A New Framework for Video-based Frequent Iris Movement Analysis towards Anomaly Observer Detection

Md. Minhaz Ur Rahman[1]
[1]Ahsanullah University of Science and Technology/CSE, Dhaka, 1215, Bangladesh
Email: minhaz.acps@gmail.com

Mahmudul Hasan Robin[2], Abu Mohammad Taief[2]
[2] Ahsanullah University of Science and Technology/CSE, Dhaka, 1215, Bangladesh
Email: {mahmudulrobin17, taiefmaiden}@gmail.com

Abstract: This paper suggested a new framework for detecting abnormal behavior, specifically based on frequent iris movements. It contributed to a decision whereas an individual is dubious or unsuspected from a video. One of the key components of questionable observer detection is to detect some specific suspicious activity. According to the writer, various areas of the body movement and human behaviors may be an indicator of suspicious behavior. In this research, we considered the movement of human eyes to identify suspicious activity. This working field is also a significant aspect of machine vision and artificial intelligence, and a big part of the understanding of human behavior. The system framework comprises three parts to monitor suspicious video activities. First, we used the Multi-task Cascaded Convolutional Networks (MTCNN) classifier to detect eyes. Second, we observe irises from eye representations with the use of Circular Hough Transformation (CHT). Finally, we calculated the average distance of iris movement from eye images using a new morphological method called TRM using some properties of the iris movement. We have observed a particular phenomenon of frequent iris movement. Hence, we are making a case of someone being an abnormal person and referring it to a suspicious observer. To vouch for our work, we created our data set with 100 videos where 30 individuals volunteered to validate this research. Each video comprises 200 frames with a duration of 6-10 seconds. We've reached an accuracy of 94% on detecting a frequent iris movement. Rather the goal is to minimize people's burdens so they can focus on a small range of cases for investigation in more depth. This research's sole purpose is to indicate a person's anomalous behavior on the basis of frequent iris movement. Our research outstrips much of the current literature on abnormal iris movement and dubious investigator identification.

Index Terms: Suspicious activity detection, frequent iris movement, eye detection, iris detection, iris movement, TRM iris dataset

1. Introduction

In recent years, the implementation of automated monitoring technologies for video recognition systems to recognize suspicious behavior has become an increasing phenomenon. Such methods remove human factors influencing safety violation device performance [1]. Although several of research has been performed on simple human behavior detection (strolling [2-6], running [7-10], sit idle [11-14], etc. [15-19]); as far as we know, none of the behaviors listed above has been identified for dubious observer detection. Very little research has been conducted on irregular behavior identification [21] and dubious spectator detection [20, 22, 23] and the identification of disruptive crowd movements [24].

In this study, we suggested a technique that could suggest irregular activity and point to unusual human actions by repeated movement of a person's iris. Maximum accused people seek to prevent touch with their eyes and attempt to perceive the environment or circumstances while using the side of their eyes. Such people typically switch their iris vertical or horizontal more frequently than a normal person [20]. When a questionably suspect observer tries to hide his or her goal or remorse, he or she may most of the time stop eye contact, which may be a good sign for a dubious observer. Our research main objective is to help a problematic observer detector from occurring and also a suspicious individual is identified before the offense is committed. The intelligent mechanism that can identify dubious observers; the intelligent machine can track and identify all the questionable observers within its range.

We followed some steps to construct the proposed method. First, we identified the eyes from videos by applying our dataset to the MTCNN method. After that, we used Circular Hough Transformation (CHT) and Canny Edge Detection (CED) to identify iris from eye images. Finally, we have established a novel morphological approach to the detection of irregular human activity that leads to a suspect observer. We have addressed this morphological method as the TRM method. We have observed one of the many gestural cues of a suspicious human being. It's not going to be enough to categorize someone as questionable. Certain non-verbal signs that are questionable must always be dealt with. If a person does all this, he will be called a dubious person. Yet we can definitely say that the TRM system can identify irregular iris changes in a particular and special manner, which is one of the conditions for finding anyone suspicious. An Efficient Framework Using Morphological Modeling for Frequent Iris Movement Investigation towards Questionable Observer Detection was introduced in [21]. In that research, the author used a morphological method of his own to detect frequent iris movement. But our accuracy rate surpasses its result. There are few limitations in our study such as we have used a pre-trained data set to detect iris for our research work. Hence in the future, we will surely overcome this obstacle and try to make this research trustworthy so that it can help human investigators with their trials and making definitive decisions towards the questionable observers.

2. Related Works

Several research papers in the areas of identification of human activity, behavioral understanding, and messy scenario analysis have been reported that apply to video-based anomaly detection clearly or indirectly [25‑27].

Facial expression and human action recognition are also relevant to this study since any particular facial expression and actions can suggest suspicious conduct. These works are also entirely based on computer vision [23, 29].

Noureddine [30] proposed a robust attribute recognition method for frontal face images based on a circular Hough transformation. The edges in the input image, using edge detection techniques, can be detected in the first case. The second is to establish the identifying points of rings from which the symbol of iris is identified. For 396 individuals with 24 frames per human, the proposed algorithm is tested for particular situations in the Caltech Database. 379 pictures of 396 photographs were easily seen, with a performance rate of 95 percent. Significant gains are achieved in a variety of color models. Poor experiments were found for observable objects where Gaussian noise was greater than 0.1.

Depending on deep convolutional network techniques, Zhang [31] has adopted three phases of deep convolutional networks that can predict the gross-to-fine position of the face and landmarks in a large way. A new analysis has shown that in this field, deep learning techniques can produce tremendous results. The author has proposed CNNs for eye detection consisting of three phases: the Proposal Network (P-Net), the Refinement Network (R-Net), and the Output Network (O-Net). Experimental results find that these methods exceed the state-of-the-art methods over multiple demanding tests while maintaining real-time efficiency.

Xiang et al. [32] model both uncommon and natural behaviors. The whole system conserves both the regular and unusual structures. Both normal and abnormal behavioral classes are amended as soon as one of those categorizes the detected behavioral examples. To determine whether a behavioral instance is irregular, the Likelihood Ratio Test (LRT) measures the chance of normal or irregular activity. All dysfunctional and regular behavior trends will lead to their equivalent (e.g. abnormality is usual). This system relies on the weight of each norm and the weight of the two thresholds (i.e. the lowest weight for the normal form of conduct and the highest weight for the abnormal behavior). This work examines more systematically the use of contexts. The study, unfortunately, since the developers have argued that the thresholds are not immune, using such thresholds is one of the system's disadvantages. And if the weight of the action is weighted to one, naturally, the weights decrease with the rise in the number in compartment classes. The threshold values also need to be adjusted accordingly.

Although some approaches [33, 34] suggest that contextual information can be used to further enhance system performance, the use of contextual information is still limited in suspicious behavioral detection. Situation information may be used to classify system awareness into awareness groups. In the same context, every other category reflects information gained. The system performance can only be enhanced by using a knowledge culture that is relevant to the current context. The result of failing to consider contextual information may not be that the technique could detect subjective suspicious behavior. Situational suspicious behavior is characterized as behavior that is perceived to be questionable only in a specific context.

From the above discussion it can be said that the respected authors have worked on computer vision regarding human behavior detection which gave us the motivation to start our work in the first place. Hence we have invented our new morphological (TRM) method for the detection of abnormal and anomalous activity based on specific behavioral action is the first of its kind in the field of computer vision.

3. Proposed Research Methodology

In this article, we presented a new paradigm for the method of identifying a person's irregular behavior from the iris movement. In this step, we detected eyes using the MTCNN method and detected iris using the Circular Hough Transformation (CHT) system. Later, using the morphological model (TRM), we measured the irregular and frequent

activity of the iris. MTCNN works perfectly under critical lighting conditions thus surpassing numerous object detection algorithms while applying to our dataset. CHT is another perfect circle detecting method which can easily detect Circle from eyes which is the iris. And finally as a normal human's iris can move to a certain point thus we proposed our new morphological method using this trait of Iris. The key phases of our proposed model are outlined in Fig.1.

```
                    ╭─────────────╮
                    │   Dataset   │
                    │ Collection  │
                    ╰─────────────╯
                           │
                           ▼
      ┌────────────────────────────────────┐
      │ Pre-processing                     │
      │ ---                                │
      │    1. Resolution Reduction         │
      │    2. Histogram Equalization       │
      │    3. Noise Reduction              │
      │    4. Resolution Reduction         │
      │    5. Image Pyramid                │
      └────────────────────────────────────┘
                           │
                           ▼
      ┌────────────────────────────────────┐
      │ Eye Detection Using MTCNN          │
      │ ---                                │
      │    1. First Proposal Network       │
      │       (P-NET)                      │
      │    2. Refinement Network           │
      │       (R-NET)                      │
      │    3. Output Network               │
      │       (O-NET)                      │
      └────────────────────────────────────┘
                           │
                           ▼
      ┌────────────────────────────────────┐
      │ Iris Detection                     │
      │ ---                                │
      │    1. Pre-processing Dataset       │
      │        a. Median Filtering         │
      │        b. Gamma Correction         │
      │    2. Hough Transformation         │
      └────────────────────────────────────┘
                           │
                           ▼
      ┌────────────────────────────────────┐
      │                                    │
      │ Calculating the average distance   │
      │ of iris center point for iris      │
      │ movement by using morphological    │
      │ model and detecting recurrent      │
      │ iris movement.                     │
      │                                    │
      └────────────────────────────────────┘
                           │
                           ▼
                    ╭─────────────╮
                    │     END     │
                    ╰─────────────╯
```
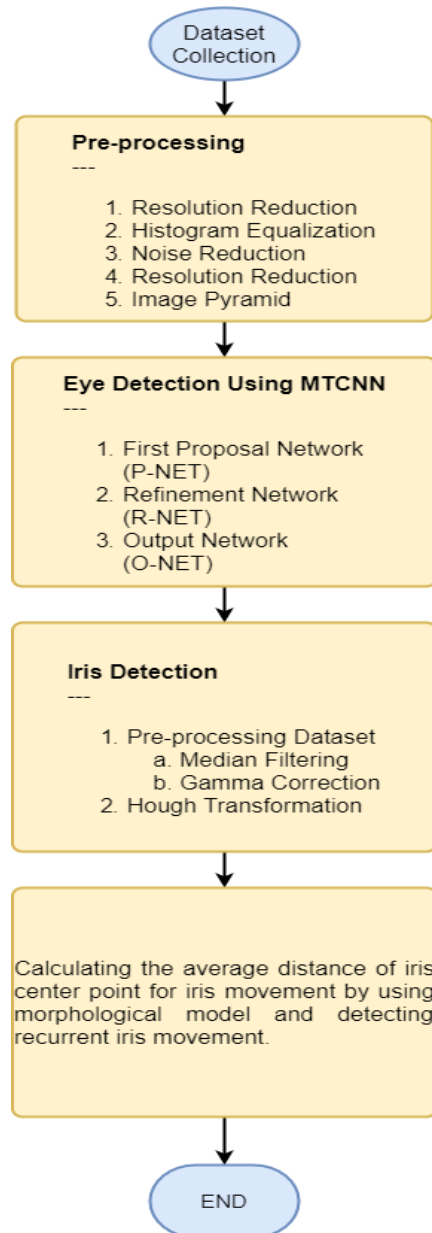
Fig.1. Flow chart diagram for proposed method

The Flowgraph of our morphological model for identifying abnormal and suspicious human behavior has been illustrated in Fig.2.
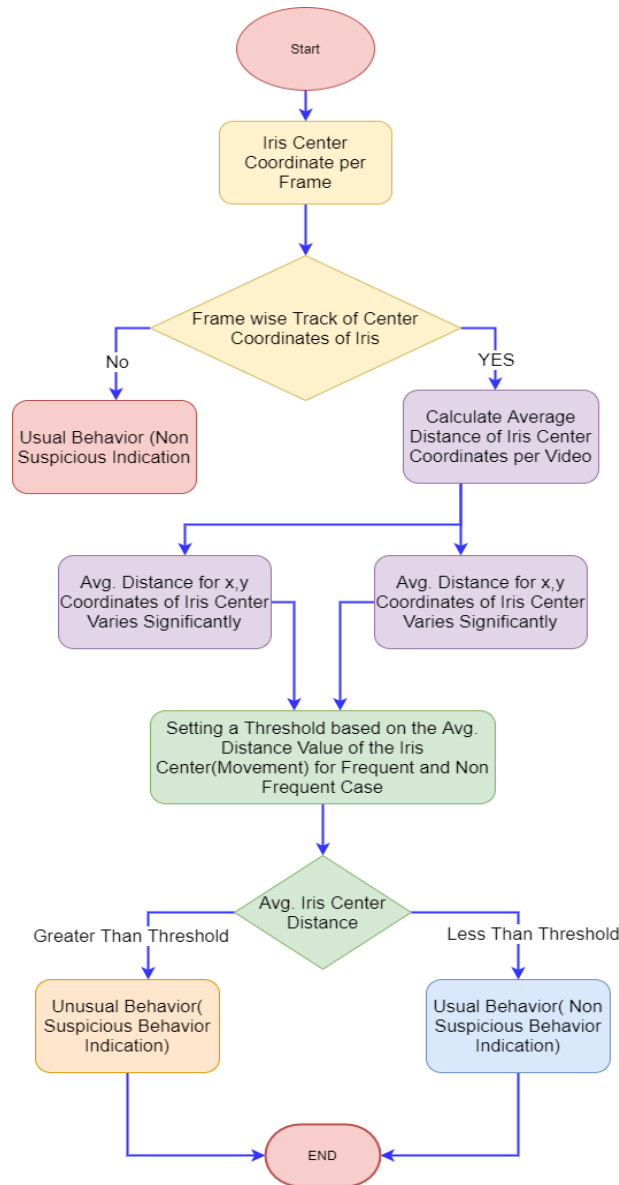
Fig.2. Flow graph of suggested morphological (TRM) method

A. Data Collection

Our dataset has been generated and comprises 100 videos. Such videos were taken from volunteers. Light requirements were correct, and then all videos were taken from a single phone to show voluntary facial vision. There were also people with unusual eye disorders. Each video has about 150-200 frames and is 6-10 seconds long. Also, every video was taken using only a 12-megapixel cell phone with a resolution of 720x1080 pixels. The distance from the object to the human being is about 1 to 3 meters. Growing the gap is going to make the machine unstable, so we took all the videos from a closer range. Fig.3 illustrates sample images of our dataset.

Fig.3. Sample images of our dataset

B. Data Preprocessing

Data preprocessing is a very important component of image analysis. Every other frame in each video has a resolution of 720 x 1080p. The histogram was used to evaluate the range of the lightness of every bottle. Alternatively, noise control is conducted with the help of certain morphological operations: closing, erosion, and opening. We have created an image pyramid, an input for the detection of eyes using MTCNN. And after that, we re-dimension the pyramid at a higher size.

C. Eye Detection using MTCNN

Once the frames have indeed been preprocessed, the input images in the MTCNN are used to identify the eyes. It has to go through three phases [31]:

1) Stage-I: This phase is known as the Proposal Network (P-Net) which receives the openings of the candidate and the vectors of its boundary box. So we use bounding box regression vectors that are estimated to control the participants. Non-maximum suppression (NMS) may then be used to merge strongly reflected participants.

2) Stage-II: The next CNN, dubbed the Refine Network (R-Net), also removes a significant portion of the undesirable contestants, advocates boundary box regression and incorporates NMS members.

3) Stage-III: This step is the same as the previous phase, except that the mask seems to be better defined at this step. The circle is drawn to highlight and, eventually, to differentiate the eyes from the face.

D. Iris Detection using Circular Hough Transformation

After having successfully defined the eyes of our study, we planned to detect the true border of the iris and finally detect the iris. This is accomplished by the Canny Edge Detection and Hough Transformation method. The key stages for iris detection: median filter and gamma correction implementation and applying CED & CHT.

1) Median Filter: This median filtering before processing is used to eliminate noise from images. Here, the median is determined by combining all pixel values from the screen to the quantitative series then switching the pixel to the middle (median) pixel value. The example of applying a median filter to an eye image from our dataset to reduce image noise is given in Fig.4.
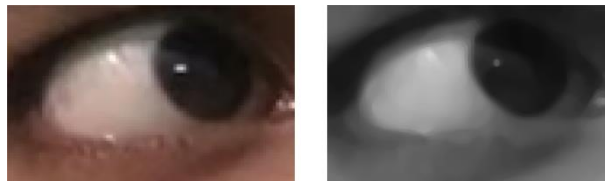

Fig.4. Sample output of median filtered image.

2) Gamma Correction: As our system does not capture and display the luminescence properly, the gamma correction is being used to correct the pixel fluorescence. It can also contribute to low contrast if the gamma value is too high or too small. An example of applying gamma correction to an eye image from our data set to rectify the luminescence of the image is illustrated in Fig.5.
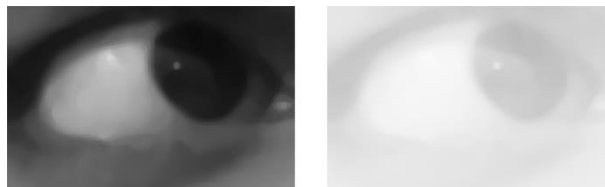

Fig.5. Sample output image after applying gamma correction.

3) Performing CED and CHT: Initially, edges are detected using Canny Edge Eye Image Detection, which has been analyzed and extracted from our dataset. The iris is determined from the eye images after the Hough Transformation process has been applied.

In our proposed approach, the Canny Edge Detection algorithm was used to categorize edges and at the same time suppressing noise. A Gaussian filter has been used to eliminate noise and unwanted bits.

After the execution of the CED, we applied the Circular Hough Transformation, which further allowed circular objects to be ejected from the image, even if the circle is incomplete. The Hough transformation can be defined as a transformation of the X, Y-plane space parameter. Statistically, the circle equation in x, y-plane is given by

$$(x + a)^2 + (y + k)^2 = r^2 \tag{1}$$

Where the circular radius is r and the center of the circle is a, b. The parametric method of the sphere is given by

$$x = r\cos(t), y = r\,sint(t)$$

When the circle's coordinates are x, y, the circle's radius is r, and t is the parameter that defines the angle of the circle at the center of the circle. The circle is shaped at the point of the bottom in the middle of the point, taking the optimal distance into account. Miscellaneous circles may cause false centers to be observed in the image of the eye. Across the whole phase of effective iris detection, we identified that the middle point of the first circle was between all the other center points of the circles created for the edge points [35]. Throughout this process, we regarded parameters such as the Hough parameter, canny parameter, Maximum radius, and Maximum distance, where a certain value of the parameters is adjusted to locate the iris with the accuracy of the eye pictures.

E. TRM Model to Calculate the Average Distance of Iris Center Point and Frequent Iris Movement Detection

Upon detection of the iris from eye images, we noted the change in the x, y coordinates of the iris center point for consecutive frames in the video. Then, finally, a new morphological model (TRM Method) was used to detect the anomaly and abnormal iris movement from one eye. As a result, we can indicate and detect abnormal human behavior through frequent iris movements which try to avoid eye contact eventually. This research will divide the unusual movement of iris from the normal movement of the iris by checking the significant change in the distance from iris centroid. Initially, we took the center of the iris to measure the distance from the initial position of the iris to the position of the next frame. Then check all the frames of more than 50 videos. We have calculated the average change in distance for iris centroid point from the initial center point. The average center coordinates distance (Center X, Center Y) of the iris is detected by using the formula:

$$CenterX = \frac{\sum_{i=0}^{n}(|frameX_i - frameX_{i+1}|)}{n} \tag{2}$$

$$CenterY = \frac{\sum_{i=0}^{n}(|frameY_i - frameY_{i+1}|)}{n} \tag{3}$$

Where, n = Total number of frames,

$frameX_i$ = Current frame x coordinate value,
$frameY_i$ = Next frame x coordinate value,
$frameX_{i+1}$ = Current frame Y coordinate value,
$frameY_{i+1}$ = Next frame Y coordinate value

The x coordinate of the average distance of the center point is determined by dividing the sum of the distance crossed by the x coordinates of the iris by the total number of frames. In the same way, the y coordinate of the average distance of the center point is determined by dividing the sum of the distance crossed by the y coordinates of the iris by the total number of frames. By using this average distance from the iris center point, we observed our detection in the next step.

In all successive frames, we recorded the range of motion of the center point of the iris. Calculating the average distance value of the motion of the center point of the iris, we set the threshold value by testing our research with 50 videos. After viewing these 50 videos effectively, we found that when a person attempts to prevent visual contact or changes the iris rapidly from right to left, or vice versa, the x-coordinate range will fluctuate significantly from frame to frame, and the average distance of iris center point's x coordinate will be greater than that of a person who does not try to avoid eye contact or shifts iris frequently. In the same way when a person attempts to prevent visual contact or changes the iris rapidly from right to left, or vice versa, the y-coordinate range will fluctuate significantly from frame to frame, and the average distance of iris center point's y coordinate will be greater than that of a person who does not try to avoid eye contact or shifts iris frequently. Therefore, when measuring the average distance value of the iris center point of the video, we detected irregular and unusual iris motions. This work will group the behavior of avoiding eye contact and significant movement of the iris if the average distance value of the iris center point of the video is greater than the threshold value that we have assigned.

Fig.6 elucidates the average distance of the iris center point movement of (x, y) coordinates in our dataset for an avoiding and significant iris movement case. It represents the variation of average distance of iris center point of both x and y coordinates for the case of frequent iris movement.

```
Center Coordinates: [(24.156, 11.468), (16.339, 8.830),

(30.807, 24.947), (23.318, 10.964), (17.158, 12.486),

(14.916, 10.813), (17.143, 11.158), (16.056, 13.380),

(29.3, 13.96), (10.45, 14.89), (12.03, 18.88),

(14.09, 20.99), (19.47, 8.68), (27.94, 8.90),

(13.56, 26.70), (18.58, 13.81), (10.6, 20.67),

(11.24, 24.63), (12.42, 18.49), (15.07, 18.19)]
```

Fig.6. The average distance of iris center point (x, y) fluctuations for frequent iris movement case

From Fig.6 we observe that in case of frequent iris movement the average distance of iris center point in consecutive videos vary with a range of 20. For x coordinates the average distance of iris center point in consecutive videos varies from 10 to 30 and the y coordinates of the average distance of iris center point in consecutive videos vary from 7 to 27.

```
Center Coordinates: [(11.53, 9.07), (11.96, 8.964),

(12.50, 7.53), (6.18, 5.98), (10.86, 6.49),

(12.18, 9.75), (7.68, 6.69), (13.04, 12.152),

(7.72, 5.71), (11.725, 7.33), (11.47, 6.21),

(7.55, 4.79), (5.32, 4.79), (6.83, 4.69),

(6.23, 5.21), (8.19, 7.14), (10.36, 9.44),

(8.62, 5.88), (15.93, 8.49), (2.68, 8.67),

(3.56, 7.88), (4.14, 10.24), (7.12, 9.09),

(9.18, 4.02), (5.90,1.44), (10.90, 8.82),

(8.88, 3.15), (6.12, 1.56), (8.91, 1.89),

(7.29, 2.59)]
```

Fig.7. The average distance of iris center point (x, y) fluctuations for non-frequent iris movement case

Fig.7 shows the average distance of iris center point movement of (x, y) coordinates in our dataset for a non-avoiding and normal iris movement case. It represents the variation of average distance of iris center point of both x and y coordinates for the case of non-frequent iris movement.

From Fig.7 we observe that the x coordinates of the average distance of the iris center point in consecutive videos vary from 2 to 15 with a range of 13 and the y coordinates of the average distance of the iris center point in consecutive videos vary from 1 to 9 with a range of 8.

From these two above instances, discrepancy in the distribution of the average distance of the iris center point coordinates variance can be seen. In avoiding cases and frequent iris movement cases, the range is much higher than that of non-avoiding cases and normal iris movement cases. Upon carefully reviewing the 50 videos, which involve both regular iris movement and non-frequent iris motion images, a threshold value is set for the average distance of both x and y coordinates of the iris center point. Taking into consideration the threshold value, a video is identified as "frequent iris movement" and "avoid case" if the average distance value of the iris center point for either x or y of the video coordinates remains higher than the threshold value. On the contrary, a video is classified as "normal iris movement" and "non-avoiding case" if the average distance value of the iris center point for either x or y of the video coordinates remains lower than the threshold value. This threshold value was used in our further experiments to successfully identify suspicious and normal human behavior on the basis of frequent iris movement criterion.

4. Experimental Results and Discussion

To explain our proposed framework, a comparative study of our proposed unusual and irregular observer identification program with other implementations is seen after our dataset is used as a test data. Our analysis is performed in two steps. 100 frames per video were considered to be monitoring the outcome of abnormal behavioral detection in the first experiment, while in the second experiment, 150-200 frames per video were taken into account for observing the outcome of abnormal behavioral detection. Our analysis is performed in three major phases. By using MTCNN, we have attained 98 percent precision of eye detection where our dataset has been used. The outcome of the previous phase is used to identify the iris of the eye pictures. 73% accuracy is achieved for our dataset when CHT is used to detect iris. In the last phase of abnormal behavior detection using frequent iris movements, we achieved 94% accuracy by using the TRM model and our dataset. This significant result has been achieved in terms of frequent iris movement and TRM model. Later, our results have been compared with other established research that concentrated on accuracy, consistency, and recall. Finally, our work is combined in the relevant field of object detection with existing work. Parts A to D give a detailed review of our study outcomes and discussion which has been reflected in Fig. 8.
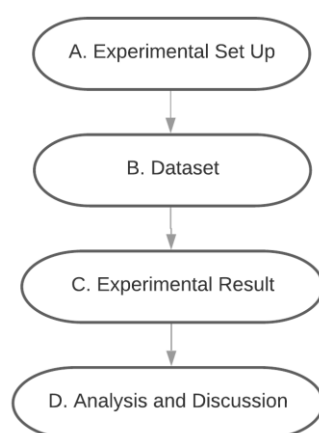

Fig.8. Four levels of study findings

Fig.9. Iris detection using CHT

A. Experimental Set Up

At first, the movements of iris were detected and observed from the eye images that we found after the eyes were detected using our dataset. Then the TRM method was used to find out the abnormal behavior in the iris movement. Every one of the implementations has been performed with OpenCV on the python platform. Fig.9 displays the reference images from our dataset where both irises from the images have been successfully determined.

After detecting both iris from eye images the movements of iris were identified by observing the coordinates of iris center points for consecutive frames in a video. Here, one iris was considered for further process for the simplification of our model and implementation. Besides, because a person's both iris move in the same manner always while the movement of the iris, one iris is taken to find out the abnormal behavior of an observer.

Iris movement is plotted in Fig.10 to show the variance between iris center points for our dataset. It elucidates that iris are continuously moving from its initial point and it will lead us to our morphological model to detect abnormal and suspicious frequent iris movement.
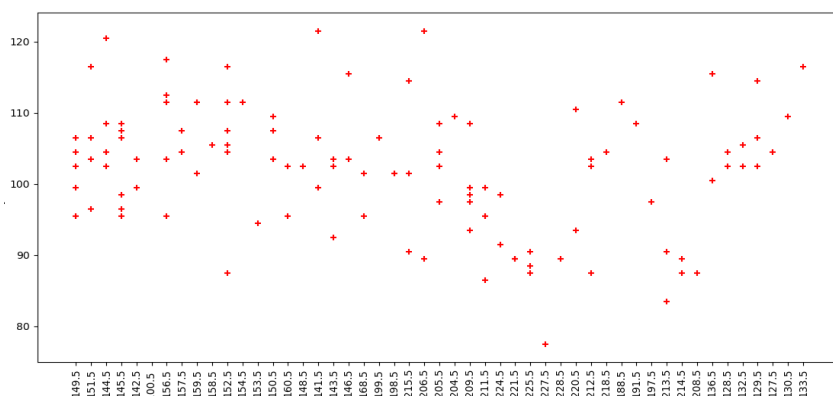
Fig.10. Iris coordinates of a video containing sequential iris movement

Then using our morphological model (TRM) the average distance of iris center point for each video was calculated and a threshold was set and while testing the rest of our dataset we observed frequent iris movements of different persons who were trying to avoid eye contact whereas some normal iris movements are detected where none were trying to avoid eye contact and were normal.

Fig.11. Frequent movement of one iris in successive frames for a person who wishes to escape eye contact

In Fig.11 it is observed that an individual tries to prevent eye contact and one iris is constantly moving which points to an abnormal and suspicious characteristic of humans.

B. Dataset

Due to the uncertainty of anomalies, the context in which the incidents occur, the lack of appropriate sizes of the abnormal ground truth training data, and other factors relevant to changes in other environments and lighting conditions, the detection of abnormal characteristics from a video are very difficult. Since we do not find any appropriate datasets available particularly to verify our research, we have generated our dataset. Hardly any accessible dataset contains videos that often allow participants to transfer iris from left to right or from up or down or conversely. We generated a dataset called "TRM Iris Dataset" for this reason which includes a total of 100 videos with about 18265 frames. The participants frequently switch their irises in 44 videos, and 56 videos either don't switch or, in a variety of words, are in a constant position. For the first 100 (2-4 seconds) and 200 (6-10 seconds) frames of each video, we conducted our experiment.

C. Experimental Results

Based on the threshold value we inspected the abnormal and suspicious behavior after the evaluation of frequent iris movement and avoiding eye contact by using the TRM method and TRM Iris Dataset.
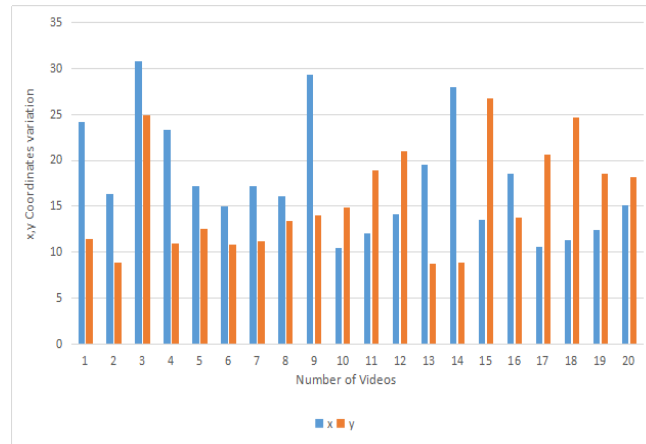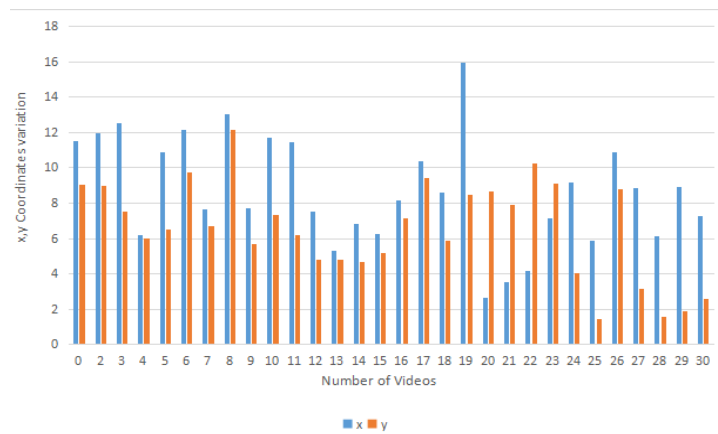
Fig.12. Histogram of the average distance of iris center point (x, y) fluctuations for abnormal iris movement and avoiding eye contact case

In Fig.12 we have illustrated the average distance of iris center point (x, y) fluctuations for frequent iris movement cases as a histogram which eventually indicates to the suspicious observer who tries to evade contact with the eyes.

In this case, since x and y coordinates are higher than the threshold for the average iris center point for most videos, the videos were described in our system as "frequent iris movement" and "eye contact avoidance".

Fig.13. Histogram of the average distance of iris center point (x, y) Fluctuations for normal iris movement and non-avoiding eye contact case.

Similarly, In Fig.13 we have illustrated the average distance of iris center point (x, y) fluctuations for non-frequent iris movement cases as a histogram which eventually indicates to the normal observer who does not try to evade contact with the eyes.

In this case, the videos are referenced as "non-frequent iris movement" and "non-avoiding eye contact" in our system, since x and y coordinates are less than the threshold of the average iris center point for most videos.

By observing the above two cases, a feasible threshold was taken for further experiments and evaluation tests.

The experiments were conducted in two parts. Firstly, the confusion matrix of abnormal and suspicious iris movement was calculated by considering the 100 frames (2-4 seconds) from each video from the TRM Iris Dataset. It was drawn in Table 1. Here, the confusion matrix was generated using 50 videos where the first 2-4 seconds were counted from each video.

Table 1. Confusion Matrix for the First 100 Frames

	Positive	Negative
Predicted Positive	18	4
Predicted Negative	2	26

We can observe here, for the first 100 frames or the first two to four seconds, that our model identified 18 videos containing frequent iris movements and 26 videos containing normal iris movements.

Table 2. Calculation of Confusion Matrix for First 100 Frames

Measure	Value (%)
Precision	81.82
Recall	90.00
Specificity	86.67
F-Measure	85.71
Accuracy	88.00
Error Rate (ERR)	0.12

The accuracy rate was 88%. Where the precision was 81.82% and recall was 90%. The specificity rate was 86.67% and finally, the F-1 Score was 85.71%. However, the error rate of this experiment was 0.12.

Additionally, the first experiment could point out a suspicious or normal video with lower accuracy.

Secondly, we have calculated the confusion matrix of abnormal and suspicious iris movement by considering the 150-200 frames (6-10 seconds) from each video from TRM Iris Dataset. It was drawn in Table 3. Here, the confusion matrix was also generated using 50 videos where the first 6-10 seconds were counted from each video.

Table 3. reveals that, in the first 150-200 frames or the first 6-10 seconds, our proposed model identified 19 videos containing recurrent iris movements and 28 videos containing non frequent iris motions correctly which is significantly better than the one based on the first 100 frames per video. Here the True positive rate and True negative rate are 19 and 28 consecutively. And the False Positive and False Negative rate are 2 and 1 one after another.

Table 3. Confusion Matrix of the First 150-200 Frames

	Positive	Negative
Predicted Positive	19	2
Predicted Negative	1	28

From Table 4. We observed that the accuracy rate was 94%, precision was 90.48% and recall was 95%. The specificity rate was 93.33% and finally, the F-1 Score was 92.68%. In addition, the error rate of the second experiment was 0.06.

Table 4. Calculation of Confusion Matrix for First 150-200 Frames

Measure	Value (%)
Precision	90.48
Recall	95.00
Specificity	93.33
F-Measure	92.68
Accuracy	94.00
Error Rate (ERR)	0.06

Hence, it is evident that the error range varies from 0.06 to 0.12 according to the experiments. Additionally, the second experiment could point out a suspicious or normal video with higher accuracy.

D. Analysis and Discussion

In this segment, we compared our research with previous works on dubious observer detection. In Table 5. We discussed the confusion matrix of the previous morphological method and the TRM method.

Table 5. Comparison of confusion matrix of TRM and other morphological method for identifying frequent iris movement and suspicious activities

Comparison of Methods	Recall (%)	Precision (%)	F-Measure (%)	Accuracy (%)
TRM Method	95.00	90.48	92.68	94.00
Morphological Method [21]	93.18	93.18	93.18	93.02

Here we can see our recall rate is higher than the existing morphological method whereas the precision rate is lower than the previous morphological method. The F-1 score is identical. But our accuracy rate is slightly higher than the existing method.

A comparison between our two experiments to detect abnormal and frequent iris movement for 100 frames and 150-200 frames has been illustrated in Fig.14. This figure perfectly describes the significant change in result in terms of identifying abnormal and frequent iris movement for our two different experiments which were considered by taking the first 100 frames per video and the first 150-200 frames per video consecutively.
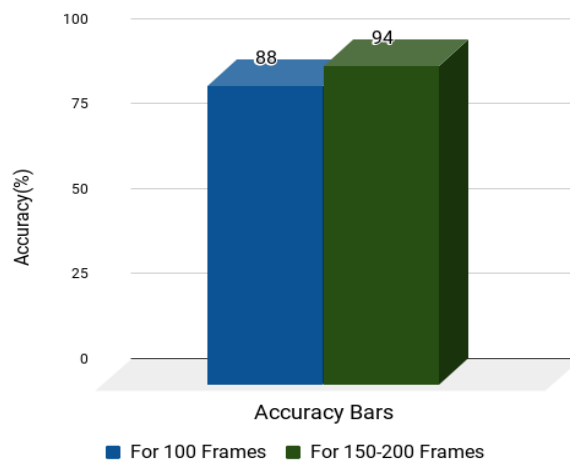

Fig.14. Comparison between our two experiments for 100 frames and 150-200 frames.

In Fig.15 a detailed comparison has been observed between our TRM method with the previous morphological method in terms of identifying frequent iris movement.
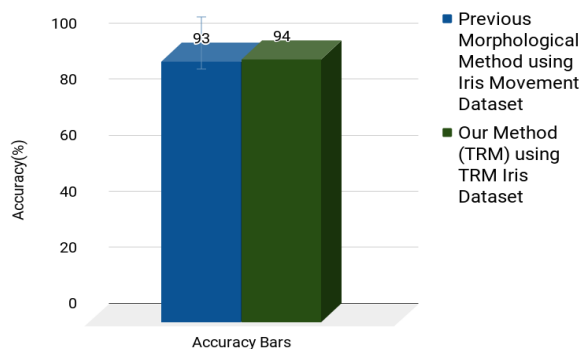

Fig.15. Comparison between existing method and TRM method in terms of identifying frequent iris movement and suspicious activities.

After evaluating the outcome for our two experiments we can come to the conclusion that the result for 150-200 frames is better than the experiment for 100 frames. And it indicates to detect someone's behavior as abnormal and frequent iris movement we should consider 6-10 seconds of video. So, we need to observe a particular person for about 6-10 seconds to get the desired output.

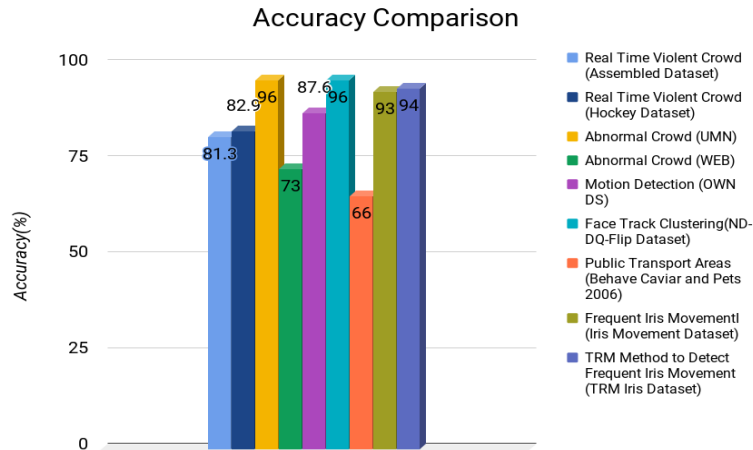Finally, the contrast of our findings with known works was demonstrated in Fig.16.

Fig.16. Comparison of our experiment with existing works

In Fig.16, the first two labels reflect the real-time identification of aggressive crowd activity dependent on the Assembled database and the Hockey database, where the accuracy was 81.3 percent and 82.9 percent successively for the Assembled database and the Hockey database [24]. The labels UMN and WEB constitute 96 percent of the UMN dataset and 73 percent of the web data of work conducted by Mehran and others [36]. They developed a new approach for detecting and evaluating odd behaviors in crowd videos based on the concept of visual flow and using the social force model. The mark Motion Detector represents Yasin et al.'s accuracy in the research [37]. Their human skepticism and irrational motion have been categorized based on instant invariants with an accuracy of 87.6 percent on their own data collection. The segmentation mark of Face Track identifies Barr et al's attempts to identify suspicious observers on the ND-QO-Flip with a precision of 96 percent using their face track grouping method [22].

The Fig.16 mark for the Public Transportation Field shows the precision of the identification work of Elhamod et al. The BEHAVE CAVIAR and PETS 2006 data collection has achieved an accuracy of 66 percent [38]. The second last bar represents accuracy [20] with a data set of its own. They also reached an accuracy of 93 percent. The experiment was carried out with the first 150 frames of their dataset. Finally, the last bar represents a new morphological TRM method for the identification of behavioral abnormalities based on frequent iris movements where 94 percent accuracy was achieved. The experiment was carried out on the basis of 150-200 frames.

Evaluating all the results of the existing works, we can conclude that we have successfully detected iris from the eye whether it is a suspicious person's iris or not based on specific criteria. In this regard, we have used our own generated method called TRM and achieved an accuracy of 94%.

5. Conclusion

In terms of identifying a suspect individual, the regular monitoring of iris activity is one of the most significant ways of identifying abnormal human behavior. The detection of the dubious observer has several important connections with law enforcement personnel. Not only can fraud be stopped and identified in shopping centres or elsewhere, but even violence or events in public spaces can be avoided. Saving human lives and a wide variety of goods and resources will also be possible. Presently, no work has been conducted using TRM method, and so this article will pave a new way to detect dubious actions of a person. This research described the methodology and techniques used for the recognition of activities, and also their data set and accuracy. In comparison, this work identified two basic concepts for the definition of an abnormal observer to reflect one of the features of a questionable observer, while some other odd acts and actions may categorize one individual as a doubtful observer. Frequent iris movements and avoiding eye contact have been concentrated here as unusual and suspicious actions. We have successfully achieved an accuracy of 94 percent because of these specific actions. The creation of detailed datasets in various circumstances to validate the proposed context is one of the most challenging tasks in this research. The datasets used in this research include a different environment form, lighting, facial expression, and eye movement for testing the proposed structure. Taking into account the vision-based tracking function, all sensor-based response identification limitations were resolved. This research has not only introduced a new and significant method (TRM), but has also constructed a new data set for the

validation of frequent iris movements, which is the first and foremost approach to the eventual detection of dubious observers.

References

[1] A. Wiliem, V. Madasu, W. Boles, and P. Yarlagadda, "A suspicious behaviour detection using a context space model for smart surveillance systems," *Computer Vision and Image Understanding*, vol. 116, no. 2, pp. 194–209, 2012.

[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] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] Matsui, Shinya, Nakamasa Inoue, Yuko Akagi, Goshu Nagino, and Koichi Shinoda. "User adaptation of convolutional neural networks for human activity recognition." In Signal Processing Conference (EUSIPCO), 2017 25th European, pp. 753-757. IEEE, 2017.

[12] 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).

[13] Sunkad, Zubin A. "Feature Selection and Hyperparameter Optimization of SVM for Human Activity Recognition." In Soft Computing & Machine Intelligence (ISCMI), 2016 3rd International Conference on, pp. 104-109. IEEE, 2016.

[14] 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.

[15] 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.

[16] 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.

[17] Chen, Wen-Hui, Carlos André 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.

[18] 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.

[19] 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.

[20] D. M. Anisuzzaman and A. F. M. S. Saif, "Efficient Framework Using Morphological Modeling for Frequent Iris Movement Investigation towards Questionable Observer Detection," *International Journal of Image, Graphics and Signal Processing*, vol. 10, no. 11, pp. 28–37, Aug. 2018.

[21] 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.

[22] 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.

[23] D. M. and A. F., "A Study of Activity Recognition and Questionable Observer Detection," *International Journal of Computer Applications*, vol. 182, no. 15, pp. 35–42, 2018.

[24] 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.

[25] K. Pawar and V. Attar, "Deep learning approaches for video-based anomalous activity detection," *World Wide Web*, vol. 22, no. 2, pp. 571–601, Mar. 2018.

[26] V. J. Kok, M. K. Lim, and C. S. Chan, "Crowd behavior analysis: A review where physics meets biology," *Neurocomputing*, vol. 177, pp. 342–362, 2016.

[27] B. Yogameena and C. Nagananthini, "Computer vision based crowd disaster avoidance system: A survey," *International Journal of Disaster Risk Reduction*, vol. 22, pp. 95–129, 2017.

[28] M. S. Zitouni, H. Bhaskar, J. Dias, and M. Al-Mualla, "Advances and trends in visual crowd analysis: A systematic survey and evaluation of crowd modelling techniques," *Neurocomputing*, vol. 186, pp. 139–159, 2016.

[29] 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.

[30] N. Cherabit, F. Z. Chelali, and A. Djeradi, "Circular Hough Transform for Iris localization," Science and Technology, vol. 2, no. 5, pp. 114–121, Jan. 2012.

[31] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks," IEEE Signal Processing Letters, vol. 23, no. 10, pp. 1499–1503, 2016.

[32] T. Xiang and S. Gong, "Incremental and adaptive abnormal behaviour detection," *Computer Vision and Image Understanding*, vol. 111, no. 1, pp. 59–73, 2008.

[33] B. Yao and L. Fei-Fei, "Modeling mutual context of object and human pose in human-object interaction activities," *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010.

[34] T. Lan, Y. Wang, W. Yang, S. N. Robinovitch, and G. Mori, "Discriminative Latent Models for Recognizing Contextual Group Activities," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 8, pp. 1549–1562, 2012.

[35] May Phyo Aung, Soe Yu Maw, "IRIS SEGMENTATION SYSTEM USING THE HOUGH TRANSFORM," 2012.

[36] 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.

[37] Yasin, Hashim, and Shoab Ahmad Khan. "Moment invariants based on 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.

[38] 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.

Authors' Profiles

Md. Minhaz Ur Rahman, male, is a Software Engineer at ICT Center, Ahsanullah University of Science and Technology, Dhaka, Bangladesh. He is currently working toward the development of the IUMS system of Ahsanullah University of Science and Technology. His research priorities include computer vision, machine learning, natural language processing, and image processing.

Mahmudul Hasan Robin is a Computer Science Teacher at Mastermind English Medium School, Dhaka, Bangladesh. He previously worked as an Artificial Intelligence Engineer at Cwork Microjob Limited, Dhaka, Bangladesh. He has received his BSc from Ahsanullah University of Science and Technology, Dhaka, in 2020. His research priorities include computer vision, machine learning, image processing, and bioinformatics.

Abu Mohammad Taief, male, is working as an Assistant Project Manager at Bit Mascot (Pvt.) Ltd., Dhaka, Bangladesh. He also worked as a Project Manager at ShobHobe, Dhaka, Bangladesh. His research priorities include computer vision, machine learning, and image processing.