## A Study of Activity Recognition and Questionable Observer Detection

#### D. M. Anisuzzaman

Department of Computer Science and Engineering, Ahsanullah University of Science and Technology, Dhaka, Bangladesh

## ABSTRACT

Detecting some specific suspicious activities is a core part of questionable observer detection. Activity recognition and questionable observer detection are important fields of study in artificial intelligent system building and computer vision. Questionable observer detection will not only reduce the workload of thousands of workers but also can prevent crimes. This research investigated existing methods and presented a framework to detect a specific activity. This work has shown the tabular study of algorithms, detected actions, datasets used and accuracy for each type of activity recognitions. This research has also proposed a framework to detect a questionable observer from video on basis of a specific action named avoiding eye contact. The algorithms to detect face, eyes and irises are also described here. This research has also proposed that determining the location of iris in consecutive frames can detect if a person is trying to avoid eye contact.

## Keywords

Iris detection, Suspicious activity detection, Activity recognition, Questionable observer detection.

#### **1. INTRODUCTION**

Now a day's surveillance video monitoring has become an important issue of interest. Not only it 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. But as the old saying goes: "prevention is better than cure"; it will be a much peaceful scenario if one can detect a crime before its occurring rather than take action after its happening. This research is focused on questionable observer detection. It is hard enough to say someone questionable for a human being, so it will be a lot harder to make a system that can say someone questionable intelligently. But if a system can be build that can detect questionable observer intelligently; it will be much more efficient than human as surveillance video monitoring is a boring and tiresome work and a human can scrutiny one person in particular or in some cases several persons; whether an intelligent system with this capacity can monitor and detect all the questionable observers in its range. To say someone questionable (in some cases criminal) some certain activities have to be detected on basis of which one can be said questionable. Some questionable activities are: 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.

A. F. M. Saifuddin Saif

Department of Computer Science, American International University-Bangladesh, Dhaka, Bangladesh

Most of the human activity recognition (HAR) system considers dataset collected from wearable devices that take certain measurements to detect an activity. Two wearable modules were attached on wrist and ankle to detect the acceleration and angular velocities generated from activity movement in [4]. In [8] wireless sensing devices are mounted on different parts of the human body which periodically transmit biophysical streams of data to a mobile device. Walking, running, stationary state, riding a train, riding a car or a bus, and cycling was recorded by using an accelerometer, a magnetometer and a gyroscope [14]. Volunteer's (with an age range of 19 to 48) smart phone sensor data was collected, who performed the following activities: standing, sitting, laying, walking, walking downstairs and walking upstairs and a protocol was also designed for data collection [15]. In [16] authors used a dataset that consists of accelerometer and gyroscope value. Authors used a dataset named WARD where five motion sensors were placed at five different body locations (two on the wrists, one on the waist and two on the ankles, respectively) and for each motion sensor, a two-axis gyroscope and a three-axis accelerometer are used to sense the human activity signals [20]. Accelerometer data was collected and used for human activity recognition in [23]. Accelerometers and gyroscopes data were used and processed for human activity recognition in [26]. Authors consider a Body Sensor Network (BSN) architecture, comprised of wearable sensing platforms, deployed on several body locations and record a number of sensing modalities (e.g., acceleration, angular velocity) in [27]. UCI HAR dataset was used in [28], which consists of accelerometer and gyroscope values collected from a smart phone worn by 30 volunteers.

The other HAR systems considers vision based dataset or non wearable sensor based datasets for activity recognition. From a web camera the video is captured and then converted into sequence of frames having a fixed size of 240 in [3]. Spatiotemporal features (STIPs) are used to determine activities in [5]. In [10] authors used two datasets consist of over 100 videos and 51 types of actions. Video taken in a controlled environment and depth video is used for the activity recognition in [11] and [12] respectively. Channel State Information (CSI) data is used in [18], which is developed using the WiFi networks. A dataset which was built using Microsoft Kinect RGBD is used in [22].

All the aforementioned datasets go through some approaches or methods or algorithms to detect an activity correctly. Some popular approaches and algorithms for activity recognition are: Convolutional Neural Network (CNN) [12] [14], Long-Short Term Memory (LSTM) neural network [15] [22] [26], Support Vector Machine (SVM) [16] etc.

Other than activity recognition some people tried to detect suspicious activity [1], [13], [17] from surveillance camera. There approaches are completely vision based without having any influence of wearable devices. Facial expression and human behavior detection is also related with this research as some specific facial expression and behavior can indicate questionable activity. These works are also completely computer vision based [6], [19], [21]. Some work has been done to detect violent crowd behavior [7], [9].

As this research is working mainly with surveillance video to detect questionable observer, the camera also plays an important role in the questionable observer detection. Some work has been done [24], [25] to capture better video which will help us to detect questionable observer intelligently.

In this paper, a state survey has been presented on action recognition. Section 2 contains the current challenges of action recognition. Section 3 discusses Previous Research on Methods and Algorithms. Section 4 contains the Research methodology where discusses proposed frameworks, expected result. Finally, the conclusion of this research is drawn in section 5.

## 2. CURRENT CHALLENGES

This section discuss about the main challenging factors in questionable observer detection and as a sub class of this domain also the challenges of activity recognition, suspicious activity detection, facial expression and human behavior detection and the detection of violent crowd behavior.

The activity recognition can be done by using wearable devices dataset, through video dataset or using some other wireless devices dataset that does not need to be worn as aforementioned. Table 1 briefly describes the challenges and limitations in activity recognition according to dataset.

Table 1. Challenges and limitations according to dataset

Datasets	Challenges And Limitations				
	Have to wear gadgets on the body which is highly impractical in real world [4].				
With Weird Devices	Not real time detection [4].				
	Users wear their smart phones quite differently (some hold them in one hand, others put them in their pockets, or in their bags etc.). Such differences in holding styles make sensor outputs significantly different, and thus, would degrade the HAR performance [14].				
	The variable values can change according to the smartphone's movement, which in turn corresponds to the user's action. However, since the acceleration data include a rotation component, such rotation component might interfere with the correct activity recognition [23].				
	Sensor readings may contain incorrect or noisy data [26].				
	The time series of smart phone sensor data cannot be directly used for shallow models to identify different human activities [15].				
	Some similar activities like "climbing upstairs" and "climbing downstairs" are hard to differentiate [16].				
	Wearable sensors require extra costs for hardware and are intrusive for users [15].				
Computer	Camera based approaches have fundamental				

Datasets	Challenges And Limitations			
Vision Based and	limitations: the line-of-sight detection, good illumination, and potential privacy leakage [18].			
Non Waerable Devices	Sometime it is critical to obtain the accurate outline of moving object in a video [3].			
	In some experiments video background has to be known [3].			
	In some situations, finding trajectories from videos might not be discriminant enough [11].			
	CSI often request specific facilities: GPS clock, RFID readers to eliminate the impacts of ambient noises [18].			
	CSI is extremely sensitive to the channel variance and position changes. If the propagation path of WiFi signals changes merely half wavelength, the received CSI would be severely irrelevant [18].			
	The variant gains make different CSI values once the transmitting or receiving devices are changed to detect the same activity [18].			
	Some experiments are only realized in indoors environment because skeleton data captured by Microsoft kinect is restricted in indoor scene [22].			
	Some experiments need expensive GPU for computing [22].			

These data are passed through some methods and algorithms to correctly recognize an activity. Some challenges and limitations of some popular methods and algorithms are briefly described in Table 2.

and argor tunns	
Method / Algorithm	Challenges And Limitations
	DNN needs much more training time than typical neural networks [12].
	Deep learning based approaches often suffer from noise effect and limited size of data, leading to an unsatisfactory performance [15].
CNN and LSTM	A deep neural network with a large number of parameters relies on a large number of training samples to optimize [15].
	If the number of hidden units is too small, the expression ability of model is not enough. If too big, the model complexity is too high and the generalization performance will descend severely [22].
SVM	Choosing a good kernel function is not easy.
	It needs long training time on large data sets.

Table 2. Challe	enges and	limitations	according	to methods
and algorithms				

The works of suspicious activity detection, violent crowd behavior detection and facial expression and behavior detection also have some challenges and limitations of their own respective fields. The challenges and limitations of these methods are shown in Table 3.

	Table 3. C	Challen	ges and	limitations	of suspici	ous activ	ity
	detection,	violen	t crowd	behavior	detection	and fac	ial
	expression	detect	ion				
Г	36 (1 1	1					

Algorithm Challenges And Limitations	
Suspicious Activity Detection	Needed fixed camera shoots from same angle and distance [1].
	With the increase of training dataset, the success rate will be better but the time required would increase as well [13].
	Variations in pose, illumination, and facial expression throughout a single video and between different videos can affect face appearance and, hence, complicate questionable observer detection as well [17].
	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].
Violent Crowd Behavior Detection	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].
Facial Expression And Behavior Detection	Face should be illuminated by approximately equal luminance over the face for the better result [6].
	In order a facial expression to be detected, face location and area must be known [19].

The research area of questionable observer detection not only posses all the aforementioned challenges and limitations but also contains some big challenges of its own. A questionable observer in other word a suspect will not wear any sensor in his own will, so all the approaches that uses data from wearable devices and sensors are useless for this task. In this respect all approaches using smart phone data is also useless. So, most of the works of this research have to be done from video data. To build a real-time detection system that can analyze all the questionable behaviors in parallel and classify a suspect on behalf of this analysis is really a challenging task.

## 3. PREVIOUS RESEARCH ON METHODS AND ALGORITHMS

Some specific activity recognition (e.g. repetitive touching of face, excessive head-turning, avoids eye contact etc.), suspicious activity detection, violent crowd behavior detection and facial expression and behavior detection are the core field of interest for questionable observer detection. Many methods and algorithms have been used for these detection processes. This section briefly describes these methods and algorithms and their detected activities and also their accuracy of detection.

An inertial-sensing-based human activity classification algorithm has been used for the recognition of 10 activities with a classification rate of 90.5%. The main advantage of using this method is that illumination variations and background change will not affect the system. Convolutional Neural Networks (CNN) has been used for activity recognition in [14] with an accuracy of 88.2%. The main advantage of using deep learning method is that it includes feature extraction process in its modeling. Long Short-Term Memory (LSTM) network has been used for activity recognition in [15]. The deep LSTM network is able to learn features from raw sensory data to encode temporal dependencies and knowledge distilling mechanism can improve the recognition accuracy of deep LSTM. Grid search algorithm and SVM classifier was used with an impressive accuracy of 99% in [16]. As random projection based method can optimally, linearly and quickly reconstruct the test sample using as few training samples as possible; an accelerated sparse representation classification method based on random projection and k-nearest neighbor classifier has been used in [20]. One of most significant reason to use this method is that, the runtime of this method is significantly reduced in comparison with the traditional sparse representation classification methods. To fill the missing data and exploit data redundancies matrix and tensor completion techniques have been used in [27]. Feature selection based on Choquet integral is used in [28], as this approach overcomes the classical feature selection methods that are generally linear and sequential and they do not consider existing dependencies and interactions among activities (classes). This method also improves the learning time compared to the others feature selection methods. A brief description of all these methods and algorithms that uses wearable sensor data as input; along with their dataset and accuracy and detected activities are shown in Table 4.

Table 4. Methods / Algorithms using wearable sensor dataset

Method / Algorithm	Detected Activities	Used Dataset	Accura cy
An inertial- sensing-based human activity classification algorithm [4].	Walk, Run, Upstairs, Downstairs, Stand Up and Squat, Drink, Take elevator, Still, Sit, and Lie	Their own developed dataset collected in the indoor environments.	90.5%
A novel online architecture for implementing feature selection on mobile devices [8].	Stand, Sit, Sit & Talk, Walk, Walk & Talk, Climb, Climb & Talk.	Their own database (FORTH- TRACE.db); from the wearable sensors to the mobile device.	84%
A user adaptation method using Learning Hidden Unit Contributions (LHUC) for Convolutional Neural Networks (CNN) [14].	Walking (Walk), Running (Run), Stationary State (Still), Riding a Train (Train), Riding a Car or a Bus (Car/Bus) And Cycling (Cycle)	Collected their own evaluation data.	88.2%

Method / Algorithm	Detected Activities	Used Dataset	Accura cy
An efficient shallow network is utilized to assist a deep Long Short-Term Memory (LSTM) network [15].	Standing, Sitting, Laying, Walking, Downstairs and Walking Upstairs	collected their own dataset by using smart phone sensor.	97.7%
Grid search algorithm and SVM classifier [16].	Walking, Climbing, Jumping, Sitting, Standing	USC-HAD database.	99%
An accelerated sparse representation classification method based on random projection and k-nearest neighbor [20].	Stand, Sit, Walk Forward, Turn Left, Turn Right, Go Upstairs, Go Downstairs, Jog, Jump.	Wearable Action Recognition Database (WARD).	92.56%
Long short- term memory (LSTM) recurrent neural networks [26].	Lying, Sitting, Standing, Walking, Running, Watching TV, Computer work, Ascending stairs, Descending stairs, Vacuum cleaning, Ironing, Folding laundry, House cleaning.	PAMAP2 (Physical Activity Monitoring) dataset	82.6%

Some methods and algorithms use video dataset or some wireless dataset where the user does not need to wear any device in his body, for activity recognition. Frame subtraction, background subtraction and optical flow method have been used for activity recognition, on a dataset made by the authors that contains real time videos in [3]. Minimum Redundancy Maximum Relevance (mRMR) feature selection method has been used to reduce the number of features in [2]. Hierarchical spatio-temporal model (HSTM) [5], Combinatorial Deep Belief Networks [10], Trajectory extraction methods: Lucas-Kanade (LK), Farnback (FB) and Interest Point tracking (IP), topic model in Channel State Information (CSI) based system [18] and Three-layer Long Short Term Memory (LSTM) is also used for activity recognition. A brief description of these methods and algorithms has been given in Table 5.

Table 5. Methods /	Algorithms using	video	dataset or no	n-
wearable sensor dat	taset			

wearable senso	r dataset	<b>T</b> T <b>1</b>	•
Algorithm /	Activities	Used Dataset	Accuracy
Hierarchical spatio-temporal model (HSTM) [5].	One-person action, Human- human interactional activity and Human-object interactional activity	KTH, UCF, CASIA and Gupta	KTH: 98% UCF:89.3 3% CASIA: 95.24% Gupta: 96.30%
Combinatorial Deep Belief Networks [10].	More than 51 types of actions.	HMDB51 and Hollywood2	HMDB51 : 80.48% Hollywoo d2: 91.21%
Trajectory extraction methods: Lucas- Kanade (LK), Farnback (FB) and Interest Point tracking (IP) [11]. A novel method from depth videos using robust spatiotemporal features with convolutional neural network [12].	Hand Clapping, Jumping, Skipping, Picking up, Pushing, Running, Sitting down, Standing up, Walking and Jumping Jack. Left Leg Moving, Right Leg Moving, Both Hand Waving, Right Hand Waving, Sitting-Down, and Standing- Up.	Created a stereo dataset that involves 11 actors performing many different activities. Their defined database and MSRC-12 Gesture database.	LK: 96.69% FB: 97.35% IP: 80.13% Their defined database: 98.17% MSRC-12 Gesture database: 98.27%
Three-layer Long Short Term Memory (LSTM),	Swipe Left and Right, Wave, Clap, Throw, Arm Cross, Basketball Shoot, Draw X, Draw Circle, Draw Triangle, Bowling, Front Boxing, Baseball Swing From Right, Tennis Forehand Swing, Arm Curl, Tennis Serve, Push, Knock, Catch, Pick Up And Throw, Jogging, Walking, Sit To Stand, Stand To Sit, Forward Lunge, and Squat.	UTD multimodal human activity dataset.	Experime nt #1: 95.31% Experime nt #2: 88.58%

Method /	Detected	Used	Accuracy
Algorithm	Activities	Dataset	
Use topic model in Channel State Information (CSI) based system [18].	Push, Pull, Drag, Circle, Kick and Squat.	Collected their own dataset by asking a number of volunteers to perform a set of actions.	> 90%

Some methods and algorithms are also used for facial expression and behavior recognition, violence in a crowd detection and suspicious activity detection. A brief description of all these methods and algorithms are given in Table 6.

 Table 6. Methods / Algorithms for human behavior, suspicious and violent activity detection

Method / Algorithm	Detection	Used Dataset	Detection Rate
Viola-Jones AdaBoost method, PCA algorithm and a novel algorithm using Kalman filter estimation [6].	Sad, Surprise, Fear, Anger and Disgust emotions.	own developed dataset by placing camera at minimum two places (entry and exit gate).	Threshold set to 60%
A Proposed Algorithm by clustering sequences of face images [17].	Questionable Observer.	ND-QO-Flip dataset comprised by fourteen 25-59 second crowd video clips.	False Positive Rate: 0.056 False Negative Rate: 0.00
A new Moment Invariant based Classifier (MIBC) has been proposed [13].	Bending Down, Gun Shot, Jumping Up, Kicking Front and Punching Forward.	Own developed dataset collected for each specified action at different backgrounds.	87.6%
A novel method that considers statistics of flow-vector magnitudes and linear SVM [7].	Violent Crowd Behavior.	Own assembled database, Hockey Database, and ASLAN	Their assembled database 81.30 % Hockey database: 82.90 % ASLAN database
		Database.	56.57 %
A novel method using Social Force model	Abnormal Crowd Behavior	UMN dataset and	UMN dataset:
[9].	Denuvior.	Web dataset.	96 %
			Web dataset:

Method Algorithm	/	Detection	Used Dataset	Detection Rate
				73%

Only one work is done on questionable observer detection [17], where the whole work is done on the basis of: a person who appears in most of the different videos in other words who is found in most of their video cluster is questionable. Some suspicious activities like: bending down, gunshot, jumping up, kicking front and punching forward is also detected in [1] and [13]. But the main focus of this research is to detect a person as questionable if he/she performs some specific activities like: repetitive head turning, blinking eyes a lot, excessive fidgeting of legs etc.

## 4. RESEARCH METHODOLOGY

Questionable observer detection is a huge process. It includes several activity detection and computation of all these activities in parallel. This section is focused about one specific questionable or suspicious activity called avoiding eye contact detection. A suspicious person tries to avoid his eye contact with other people and also move his eye frequently with hesitation. To the best of our knowledge no work had been done yet to detect this specific activity. This activity detection is done in 6 steps and a block diagram of this process is shown in Fig 1.

## 4.1 Proposed Framework

After collecting the data from the source (surveillance video or YouTube or any other source), all data will go through some processing to build the dataset. The face will be detected by using neural network from each frame of every video. Then the eyes will be detected and from that this work will detect the locations of both irises. After that this work will detect the position of two irises and will perform some calculation to detect the specific activity named "avoiding eye contact". The whole process is described elaborately in the following sections (4.2 to 4.6).

## 4.2 Video Processing

Each video collected from surveillance video or YouTube or any other source will consist of 50 frames and will be organized with a fixed height and width. These frames are the input for the next step in the neural network.



Fig 1: A block diagram of "Avoiding Eye Contact" activity recognition

## 4.3 Face Detection using Neural Network

Many approaches have been used for face detection. There are mainly two ways for face detection, one is feature based approach and the other one is image based approach. Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Factor Analysis (FA), Support Vector Machine (SVM), Neural Network (NN) all are image based approaches [29]. All these methods has pros and cons; but this work is going to use Neural Network for the face detection part as it provides the highest accuracy for the face recognition [30].

# 4.4 Eyes Detection using Physiological properties, Dynamics and Appearance

Structured infrared lighting is used to capture the physiological properties of eyes, Kalman trackers are used to model eye/head dynamics, and a probabilistic based appearance model is used to represent eye appearance. By combining three separate modalities, with specific enhancements within each modality, [31] allows eyes to be treated as robust features. This research is going to use this method for the eye detection part. This method can detect and track multiple eyes in a single frame also.

## 4.5 Iris Detection

The iris image is of radial nature, rather than rectangular, which makes it unsuitable to be processed by any mathematical or statistical model of linear nature. The main idea here is that iris is the most roundish object of the eye, so if this work can detect the most roundish object in the eye region it can detect the iris. The process is done in the following steps:

- First this work processes the image to be fit for finding the most roundish object by: converting the RGB image into grayscale image, converting the image to black and white, removing the noise by removing all object containing fewer than 10 pixels and filling the gaps.
- Then the boundaries have been detected.
- At last this work determines which objects are round by estimating each object's area and perimeter

and using these results to form a simple metric indicating the roundness of an object. This process can determine any circular object and by fixing a threshold value of 0.65 roundness this work can detect the iris. The threshold value to 0.65 has been set as the eye is not wide open and the iris is not fully visible in all cases.

# **4.6 Position of the Iris Detection and Calculations**

This work will split the image into 9 grids and find the position of the iris by using getpixelposition(handle) function, which gets the position, in pixels, of the component specified by handle. This function returns the position as a four-element vector that specifies the location and size of the component: [distance from left, distance from bottom, width, height]. So using this process this work can find in which grid the iris is located in. Then the deflection of the iris between two consecutive frames will be calculated. If the number of deflection between two consecutive frames is greater than a threshold value this work will categorize the action under avoiding eye contact.

## 5. CONCLUSION

Questionable observer detection has some great significance: it can be used by law enforcement officials such as Police force; RAB, CID etc.; it can prevent and detect theft in shopping malls and other places; it can also prevent violence or riot in public place and may lead to save a lot of life and goods. Here this research has described how to detect a questionable observer by detecting a specific suspicious action (avoiding eye contact). There are lots of suspicious actions which can categorize one person as questionable observer. These actions are not recognized yet with artificial intelligent systems which leave us with a lot of work to detect someone as questionable. Though these activities are not recognized yet, a lot of works have been done in the field of activity recognition. This research has discussed about the methods and algorithms used for activity recognition and also about their used dataset and accuracy. In near future this research will carry on to detect other suspicious activities like: repetitive touching of face, excessive head-turning, faster eve blink, widely open staring eyes, "Adam's Apple" jump, increased breathing rate, powerful grip of a bag or hand inside the bag, rubbing or wringing of hands, excessive clock watching, excessive shuffling feet, excessive leg shaking and excessive fidgeting. This research will also carry on finding out other activities and behaviors upon which one can be classified as questionable and try to include those activities and behavior on this research. By detecting all these questionable activities along with the currently proposed framework for detecting this specific activity - "avoiding eye contact" of this research; and calculating them with their respective weights this research hope to build a complete and elegant questionable observer in near future. After completing the questionable observer detector this research will focus on detecting questionable person from real time video feed.

#### 6. REFERENCES

- 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 (ISCMI), 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.

International Journal of Computer Applications (0975 – 8887) Volume 182 – No. 15, September 2018

- [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] Hjelmås, Erik, and Boon Kee Low. "Face detection: A survey." Computer vision and image understanding 83, no. 3 (2001): 236-274.
- [30] Kalavdekar Prakash, N. "Face Detection using Neural Network." International Journal of Computer Applications (0975–8887) 1, no. 14 (2010).
- [31] Haro, Antonio, Myron Flickner, and Irfan Essa. "Detecting and tracking eyes by using their physiological properties, dynamics, and appearance." In Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on, vol. 1, pp. 163-168. IEEE, 2000.