak classifier, it constructs as a strong classification and Threshold detection is used to decide whether the feature is facing part or not. Finally, we extract 2500 features for each window.

$$F(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x) + \alpha_3 f_3(x) + \ldots \qquad (2)$$

F(x) is strong classifier. α_i is a weight of the feature. $f_i(x)$ is related features of face part region. Where i=1,2…..n. In cascade classifier is used to extract only the related features from the Adaboost classifier. This process is done using 4 stages. In each stage, it extracts only the related features and other unrelated features are discarded. The Architecture of Viola-Jones algorithm is shown in Fig.4.

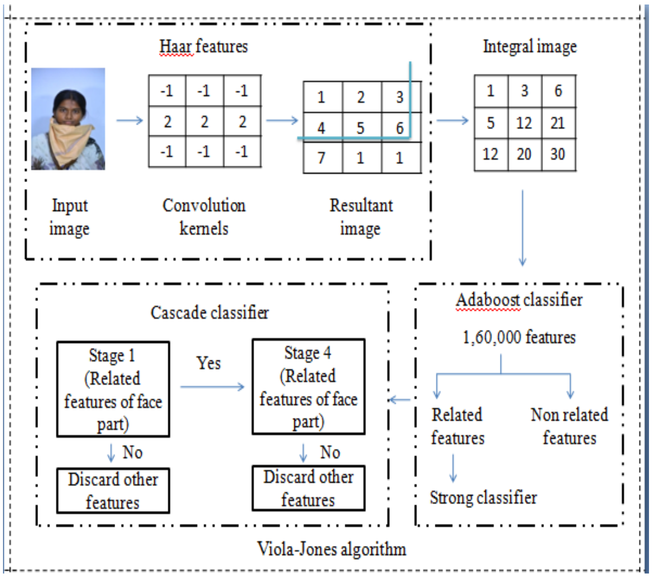

Fig.4: Architecture of Viola-Jones Algorithm

The Mask Detection is based on the state diagram of face part model as described in Fig. 4. Because the face part model is better than the Holistic Part model. Each of the final states is reached after the detection of the Facial features from the frame is determined. 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. All the features such as the Eye pair, Nose, and Mouth is detected, the algorithm reaches the state of No Mask or at least Eye pair and the Nose is visible it reaches the state of No Mask. Similarly, when All the features such as the Eye pair, Nose and Mouth is not detected, the algorithm reaches the state of the Mask. If the Facial features of Eye pair are visible and other Facial features are not visible, it also reaches the state of the Mask.

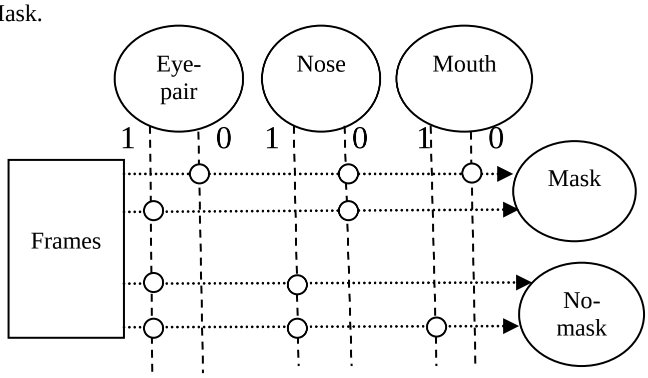

Fig.5: State Model for Mask Detection

4. Experimental Results

The videos are converted into 25 frames per second. Each video has a 593 frame and each frame have the dimension of 1080 x1920 pixels. The Foreground was extracted which is used to detect the regions of interest. The Viola-Jones algorithm will easily detect the Face when the Eyes and Nose are visible. And also this algorithm is applicable with visible forehead, Eyes, and Nose. If face some of the Facial features not visible, it's considered as a Mask. The results of face detection from holistic approach are shown in Fig.6. The results of face part detection are shown in Fig.7. The Comparison of face part model is better than the Holistic part model is shown in Table.1

Table 1. Results of Holistic and Face part Detection

	Detection of Face	Detection of Eye-Pair	Detection of Nose
Nose and Mouth are visible	-	-	80
Forehead, Mouth are visible	-	-	-
Forehead, Nose are visible	1.7	-	18.3
Facial Features are not visible	3.4	-	-
EyePair and Mouth are visible	6.7	90	-
EyePair and Nose are visible	15	98.3	88.3
Forehead, EyePair are visible	40	70	-
Forehead, EyePair, Nose is visible	90	100	88.3
All Facial Features are visible	100	100	100

Fig.6: Results of Face Detection

Fig.7: Results of Face part Detection

The Viola-Jones algorithm works well when all the Facial features are visible. Otherwise the Masked Face is detected using Face part model. The confusion matrix is shown in Table II. In confusion matrix where TP is defined as correctly classified positive examples.TN is represented as correctly classified negative examples. Similarly, FP is wrongly classified as positive examples and FN is wrongly classified negative examples.

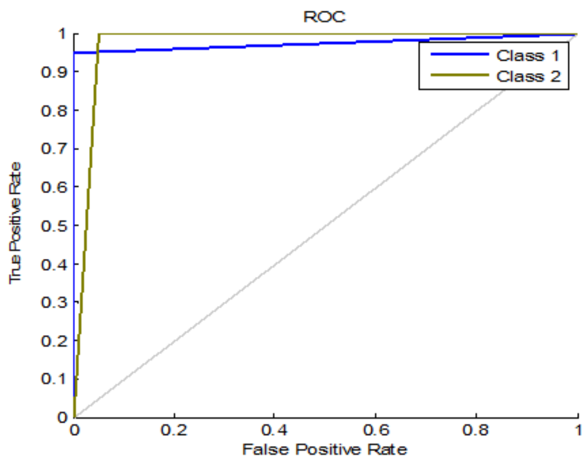

Fig.8: Results of Mask Detection

Table 2: Confusion Matrix

	Classified Positive	Classified Negative
Actual Positive	TP	FN
Actual Negative	FP	TN

$$precision = \frac{TP}{(TP + FP)} \quad (3)$$

$$sensitivity = \frac{TP}{(TP + FN)} \quad (4)$$

$$accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (5)$$

$$errorrate = 1 - accuracy \quad (6)$$

The precision and sensitivity value of Mask detection is 0.9524 and 1.Similarly, the accuracy value of Mask detection using equation (5) is 0.9750.The error Rate of Mask detection is 0.0250.From the ROC curve, class1 and class 2 means nomask and mask detection.

5. Conclusion

A threat in ATM centers due to mask wore robbers can be reduced by an alert system for mask detection. In this paper, an automatic state model based mask detection algorithm is proposed. Face is detected from the surveillance video frame by torso partitioning and application of Viola Jones algorithm. A state model is coined after facial part extraction to detect the presence of the mask. The receiver operating characteristics of the proposed mask detection algorithm has attained better performance. The future work of the paper is to overcome the false detection due to face orientation.

References

[1] Sambarta Ray, Souvik Das, Dr. Anindya Sen,"An Intelligent Vision System for monitoring Security and Surveillance of ATM", IEEE INDICON 2015

[2] Y. Xia, and F. Coenen, "Face Occlusion Detection Based on Multitask Convolution Neural Network," in Proceedings of 12th International Conference on IEEE Fuzzy Systems and Knowledge Discovery (FSKD), pp.375-379, 2015.

[3] A. Maghraby M. Abdalla O. Enany, M.Y.El Nahas "Detect and Analyze Face Parts Information using Viola- Jones and Geometric Approaches" International Journal of Computer Applications (0975 – 8887) Volume 101– No.3, September 2014.

[4] 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.

[5] C. Wen, S. Chiu, Y. Tseng, and C. Lu, "The Mask detection technology for occluded Face analysis in the surveillance system," Journal of Forensic Science, vol. 50, no. 3, pp. 1-9,May 2005.

[6] G. Kim, J. K. Suhr, H. G. Jung, and J. Kim, "Face occlusion detection by using b-spline active contour and skin color information," In Proceedings of the International Conference on Control, Automation, Robotics and Vision, Dec. 2010, pp. 627-632.