

# Efficient Framework Using Morphological Modeling for Frequent Iris Movement Investigation towards Questionable Observer Detection

**D. M. Anisuzzaman**

Department of Computer Science and Engineering, Ahsanullah University of Science and Technology,  
Dhaka, Bangladesh  
Email: rajon99@gmail.com

**A. F. M. Saifuddin Saif**

Department of Computer Science, American International University-Bangladesh, Dhaka, Bangladesh  
Email: saif@aiub.edu

Received: 29 June 2018; Accepted: 17 September 2018; Published: 08 November 2018

**Abstract**—This research presents a framework to detect a questionable observer depending on a specific activity named “frequent iris movement”. We have focused on some activities and behaviors upon which we can classify one as questionable. So this research area is not only an important part of computer vision and artificial intelligence, but also a major part of human activity recognition (HAR). We have used Haar Cascade Classifier to detect irises of both left and right eyes. Then running some morphological operation we have detected the midpoint between left and right irises; and based on some characteristics of midpoint movement we have detected a specific activity – frequent iris movement. Depending on this activity we are declaring someone as questionable observer. To validate this research we have created our own dataset with 86 videos, where 15 individuals have volunteered. We have achieved an accuracy of 90% for the first 100 frames or 3.33 seconds of each of our videos and an accuracy of 93% for the first 150 frames or 5.00 seconds of each of our videos. No work has been done yet on basis of this specific activity to detect someone as questionable and furthermore our work outperforms most of the existing work on questionable observer detection and suspicious activity recognition.

**Index Terms**—Iris detection, frequent iris movement detection, suspicious activity detection, activity recognition, questionable observer detection

## I. INTRODUCTION

Questionable observer detection in the process of guessing someone is guilty and doing/did something wrong or going to do something harmful for a specific or a massive number of people; from the behavior or from

some specific action of that individual. Not only ‘questionable observer detector’ can support respective person to help detect criminal activities, but also can help to prevent it from happening if the questionable person is detected before doing the crime. Our research is focused on questionable observer detection. A criminal activity that occurs in the large crowd is the most harmful. So our goal is to predicting someone questionable from his/her behavior or some specific activities before he/she can do something harmful. A system that can detect questionable observer intelligently; will be much more efficient than human sitting before surveillance video monitoring as it is a tiresome and boring work and an intelligent system can monitor and detect all the questionable observers in its range whether a human can scrutiny one person in particular or in some cases several persons. We can say someone questionable on basis of some specific activities: repetitive touching of face, excessive head-turning, avoids eye contact, powerful grip of a bag or hand inside the bag, excessive clock watching etc. Although a lot of work has been done for basic human activity detection (walking[2-5], [8], running [10-12, 18], sitting still [14-16, 20] etc. [22, 23, 26-28]); to the best of our knowledge none of these activities mentioned earlier for questionable observer detection have been detected.

In this research we have detected a specific activity named “frequent iris movement”, and in basis of this activity we are accusing someone as questionable. One of the activities of a questionable observer is frequent iris movement. Most of the guilty person tries to avoid their eye contact and also tries to observe their surrounding territory or situation from the edge of their eyes. So the person moves his/her iris from left to right or from bottom to top or vice versa more frequently than a normal person. On basis of these characteristics of the eye movement or more specifically iris movement we have

detected the activity “frequent iris movement”, and claim the person as questionable. Only some works have been done on suspicious activity detection [13, 29] and questionable observer detection [17] and the detection of violent crowd movement [7, 9]. But none of them considers the behavior of frequent iris movement as a measurement of their detection. As a questionable observer tries to hide his/her intention or guilt, he/she tries to avoid eye contact. In the pecking order of the society, a questionable observer tries not to make eye contact with a superior since that would be a challenge. On basis of these facts we try to differentiate questionable observers depending in this specific activity – “frequent iris movement”.

This research proposed a new framework for questionable observer detection by detecting an action named “frequent iris movement”. Section 2 contains the background study on action recognition and questionable observer detection. Section 3 contains the research methodology where we have discussed proposed frameworks. Section 4 contains the detail of experiment, result to reveal the efficiency of the proposed framework. Finally, section 5 illustrates the concluding remarks.

## II. RELATED WORKS

The field of human activity recognition considers mostly the daily activities: Lying, Sitting, Standing, Walking, Running, Watching TV etc. Among them most of the research consider dataset collected from wearable devices that take certain measurements to detect an activity. Two wearable modules were attached [4] to detect the acceleration and angular velocities and wireless sensing devices [8] are attached on different parts of the human body to transmit biophysical data to a mobile device. Matsui et al., Chen et al. And Sunkad and Zubin also used the data from wearable devices [14-16]. Cheng et al. used a dataset where five motion sensors were placed at five different parts of the body [20]. Accelerometers and gyroscopes data were used and processed for human activity recognition in different researches [23, 26-28]. All these works have some common problems: the users have to wear gadgets on the body which is highly impractical in real world [4]; some of them are not real time detection [4]; in the cases where the data was collected from smart phones of the users, each users wear their smart phones quite differently (some hold them in one hand, others put them in their pockets, or in their bags etc.), which makes sensor outputs significantly different, and thus the HAR performance is degraded [14, 23], most sensor readings may contain incorrect or noisy data [26], some similar activities like “climbing upstairs” and “climbing downstairs” are hard to differentiate by using sensor data [16] and wearable sensors require extra costs for hardware and are intrusive for users [15].

The other HAR systems use non wearable sensor based datasets or vision based dataset for activity recognition. Videos from web camera [3]; spatio-temporal features (STIPs) [5] and videos from different dataset which

includes different type of activities [10-12] are used for vision based activity recognition. Channel State Information (CSI) data developed by using the WiFi networks is used in [18]. Different dataset that were developed by the authors for their own works are also used [22]. The taken from videos or from non wearable devices also have some limitations: the line-of-sight detection, good illumination, and potential privacy leakage [18]; in some cases it is critical to obtain the accurate outline of moving object from a video and also the video background has to be known [3]; finding trajectories from videos might not be discriminate enough [11]; CSI often request specific facilities: GPS clock, RFID readers and it is also extremely sensitive to the channel variance and position changes [18] and some experiments are only realized in indoors environment and also needed expensive GPU for computation [22].

Whatever the dataset is all of them go through some approaches or algorithms to detect an activity correctly and some popular approaches and algorithms for activity recognition are: Support Vector Machine (SVM) [16], Convolutional Neural Network (CNN) [12, 14], Long-Short Term Memory (LSTM) neural network [15, 22, 26] etc. These methods and algorithms also have their own limitations: in SVM choosing a good kernel function is not easy and it needs long training time on large data sets [16]; deep learning based approaches often suffer from noise effect and limited size of data, leading to an unsatisfactory performance [15]; and for deep neural network if the number of hidden units is too small, the expression ability of model is not enough and if too big, the model complexity is too high and the generalization performance will descend severely [22].

A very little work has been done on detecting suspicious activity and questionable observer [1, 13, 17] from surveillance camera. Some work has been done to detect violent crowd behavior [7, 9]. Facial expression and human behavior detection is done in [6, 19, 21], which are also vision based approaches. These works also contains some challenges and limitations: some works needed fixed camera shoots from same angle and distance [1]; in some cases with the increasing training dataset, the time required also increased by some margin [13]; some detection is harmed by variations in pose, illumination, and facial expression throughout a single video and between different videos, that can affect face appearance and, hence, complicate questionable observer detection as well and also the video evidence may be recorded by camera phones or surveillance cameras and so the quality of the face image sequences can be very low [17] and the “object-based methods” for understanding crowd behaviors faces considerable complexity in detecting objects, tracking trajectories and recognizing activities in dense crowds where the whole process is affected by occlusions [9].

The camera also plays an important role in the questionable observer detection as the detection of questionable observer will be done from surveillance videos. Some work has been done [24, 25] to capture better video which will help us to detect questionable

observer more accurately.

From the above discussion we can state that, no questionable observer detector has been built yet depending on some specific behaviors and actions we have mentioned earlier and our work on questionable observer detection by using the activity “frequent iris movement” is a new angle to look upon this field of work.

### III. PROPOSED RESEARCH METHODOLOGY

Questionable observer detection involves several activity detection and computation of all these activities in parallel. Among these activities, one is “frequent iris movement”. When a person tries to avoid his eye contact with other people he moves his eyes frequently from left to right or from top to bottom or in vice versa with hesitation; which indicates something suspicious is going to happen and the person is involved in it. To the best of our knowledge no work had been done yet to detect this specific activity. This activity detection is done in five steps and a block diagram of this process is shown in Fig. 1.

#### A. Framework

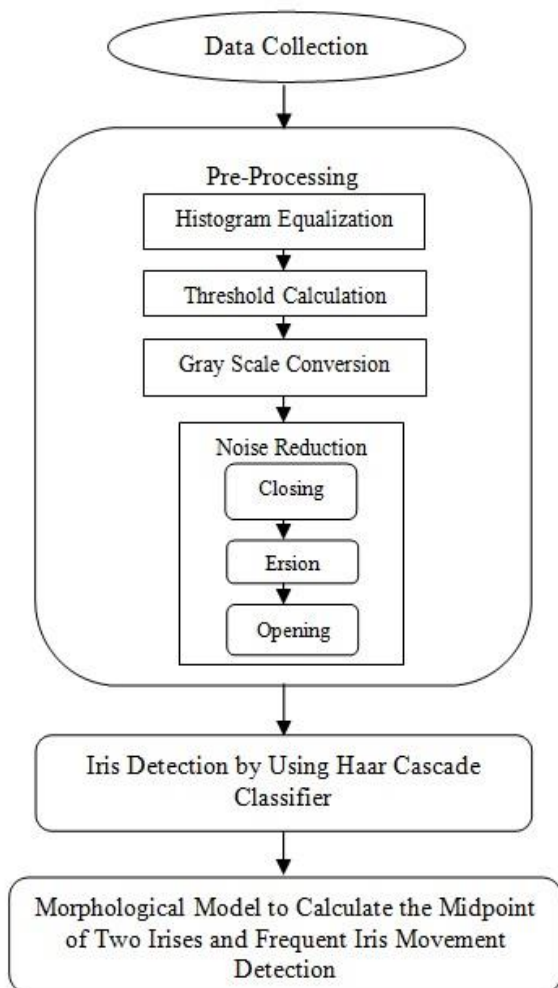


Fig. 1. A Block Diagram of “Frequent iris movement” Activity Recognition

After collecting the data from the volunteers, all data go through some processing to build our dataset. From each frame of every video we have found the locations of both irises by using haar cascade classifier. After that we have found the midpoint between these two detected irises by using some morphological operations. Finally from the frequent movement of the midpoint in consecutive irises we have taken our decision. The whole process is described elaborately in the following sections (B to F).

#### B. Data Collection

All videos are taken using a mobile handset (Samsung Galaxy Note 8) in different environments. A total of 15 volunteers have been involved in this experiment. All videos contain the frontal facial view of the volunteers, and the lighting and volunteer movement was not enforced externally. In some videos volunteers move their iris more often than regular from left to right and top to bottom and vice versa. In the remaining videos they keep their eye contact in regular. Each video has a length of 5 to 7 seconds.

#### C. Pre-processing

All videos have a 30 frame per second rate, where each frame contains a resolution of  $1920 \times 1080$ . Without changing the resolution we consider original frame rate (30 fps) for our experiment. We have drawn the histogram to find the lightness distribution of each frame. Following that histogram equalization has been run to spread frequencies for uniform intensity distribution. After that threshold calculation has been done for the grayscale conversion. Here grayscale conversion is needed because cascade classifier expects grayscale images. Then noise reduction is performed by using some morphological operations: closing, erosion and opening. Closing has been used to preserve background regions that have a similar shape to the structuring element, or that can completely contain the structuring element, while eliminating all other regions of background pixels. Opening has been used to preserve foreground regions that have a similar shape to the structuring element, or that can completely contain the structuring element, while eliminating all other regions of background pixels. For opening and closing we have used  $3 \times 3$  structuring element. Finally, these pre-processed videos have been used in the next step for haar cascade classifier on our video dataset.

#### D. Iris Detection by Using Haar Cascade Algorithm

Iris is detected by using haar cascade algorithm. Object Detection using haar feature-based cascade classifiers is an effective object detection method proposed by Paul Viola and Michael Jones [30]. It is a machine learning based approach. A lot of positive and negative images were involved in the training of the cascade function; which is then used in the detection of objects in other images. We have used the eye cascade of haar cascade classifier to detect eye from each frames. From that we have done some morphological image processing to get

the pupil and found the biggest blob and get the centroid to detect iris. The process is done for both irises.

*E. Morphological Model to Calculate the Midpoint of Two Irises and Frequent Iris Movement Detection*

After detecting the irises we have also obtained the centroides of both left and right irises. The centroides (*leftX*, *leftY*) and (*rightX*, *rightY*) is obtained from the left iris and right iris respectively by applying morphological operations on haar cascade classifier for eye detection. After that by using the centroides of both left and right irises we have detected the midpoint in each frame. The midpoint (*centerX*, *centerY*) is detected by using the formula:

$$centerX = \frac{leftX + rightX}{2} \quad (1)$$

$$centerY = \frac{leftY + rightY}{2} \quad (2)$$

The x coordinate of the midpoint is obtained by dividing the sums of x coordinate of the left iris and right iris by two. Similarly, the y coordinate of the midpoint is obtained by dividing the sums of y coordinate of the left iris and right iris by two. By using this midpoint in the next step we have made our detection.

We find the range of movement of the midpoint in all the consecutive frames. We set a threshold value after testing our work with more than 100 videos. From these 100 videos we have observed that, if a person tries to avoid his/her eye contact or moves his/her eyes frequently from right to left or vice versa the range of the x coordinate will vary a lot in comparison with a person who is not trying to avoid the eye contact. Similarly, if a person moves his/her eyes frequently from top to bottom or bottom to top the range of the y coordinate will vary a lot in comparison with a person who is not trying to avoid the eye contact.

Fig. 2 illustrates the movement of (x, y) coordinate of the midpoints in consecutive frames for an avoiding case. Fig. 3 illustrates the movement of (x, y) coordinate of the midpoints in consecutive frames for a non avoiding case.

Center Coordinates: [(605, 300), (602, 315), (604, 304), (605, 305), (605, 307), (631, 319), (603, 310), (604, 318), (632, 321), (632, 323), (607, 322), (606, 323), (605, 325), (595, 345), (612, 307), (665, 329), (622, 348), (623, 349), (616, 305), (629, 347), (659, 349), (628, 348), (618, 346), (622, 352), (622, 356), (582, 362), (585, 364), (586, 364), (586, 365), (590, 368), (603, 373), (606, 375), (602, 377), (608, 378), (586, 374), (606, 380), (602, 382), (605, 396), (578, 382), (578, 384), (576, 386), (602, 402), (577, 383), (599, 390), (576, 384), (601, 404), (591, 378), (618, 392), (591, 377), (591, 377), (627, 378), (630, 375), (618, 371), (619, 374), (586, 378), (586, 378), (588, 377), (596, 376), (611, 355), (606, 363), (604, 364), (603, 365), (603, 365), (601, 366), (601, 366), (600, 356), (599, 356), (597, 356), (595, 356), (593, 358), (590, 358), (588, 357), (588, 356), (583, 356), (607, 364), (577, 354), (576, 353), (572, 354), (571, 353), (558, 360), (568, 354), (566, 354), (591, 361), (578, 385), (578, 387), (598, 384), (582, 379), (592, 383), (548, 387), (553, 389), (553, 388), (564, 386), (554, 389), (554, 390), (555, 391), (562, 393), (563, 394), (565, 396), (566, 397), (571, 400), (580, 404), (587, 406), (563, 400), (588, 406), (586, 407), (592, 404), (592, 405), (568, 400), (571, 401), (573, 402), (575, 404), (578, 406), (581, 408)]

Fig.2. The movement of midpoints in consecutive frames in an frequent iris movement case

So if a video crosses the threshold value for either x coordinates or y coordinates, we classify the video as “frequent iris movement” otherwise it is classified as “not frequent iris movement”.

From Fig. 2 we can see that the x coordinates of the midpoint in consecutive frames varies from 548 to 665 with a rage of 117 and the y coordinates of the midpoint in consecutive frames varies from 300 to 408 with a range of 108.

Center Coordinates: [(600, 318), (601, 317), (602, 317), (604, 319), (605, 320), (605, 321), (605, 322), (606, 323), (606, 324), (608, 324), (609, 325), (609, 326), (608, 325), (608, 323), (607, 322), (607, 321), (606, 321), (604, 322), (603, 322), (601, 322), (599, 323), (596, 327), (587, 332), (575, 330), (563, 332), (562, 342), (584, 359), (562, 362), (556, 358), (561, 360), (571, 358), (571, 358), (573, 359), (573, 360), (570, 360), (570, 357), (570, 353), (571, 350), (574, 348), (575, 347), (580, 345), (577, 343), (575, 343), (574, 343), (572, 342), (567, 341), (567, 340), (562, 341), (560, 341), (560, 339), (555, 338), (552, 339), (553, 338), (552, 338), (551, 338), (549, 339), (547, 341), (543, 341), (541, 342), (540, 343), (541, 344), (539, 345), (537, 345), (542, 343), (537, 346), (537, 347), (536, 347), (540, 348), (539, 347), (540, 347), (541, 347), (542, 348), (543, 349), (546, 350), (545, 350), (545, 351), (543, 351), (543, 352), (541, 353), (542, 354), (542, 355), (542, 358), (541, 360), (540, 361), (540, 361), (540, 361), (540, 360), (540, 361), (540, 361), (540, 362), (542, 362), (541, 362), (542, 361), (547, 360), (543, 359), (544, 356), (548, 356), (547, 356), (546, 354), (549, 355), (547, 354), (548, 353), (548, 353), (550, 353), (550, 354), (548, 353), (547, 354), (547, 354), (547, 355), (547, 355), (549, 356), (548, 357), (551, 356), (549, 355), (548, 355), (549, 354), (547, 352), (546, 352), (551, 352), (548, 352), (549, 352), (551, 351), (551, 351), (553, 351), (554, 351), (553, 352), (552, 352), (553, 354), (553, 355), (551, 354), (550, 354), (548, 352), (549, 351), (549, 350), (547, 350), (546, 350), (545, 350), (543, 350), (544, 350), (542, 351), (541, 351), (544, 351), (540, 351)]

Fig.3. The movement of midpoints in consecutive frames in a non frequent iris movement case

From Fig. 3 we can see that the x coordinates of the midpoint in consecutive frames varies from 538 to 611 with a rage of 73 and the y coordinates of the midpoint in consecutive frames varies from 317 to 362 with a range of 45.

From these two cases we can clearly see a difference in the range of midpoint coordinates variation. In the non frequent cases the range is much lower than the case of frequent iris movement. After a careful inspection in 100 videos we set a threshold value for the range of both x and y coordinates of the midpoint. If in a video the range of either x or y coordinate remains lower than the threshold value we classify the video in not avoiding case; otherwise we classify the video as avoiding case.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The experiment is carried out in two phases. The first experiment is done with first 100 frames of each of our videos and the second experiment is done with 150 frames of each video. All the experiment is done with Anaconda 3 platform by including OpenCV package. To validate our classification of questionable observer depending on frequent iris movement, we have created our own dataset by using a smart phone where the illumination, pose of the volunteers and expression of the volunteers are not externally forced. Then we find the accuracy, macro-average precision, macro-average recall, micro-average precision and micro-average recall for both of our experiments. Then we compare our findings of both of our experiments according to accuracy, micro-



average and macro-average precision and recall. Finally we compare our work with the existing work in the field of questionable observer detection, violent crowd behaviour detection and suspicious activity detection and recognition. Fig. 4 reflects our experimental result and discussion section; and followed by this figure a comprehensive description of our experimental findings and discussion about our findings are given in the sections A to D.

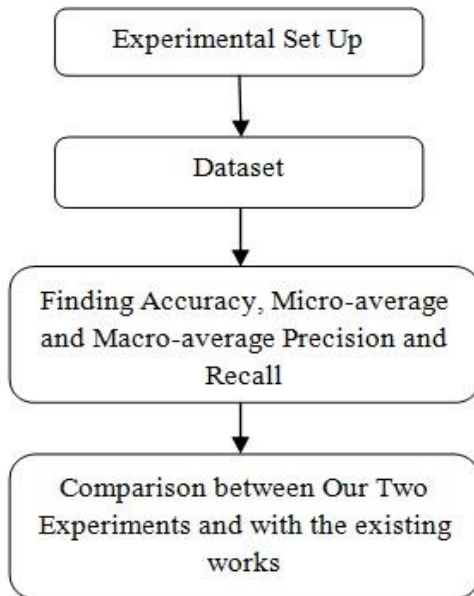


Fig.4. Four phases of experimental results and discussion

#### A. Experimental Set Up

The experiment is done with OpenCV in python platform. The detection process is done here.



Fig.5. Detection of both left and right irises

First we create eye detector with OpenCV and for this we need to load the required XML classifiers. But the XML file can only detect the right pupil. By applying some morphological operation on this XML file we create the iris detector for both left and right eyes. Then we load our input video (or frame) in gray scale mode.

Fig. 5 shows some example result of our iris detection process.

After that we find the midpoint in each frame by using equations (1) and (2). Fig. 6 shows some example of our midpoint detection process. From the consecutive frames of a video of a volunteer who is trying to avoid eye contact, we can found that the midpoint of the consecutive frames changes frequently with a larger range (Fig. 7) than a volunteer who is not trying to avoid eye contact (Fig. 8).

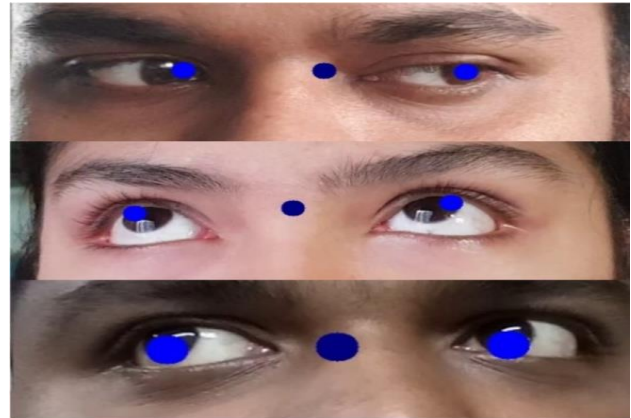


Fig.6. Detection of midpoint between two irises

#### B. Dataset

As we do not find any suitable dataset available to validate our work specifically we have to develop our own dataset. No available dataset contains such videos where the volunteers move their iris frequently from left to right or from up to down or vice versa. For this reason we have developed our own dataset named "Iris Movement Dataset". The dataset includes a total of 86 videos. In 44 videos the volunteers are frequently moving their irises and in 42 videos they are not frequently moving their irises or in other words are in their normal

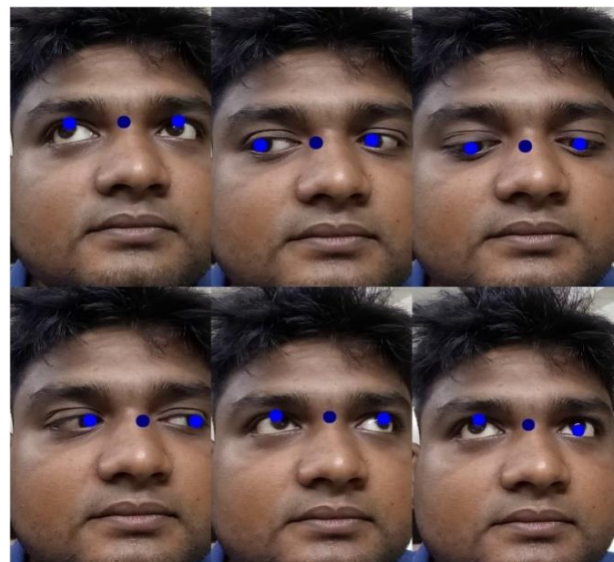


Fig.7. Movement of midpoint in consecutive frames for a person who is trying to avoid eye contact

state. Each video has a 30fps frame rate. We have done our experiment for the first 100 (3.33 second) and 150 (5 second) frames of each video.

C. Experimental Results

From the frequent movement in consecutive frames and from the range of the movement of the midpoint we detect the action “frequent iris movement”. In Fig. 9 we have shown the result of a video where the person is not trying to avoid his eye contact. Both the range of x coordinate and y coordinate is less than the threshold and our system detected the video as “Not frequent iris movement”.

Similarly, in Fig. 10 we have shown the result of a video where the person is trying to avoid his eye contact; where both the range of x coordinate and y coordinate is greater than the threshold and our system detected the video as “Avoiding”.

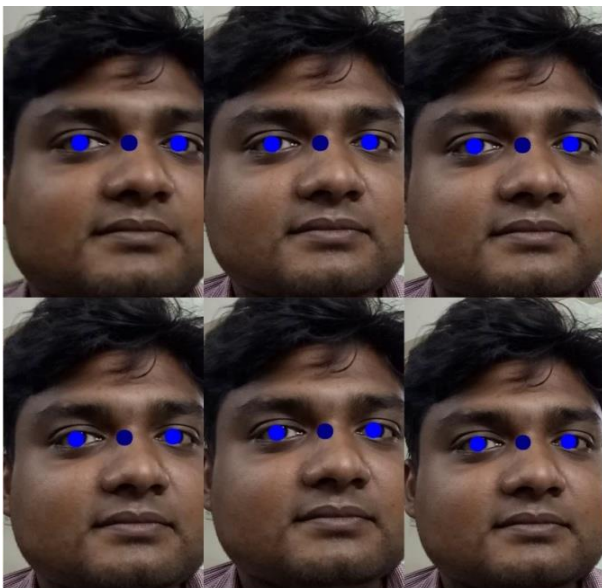


Fig.8. Movement of midpoint in consecutive frames for a person who is not trying to avoid eye contact.

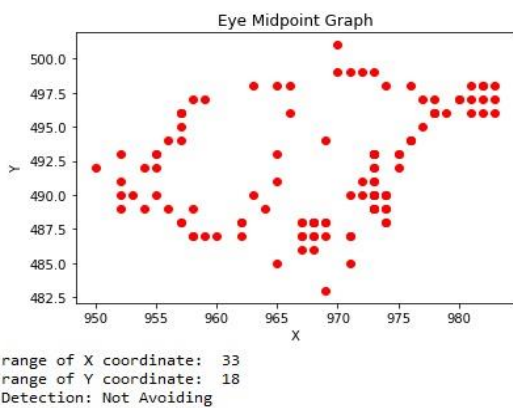


Fig.9. Result of a “Not frequent iris movement” video.

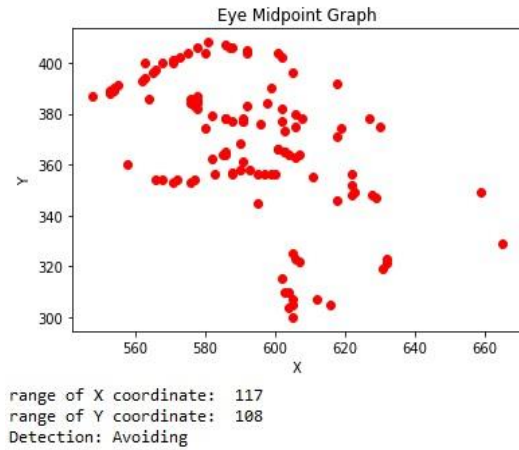


Fig.10. Result of a “Frequent iris movement” video.

The confusion matrix of our experimental result for the first 100 frames of each video is shown in Table 1. As each video has a frame rate of 30 fps, in this experiment we consider the first 3.33 second of each of our videos. The confusion matrix of our experimental result for the first 150 frames of each video is shown in Table 2. With a rate of 30 frames per second this experiment considers the first 5 seconds of each video.

Table 1. Confusion Matrix for First 100 Frames

		System Output (Prediction)		
		A	N	
gold labels	A	37	7	Recall ‘A’ = $\frac{37}{37+7}$ = 0.84
	N	2	40	Recall ‘N’ = $\frac{40}{2+40}$ = 0.95
		Precision ‘A’ = $\frac{37}{37+2}$ = 0.95	Precision ‘N’ = $\frac{40}{7+40}$ = 0.85	Accuracy = $\frac{37+40}{86}$ = 0.90
Macro-average Precision =		$\frac{0.95+0.85}{2} = 0.90$		Macro-average Recall = $\frac{0.84+0.95}{2} = 0.90$

In Table 1, the gold labels represent the actual category of the videos. Here ‘A’ represents frequent iris movement and ‘N’ represents not frequent iris movement. The system output or prediction represents the detected class by our classifier. Similarly in Table 2 the gold labels and system output or prediction represents the same thing as Table 1. Also the labels ‘A’ and ‘N’ have the same meaning as Table 1.

We can see that our classifier detected 37 frequent iris movement videos correctly and 40 not frequent iris movement correctly for the first 100 frames or in first 3.33 seconds (Table 1). But from Table 2 we can see that

our classifier can detect 41 frequent iris movement videos as ‘A’ and 39 not frequent iris movement videos as ‘N’ for the first 150 frames or in the first 5 seconds of each of our videos in the dataset.

Table 2. Confusion Matrix for First 150 Frames

		System Output (Prediction)		
		A	N	
gold labels	A	41	3	Recall ‘A’ $= \frac{41}{41+3} = 0.93$
	N	3	39	Recall ‘N’ $= \frac{39}{3+39} = 0.93$
		Precision ‘A’ $= \frac{41}{41+3} = 0.93$	Precision ‘N’ $= \frac{39}{3+39} = 0.93$	Accuracy $= \frac{41+39}{86} = 0.93$
	Macro-average Precision = $\frac{0.93+0.93}{2} = 0.93$		Macro-average Recall = $\frac{0.93+0.93}{2} = 0.93$	

Finally we compare our work with the existing works on questionable observer detection or suspicious activity recognition. Except one case our result shows the highest accuracy. The comparison is shown in Table 3.

Table 3. Comparison of Our Work with Existing Works

Detection	Dataset	Accuracy
Real-time detection of violent crowd behavior [7].	(1) Own assembled database (2) Hockey database	(1) 81.30 ± 0.21% (±SD) (2) 82.90 ± 0.14% (±SE)
Abnormal crowd behavior detection using social force model [9].	(1) UMN dataset (2) Web dataset	(1) 96% (2) 73%
Moment invariants based human mistrustful and suspicious motion detection, recognition and classification [13].	Own developed dataset	87.6%
Detecting questionable observers using face track clustering [17].	ND-QO-Flip dataset	96%
Automated Real-Time Detection of Potentially Suspicious Behavior in Public Transport Areas [29].	BEHAVE CAVIAR and PETS 2006.	66%
Questionable observer detection depending on an activity “frequent iris movement”	Iris Movement Dataset	93%

D. Analysis and Discussion

From Table 1 and 2 we can see that the accuracy obtained for the first 100 frames is 90% and the macro-average precision and recall is 90% and 90% respectively and the accuracy obtained for the first 150 frames is 93% and the macro-average precision and recall is 93% for both.

The comparison of accuracy, macro-average precision and recall between our two experiments is shown in Fig. 11. From here we can see that for 150 frames we get slightly better accuracy, as well as better macro-average

precision and recall. The main reason behind that is: iris movement in first 3.33 seconds is less than the movement in first 5.00 seconds. For this reason we get slightly better

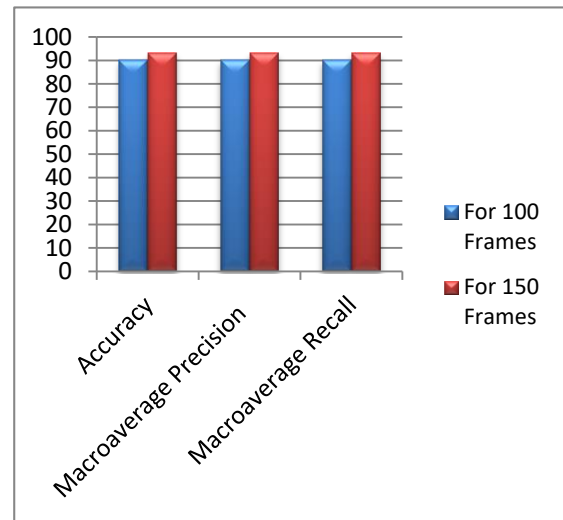


Fig.11. Comparison between our two experiments on basis of accuracy, macro-average precision and recall

result for 150 frames than 100 frames. As macro-average only computes the metric independently for each class and then take the average, it treats all classes equally. But micro-average is more preferable to find out the class imbalance, which indicates the presence of many more examples of one class than of other classes. For that reason we show the comparison of our two experiments on basis of micro-average precision and recall for both class ‘A’ and ‘F’ in Fig. 12.

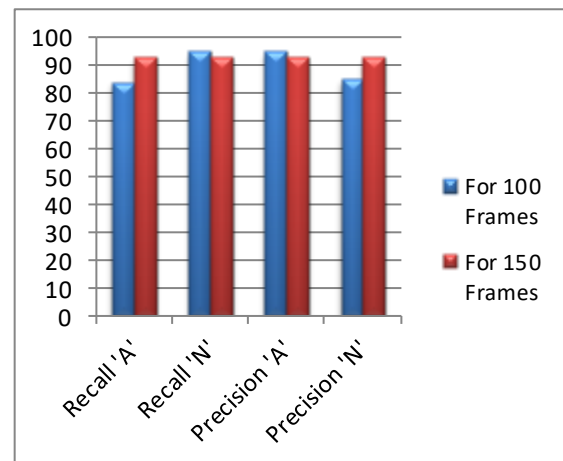


Fig.12. Comparison between our two experiments on basis of micro-average precision and recall

From here we can see that the micro-average recall of class ‘N’ and the micro-average precision of class ‘A’ are almost same for both of our experiments. But the micro-average recall of class ‘A’ and precision of class ‘N’ differs by a good margin in our experiments. The experiment for 100 frames reflects a poor micro-average recall and precision of 84% and 85% for class ‘A’ and



‘N’ respectively; where the micro-average recall and precision of class ‘A’ and ‘N’ is 93% and 93% respectively in the experiment where we consider 150 frames. This finding indicates that a large number of examples are not properly classified in our first experiments, where the second experiment indicates a balanced classification.

Finally we compare our work on basis of accuracy with the existing works of questionable observer detection and suspicious activity recognition. The comparison is shown graphically in Fig. 13.

In Fig. 13, the label 7-(1) represents the accuracy of the work performed by Hassner et al. in their first dataset as shown in Table 3. Similarly the 7-(2) shows the accuracy of their second database. They have achieved an accuracy of  $81.30 \pm 0.21\%$  ( $\pm$ SD) on their own assembled database and an accuracy of  $82.90 \pm 0.14\%$  ( $\pm$ SE) on Hockey database [7]. They have provided a novel approach to real-time detection of breaking violence in crowded scenes by using the concept of changing flow-vector magnitudes over time and by using linear SVM, they have classified their videos into two classes: violent or non-violent.

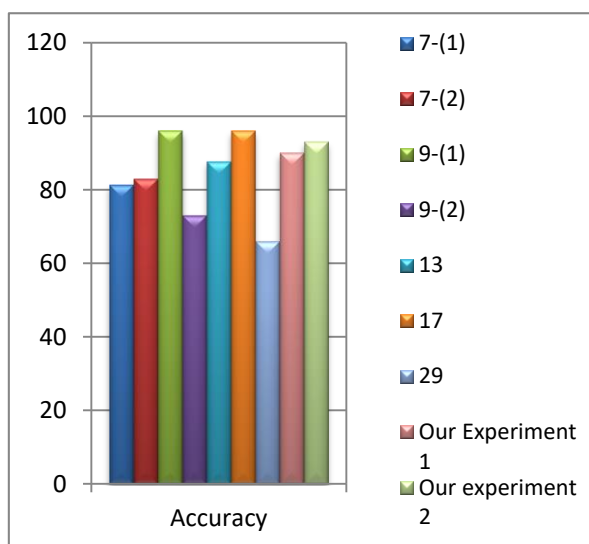


Fig.13. Comparison of our two experiments with existing works on basis of accuracy

In Fig. 13, labels 9-(1) and 9-(2) represents the accuracy of 96% on UMN dataset and 73% on web dataset of the work done by Mehran et al. [9]. On basis of the concept of optical flow and by using social force model they have introduced a novel method to detect and localize abnormal behaviours in crowd videos. The label 13 represents the accuracy of the work done by Yasin et al. [13]. They have recognized and classified human mistrustfulness and suspicious motion based on moment invariants on their own developed dataset with an accuracy of 87.6%. The label 17 represents the work of Barr et al., where they have used face track clustering method to detect questionable observers on ND-QO-Flip dataset with an accuracy of 96% [17]. In Fig. 13, the label 29 represents the accuracy achieved by Elhamod et al. in

their work of suspicious behaviours detection in public transport areas. They have achieved an accuracy of 66% on BEHAVE CAVIAR and PETS 2006 dataset. Finally the last two bars represent the accuracy of our work in our own developed dataset. We have achieved an accuracy of 90% and 93% on our experiment 1 and experiment 2 respectively. Here experiment 1 is done with the first 100 frames of our dataset and experiment 2 is done with the first 150 frames of our dataset.

On basis of the above discussion we can say that we have achieved an accuracy of 93%, which is better than most of the cases. No work has been done on the detection of questionable behaviour on basis of the specific activity we have used here. Also no classification of questionable behaviour or suspicious activity has been done by using haar cascade classifier. So by using a new approach for questionable observer detection we have achieved a reasonable accuracy of 93%.

## V. CONCLUSION

As frequent iris movement detection is a core part of questionable observer detection; it has a great importance in this field of research. The detection of questionable observer has some great significance for the law enforcement officials. It can also prevent and detect theft in shopping malls and other places and can also prevent violence or riot in public place and may lead to save a lot of life and goods. There are lots of suspicious activities and behaviors which can categorize one person as questionable observer, and among them we have detected one – frequent iris movement. We have detected this activity with an accuracy of 93%. Building of comprehensive datasets with various situations to validate the proposed framework is a huge challenge in this research. To validate the proposed framework, datasets used by this research is in different type of surroundings independent of controlled pose, illumination, and facial expression which contains noise such as camera shake. As our work is vision based detection we have overcome all the limitations of sensor based action recognition. We not only detected a new action but also made a dataset to validate our work of frequent iris movement, which plays a vital role for questionable observer detection.

## REFERENCES

- [1] Takai, Miwa. "Detection of suspicious activity and estimate of risk from human behavior shot by surveillance camera." In *Nature and Biologically Inspired Computing (NaBIC), 2010 Second World Congress on*, pp. 298-304. IEEE, 2010.
- [2] Doewes, Afrizal, Sri Edi Swasono, and Bambang Harjito. "Feature selection on Human Activity Recognition dataset using Minimum Redundancy Maximum Relevance." In *Consumer Electronics-Taiwan (ICCE-TW), 2017 IEEE International Conference on*, pp. 171-172. IEEE, 2017.
- [3] Dhulekar, P. A., S. T. Gandhe, Anjali Shewale, Sayali Sonawane, and Varsha Yelmame. "Motion estimation for human activity surveillance." In *Emerging Trends & Innovation in ICT (ICEI), 2017 International Conference on*, pp. 82-85. IEEE, 2017.



- [4] Hsu, Yu-Liang, Shyan-Lung Lin, Po-Huan Chou, Hung-Che Lai, Hsing-Cheng Chang, and Shih-Chin Yang. "Application of nonparametric weighted feature extraction for an inertial-signal-based human activity recognition system." In *Applied System Innovation (ICASI), 2017 International Conference on*, pp. 1718-1720. IEEE, 2017.
- [5] Xu, Wanru, Zhenjiang Miao, Xiao-Ping Zhang, and Yi Tian. "Learning a hierarchical spatio-temporal model for human activity recognition." In *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*, pp. 1607-1611. IEEE, 2017.
- [6] Shakya, Subarna, Suman Sharma, and Abinash Basnet. "Human behavior prediction using facial expression analysis." In *Computing, Communication and Automation (ICCCA), 2016 International Conference on*, pp. 399-404. IEEE, 2016.
- [7] Hassner, Tal, Yossi Itcher, and Orit Kliper-Gross. "Violent flows: Real-time detection of violent crowd behavior." In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*, pp. 1-6. IEEE, 2012.
- [8] Karagiannaki, Katerina, Athanasia Panousopoulou, and Panagiotis Tsakalides. "An online feature selection architecture for Human Activity Recognition." In *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*, pp. 2522-2526. IEEE, 2017.
- [9] Mehran, Ramin, Alexis Oyama, and Mubarak Shah. "Abnormal crowd behavior detection using social force model." In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pp. 935-942. IEEE, 2009.
- [10] Gowda, Shreyank N. "Human activity recognition using combinatorial Deep Belief Networks." In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on*, pp. 1589-1594. IEEE, 2017.
- [11] Boufama, Boubakeur, Pejman Habashi, and Imran Shafiq Ahmad. "Trajectory-based human activity recognition from videos." In *Advanced Technologies for Signal and Image Processing (ATSIP), 2017 International Conference on*, pp. 1-5. IEEE, 2017.
- [12] Uddin, Md Zia, Weria Khaksar, and Jim Torresen. "Human activity recognition using robust spatiotemporal features and convolutional neural network." In *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2017 IEEE International Conference on*, pp. 144-149. IEEE, 2017.
- [13] Yasin, Hashim, and Shoab Ahmad Khan. "Moment invariants based human mistrustful and suspicious motion detection, recognition and classification." In *Computer Modeling and Simulation, 2008. UKSIM 2008. Tenth International Conference on*, pp. 734-739. IEEE, 2008.
- [14] Matsui, Shinya, Nakamasa Inoue, Yuko Akagi, Goshu Nagino, and Koichi Shinoda. "User adaptation of convolutional neural network for human activity recognition." In *Signal Processing Conference (EUSIPCO), 2017 25th European*, pp. 753-757. IEEE, 2017.
- [15] Chen, Zhenghua, Le Zhang, Zhiguang Cao, and Jing Guo. "Distilling the Knowledge from Handcrafted Features for Human Activity Recognition." *IEEE Transactions on Industrial Informatics* (2018).
- [16] Sunkad, Zubin A. "Feature Selection and Hyperparameter Optimization of SVM for Human Activity Recognition." In *Soft Computing & Machine Intelligence (SCMI), 2016 3rd International Conference on*, pp. 104-109. IEEE, 2016.
- [17] Barr, Jeremiah R., Kevin W. Bowyer, and Patrick J. Flynn. "Detecting questionable observers using face track clustering." In *Applications of Computer Vision (WACV), 2011 IEEE Workshop on*, pp. 182-189. IEEE, 2011.
- [18] Zhao, Kun, Wei Xi, Zhiping Jiang, Zhi Wang, Hongliang Luo, Jizhong Zhao, and Xiaobin Zhang. "Leveraging Topic Model for CSI Based Human Activity Recognition." In *Mobile Ad-Hoc and Sensor Networks (MSN), 2016 12th International Conference on*, pp. 23-30. IEEE, 2016.
- [19] Maglogiannis, Ilias, Demosthenes Vouyioukas, and Chris Aggelopoulos. "Face detection and recognition of natural human emotion using Markov random fields." *Personal and Ubiquitous Computing* 13, no. 1 (2009): 95-101.
- [20] Cheng, Long, Yani Guan, Kecheng Zhu, Yiyang Li, and Ruokun Xu. "Accelerated Sparse Representation for Human Activity Recognition." In *Information Reuse and Integration (IRI), 2017 IEEE International Conference on*, pp. 245-252. IEEE, 2017.
- [21] De Silva, Liyanage C., Tsutomu Miyasato, and Ryohei Nakatsu. "Facial emotion recognition using multi-modal information." In *Information, Communications and Signal Processing, 1997. ICICS., Proceedings of 1997 International Conference on*, vol. 1, pp. 397-401. IEEE, 1997.
- [22] Li, Kang, Xiaoguang Zhao, Jiang Bian, and Min Tan. "Sequential learning for multimodal 3D human activity recognition with Long-Short Term Memory." In *Mechatronics and Automation (ICMA), 2017 IEEE International Conference on*, pp. 1556-1561. IEEE, 2017.
- [23] Lee, Song-Mi, Heeryon Cho, and Sang Min Yoon. "Statistical noise reduction for robust human activity recognition." In *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2017 IEEE International Conference on*, pp. 284-288. IEEE, 2017.
- [24] Aramvith, Supavadee, Suree Pumrin, Thanarat Chalidabhongse, and Supakorn Siddhichai. "Video processing and analysis for surveillance applications." In *Intelligent Signal Processing and Communication Systems, 2009. ISPACS 2009. International Symposium on*, pp. 607-610. IEEE, 2009.
- [25] Li, Wanqing, Igor Kharitonenko, Serge Lichman, and Chaminda Weerasinghe. "A prototype of autonomous intelligent surveillance cameras." In *Video and Signal Based Surveillance, 2006. AVSS'06. IEEE International Conference on*, pp. 101-101. IEEE, 2006.
- [26] Chen, Wen-Hui, Carlos Andrés Betancourt Baca, and Chih-Hao Tou. "LSTM-RNNs combined with scene information for human activity recognition." In *e-Health Networking, Applications and Services (Healthcom), 2017 IEEE 19th International Conference on*, pp. 1-6. 2017.
- [27] Savvaki, Sofia, Grigorios Tsagkatakis, Athanasia Panousopoulou, and Panagiotis Tsakalides. "Matrix and Tensor Completion on a Human Activity Recognition Framework." *IEEE journal of biomedical and health informatics* 21, no. 6 (2017): 1554-1561.
- [28] Jarraya, Amina, Khedija Arour, Amel Bouzeghoub, and Amel Borgi. "Feature selection based on Choquet integral for human activity recognition." In *Fuzzy Systems (FUZZ-IEEE), 2017 IEEE International Conference on*, pp. 1-6. IEEE, 2017.
- [29] Elhamod, Mohannad, and Martin D. Levine. "Automated real-time detection of potentially suspicious behavior in public transport areas." *IEEE Transactions on Intelligent Transportation Systems* 14, no. 2 (2013): 688-699.
- [30] Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." In *Computer*

*Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 1, pp. I-I. IEEE, 2001.

### Authors' Profiles



**D. M. Anisuzzaman** is a Lecturer with Computer Science and Engineering Department, Ahsanullah University of Science and Technology, Dhaka, Bangladesh. He is pursuing his MSc in Computer Science from American International University - Bangladesh, Dhaka. He has received his BSc from Ahsanullah University of Science and

Technology, Dhaka in 2013. His research interest includes machine learning, neural network, computer vision, natural language processing and algorithms.



**A.F.M. Saifuddin Saif** received Ph.D. from Faculty of Information Science and Technology, University Kebangsaan Malaysia (UKM) in 2016. He received M.Sc. in Computer System Engineering (Software System) from University of East London, UK, and B.Sc. (Eng.) degree in Computer Science and Engineering from Shahjalal University of

Science and Technology, Bangladesh in 2012 and 2008, respectively. Most of his contributions in Computer Vision and Artificial Intelligence Research field were published in ISI Q1 journals. He has published many papers in ISI indexed Journals; Scopus indexed Journals, Book Chapters, Conferences, and Proceedings. He served as Technical Committee Members, Reviewers, Guest Speakers, Session Chairs in many Conferences and Workshops. Currently, he is an Assistant Professor at Faculty of Information Science and Technology, American International University Bangladesh (AIUB). Before joining the university, he did Post Doctorate at Faculty of Information Science and Technology, University Kebangsaan Malaysia. He spent more than 6 years in IT industry such as Advanced Software Development, Web eMaze etc as IT researcher. His research interests include Image Processing, Computer Vision, Artificial Intelligence, Augmented Reality, 3D Reconstruction, and Medical Image Processing.

**How to cite this paper:** D. M. Anisuzzaman, A. F. M. Saifuddin Saif, "Efficient Framework Using Morphological Modeling for Frequent Iris Movement Investigation towards Questionable Observer Detection", *International Journal of Image, Graphics and Signal Processing(IJIGSP)*, Vol.10, No.11, pp. 28-37, 2018.DOI: 10.5815/ijigsp.2018.11.04