International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

# Anomalous Human Activity Perception/ Recognition on Crowd Places

*Shreya Sharma[1], Meetali Jain[2], Kunal Baraniya[3],Prof. Manoj Gupta[4], Prof Vandana Kate[5]*

[1,2,3,4]**Acropolis Institute of Technology and Research**

**ABSTRACT:-**

Given the rise in incidents of anti-social behavior, security has become a top priority in recent times. Many organizations have installed CCTV cameras for continuous monitoring of people and their actions. In a developed country with a population of 64 million, each individual is captured by a camera approximately 30 times daily. This generates a substantial amount of video data that is stored for a certain period of time (30 days in India). A 704x576 resolution image captured at 25 frames per second results in roughly 20GB of data per day. Given that monitoring this data constantly to identify abnormal events is an almost impossible task as it requires a workforce and their constant attention, there is a need to automate this process. Furthermore, there is a need to indicate which frame and which sections of it contain the unusual activity, which can aid in faster identification of abnormal behavior.

The proposed method involves generating a motion influence map for frames to represent the interactions captured in a frame. The key feature of this motion influence map is that it effectively portrays the motion characteristics of movement speed, movement direction, object size, and object interactions within a sequence of frames. The method then extracts frames with high motion influence values and compares them with the testing frames to automatically detect global and local unusual activities.

**INDEX TERMS: -** Motion influence map, Congested environments, Anomaly detection, Optical surveillance, Machine learning.

## I. INTRODUCTION: -

The rapid growth of VLSI chip technologies leads to the design of the surveillance camera and required electronics components at a low price. Hence, a surveillance camera network is established in most office premises, public places, shopping malls, streets, hospitals, and houses to monitor the environment continuously. It generates a vast amount of video containing the incidents that take place in the coverage of surveillance networks. The recorded video can be utilized to verify the activities in the future to know the status of the past incident. This recorded video also contains actions and activities patterns of people involved in the incidents [1]. The deep learning-based service-oriented architecture technology evolved to process the large data recorded in real-time applications in various situations. This architecture enabled the automation system to interact with objects such as various cameras, alarming sensors, and humans in a different environment [2].

### *Human actions and behavior*

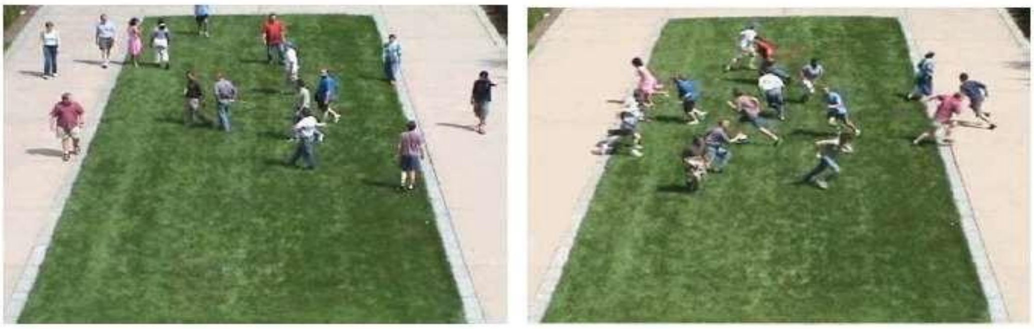
Human action is simple with little movement of the organ such as opening the window, clapping, pushing, punching, taking a water bottle, kicking, keeping a cloth on the shelf, etc. Human behavior is the combination of a sequence of motionless postures, and their corresponding temporal association such as running, fighting, playing volleyball, cycling, bike riding, etc. These characteristics are used to identify the human activities that are categorized as normal (accepted) and abnormal (unwanted) based on the applications such as robbery, violence, accidents, patient monitoring, theft, terrorism, natural disaster, chaotic, kids and elderly people monitoring. This study focuses on observing people's normal and abnormal behavior in crowded and uncrowded scenarios in an indoor and outdoor environment.

### *Human abnormal behavior*

An abnormal action is categorized into two types such as global and local abnormal actions. Local abnormal actions consist of a single or two persons whose actions would vary from the rest of the people (Loitering, falling, fighting, patient, etc). Global abnormal action includes many or groups of people whose behavior would be different such as violence, escaping behavior during a natural disaster, fighting, etc. Abnormal behavior recognition contains two modules such as action representation and measuring an abnormal action. In the view of environment surveillance, human abnormal behavior detection is the primary challenge to ensure public safety due to lengthy and numerous video datasets. Moreover, abnormal activities take place occasionally that lead to mistakes in identifying the specific frame of abnormal action. Hence, it needs more manpower to verify the video streams that make it expensive, inefficient, and time-consuming .

(a) Local usual activity

(b) Local unusual activity: bicycle in the middle of the frame



(c) Global usual activity

(d) Global unusual activity: running people across the frame

## II. LITERATURE SURVEY: -

Researchers in the field of perception-based supervision have been interested in detecting abnormal events or activities. Some analysts have concentrated on the modeling behavior of individuals and detecting anomalies based on this behavior. For example, researchers have used the probability ratio check with usual action classes, atomic events, and motion patterns to detect anomalies. Other researchers have focused on modeling the behavior of crowds, both globally and locally, by describing crowd behaviors using social force models, interaction energy potentials, motion variations, and local spatio-temporal cuboids. However, these methods have mostly concentrated on either local or global unusual activity recognition, and they have not considered the motion flows, object size, and interactions among objects, which can represent human activities in a crowded scene more effectively. In this paper, the authors propose a novel method that can efficiently detect and localize both local and global unexpected activities by considering the motion flows, object size, and interactivity among objects. Firstly the authors propose a motion influence map for displaying the activities in the crowded places and then devise a method for recognizing and containing atypical activities within a unified framework.

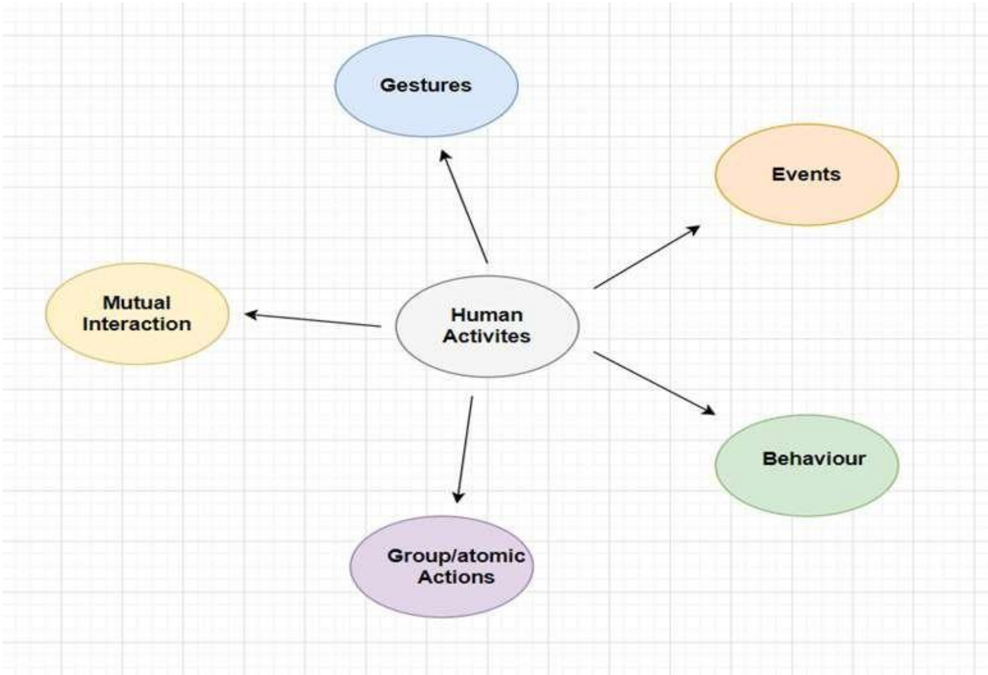
## III. RESEARCH MOTIVATION:-

Abnormal human activity detection is a major research area in the field of human activity recognition, with various applications in regions such as security, surveillance, and healthcare. The motivation for this research stems from the need to determine and prevent atypical or hazardous activities in real-time.

For example, in security and surveillance, the goal is to detect and prevent criminal activities such as theft, assault, and vandalism. In healthcare, the goal is to detect and alert healthcare professionals in case of any unusual activity that may indicate a potential health issue, such as falls, seizures, or sudden changes in vital signs.

The research in abnormal human activity detection is challenging due to the variability in human behavior and the difficulty in defining what constitutes abnormal. However, advancements in technology and increased computing power have enabled researchers to develop more sophisticated and accurate systems for detecting abnormal human activities.

Overall, the research motivation for abnormal human activity detection is driven by the need to improve safety and well-being in various applications and to provide early warning systems that can prevent harm to individuals and society.





## IV. PROBLEM STATEMENT:-

Detection of activities is a crucial issue in intelligent video surveillance. It is a basic challenge in computer vision, i.e. to identify the actions of humans in surveillance videos. These systems require real-time detection capability, but it usually takes a lot of time to detect the actual activity due to the huge size of the surveillance video clip and low analytical power. The big size is a result of the high resolution of the CCTV camera. Hence, reducing the resolution of the video clips and focusing on the activity being performed by the subjects is crucial. Numerous solutions using deep learning have been proposed until now, but none of them are efficient when there is a lot of information in the video, making it difficult to detect the actual activity. In such cases, compressing the rest of the details can help concentrate on the actual activity.

## V. PROJECT BENEFITS:-

The benefits of abnormal activity detection can be significant, as it has the potential to improve safety and well-being in various applications. Some of the key benefits of this technology include:

1. **Improved security and surveillance:** In security and surveillance applications, abnormal activity detection can help detect and prevent criminal activities such as theft, assault, and vandalism, thus improving the overall safety and security of individuals and communities.

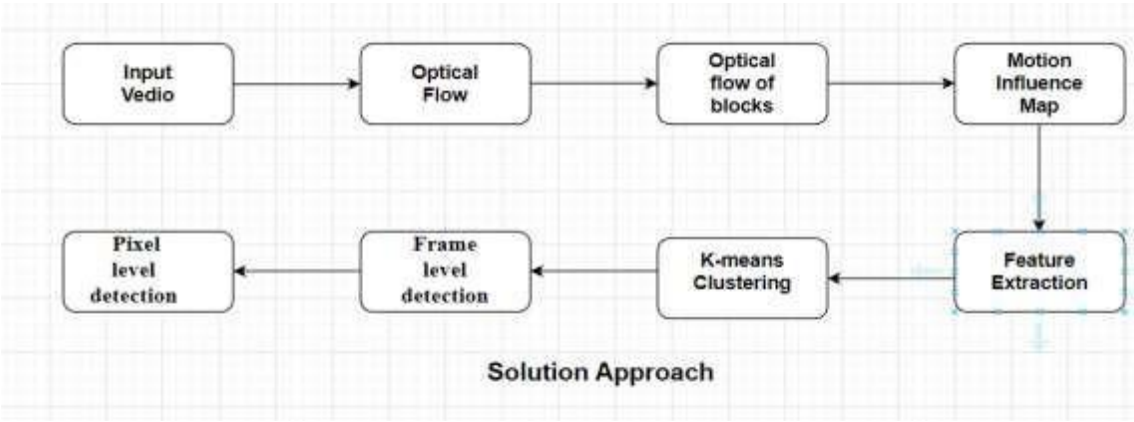
2. **Early detection of potential health issues:** In healthcare, abnormal activity detection can alert healthcare professionals in case of any unusual activity that may indicate a potential health issue, such as falls, seizures, or sudden changes in vital signs, allowing for early intervention and treatment.

3. **Enhanced decision-making and response times:** Abnormal activity detection can provide real-time information to relevant authorities, allowing them to make informed decisions and respond quickly to potential hazards, thus reducing the risk of harm to individuals and society.

4. **Increased efficiency and accuracy:** With the help of advanced technology, abnormal activity detection systems can provide more accurate and efficient results, reducing the number of false alarms and improving the overall efficiency of the system.

5. **Cost savings:** By detecting and preventing potential hazards before they occur, abnormal activity detection can help reduce the costs associated with responding to emergencies and treating injuries, thus saving money and resources.

Overall, abnormal activity detection has the potential to bring about significant benefits in various applications, making it an important area of research and development.




## VI. PROPOSED METHODOLOGY:-

The code is segregated into 5 distinct modules, including optflow of blocks, motion influence generator, create mega block, training, and testing.

This section presents a method to express movement characteristics for the recognition and detection of abnormal incidents in a densely populated vicinity. It is crucial to observe that two categories of exceptional actions are considered:

1. local and
2. global

Regional exceptional actions occur within a relatively small area. Various motion designs may arise in a portion of the view, like the unique manifestation of lifeless items or the swift motion of an individual when most of the other walkers are moving sluggishly. Universal exceptional activities transpire throughout the scene, such as when all walkers within a view abruptly begin to run to evacuate the zone.

### *VI. 1) Data Input and Pre-Processing*

The system receives a video file as an input, which undergoes pre-processing. The video is considered as a series of pictures referred to as frames and these frames are handled in order. An RGB frame is initially transformed to monochrome. A monochrome image contains solely the intensity data of the picture instead of the visible hues. RGB array is 3D (as it comprises of values of hues red, green and blue) whereas monochrome array is 1D.

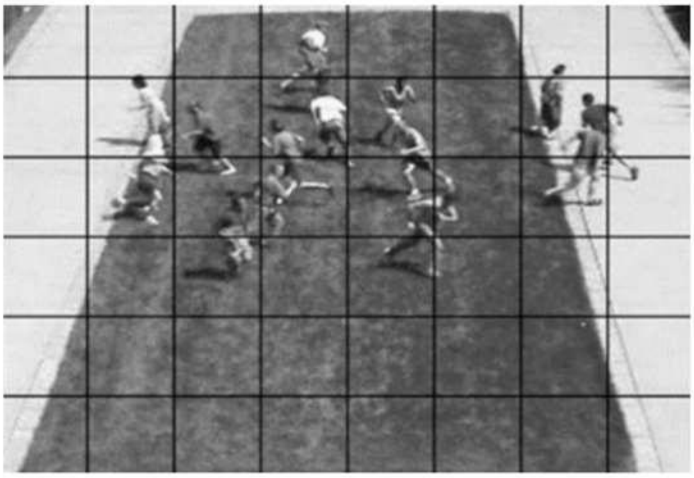
### *VI. 2) Optical Flow*

Following the pre-processing phase, for every frame within the video, optical motion is calculated for every individual pixel in a frame using the FarneBack algorithm. Optical motion denotes the arrangement of observable movement of objects, surfaces, and edges in a visual scene that occurs due to the relative movement between an observer and the scene. Optical motion is a vector of the structure (r, θ), where r signifies the size of each pixel and θ signifies the direction in which the corresponding pixel has moved in relation to the corresponding pixel in the previous frames. The function calcOpticalFlowFarneback() in openCV calculates a dense optical motion using Gunnar Farneback's algorithm.

### *VI. 3) Optical-Flow of blocks*

#### *Dividing a frame into blocks*

Once the optical motions are calculated for every individual pixel in a frame, we divide the frame into M by N identical sections, without any loss of generality, where the sections can be arranged by {B1, B2, . , , , BMN}. Figure 5.5 illustrates a frame measuring 240 x 320 being separated into 48 sections, where every section is of the magnitude 20 x 20.



***Calculation of Optical-Flow of each block***

Once the frames are divided into sections, we determine the optical motion of every section by calculating the mean of the optical motions of all the pixels constituting a section. Here, bi denotes the optical motion of the i-th section, J signifies the count of pixels in a section, and f ji indicates the optical motion of the j-th pixel in the i-th section. The optical motion of a section is a vector (r, θ) which indicates the magnitude of the movement of each section and the direction in which it has moved compared to the corresponding section in the preceding frames.

$$\mathbf{b}_i = \frac{1}{J}\sum_j \mathbf{f}_i^j$$

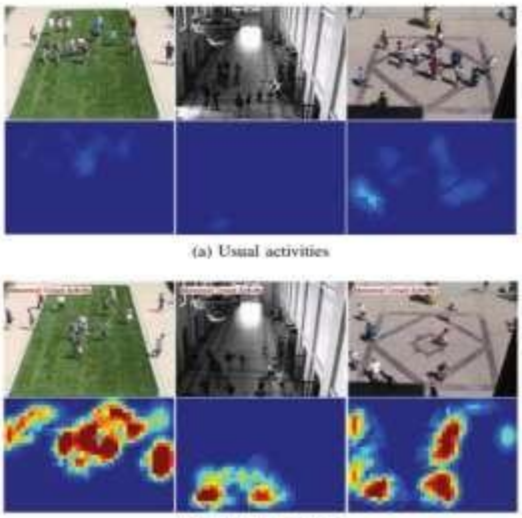
***VI..4) Motion Influence Map***

The direction of movement of a person within a group can be impacted by various elements, such as hindrances in the path, nearby pedestrians, and moving vehicles. We label this feature of interaction as motion impact. We presume that the sections affected by the movement of an object can be determined based on two factors:

The direction of motion

The speed of motion.

The quicker an object moves, the greater the number of adjacent sections influenced by it. Adjacent sections have more impact than faraway sections.



(a) Usual activities

(b) Unusual activities

***Algo/ Method for motion influence map***

INPUT: B ← motion vector set, S ← block size, K ← a set of blocks in a frame

OUTPUT: H ← motion influence map

H j (j K) is set to zero at the beginning of each frame ∈

for all i K do

Td = bi × S;

Fi/2 = bi + ? 2 ; −Fi/2 = bi − ? 2 ;

for all j K do

if i = j then

Calculate the Euclidean distance D(i, j) between

bi and bj

if D(i, j) < Td then

Calculate the angle ?in between bi and bj

if − Fi 2 < ?ij < Fi 2 then

H j ( bi) = H j ( bi) + exp− D(i,j)bi

end if

end if

end if

end for

end for

***Feature Extraction***

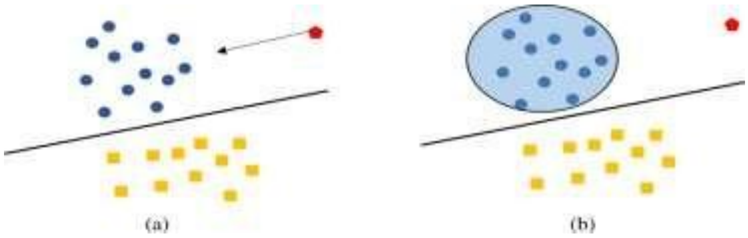
The motion influence map shows that a section where an anomalous event takes place, along with its adjacent sections, has distinctive motion impact vectors. Moreover, since an event is captured across multiple successive frames, we obtain a feature vector from a rectangular solid defined by n × n sections over the most recent t number of frames.

***Mega Block Creation***

The frames are divided into non-overlapping large blocks, each of which is a composition of several motion impact blocks. The total Motion Influence value of a large block is the sum of the Motion Influence values of all the smaller blocks that make up the larger block.

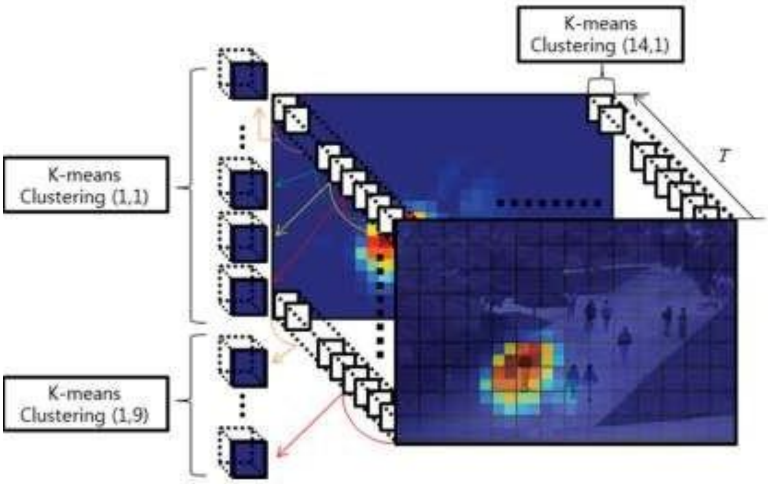
***Extracting Features***

Once the most recent 't' frames are separated into Mega Blocks, an 8 × t-dimensional merged feature vector is obtained for each megablock, encompassing all the frames. For instance, we take the mega block (1,1) of all the frames ('t' number of frames) and concatenate their feature vectors to form a combined feature vector for block (1,1).



***VI..6) Clustering***

For every mega block, we cluster using the spatiotemporal characteristics and establish the centers as codewords. Thus, for the (i,j)th mega block, we have K codewords, {w(i,j)k}K k=1. It is important to note that during the training phase, we solely employ video clips of regular activities. As a result, the codewords of a mega block represent the typical activity patterns that may arise in the corresponding region.




***Testing Phase***

Now that we have generated the codebook for normal activities, it is time to assess the developed model with a test dataset which contains abnormal activities.

***Minimum Distance Matrix***

During the evaluation phase, after extracting the spatial-temporal feature vectors for all mega blocks, we create a minimum distance matrix E over the mega blocks, in which the value of an element is determined by the minimum Euclidean distance between a feature vector of the current test frame and the codebook in the corresponding mega block.

***Frame-level Detection of Abnormal Activities***

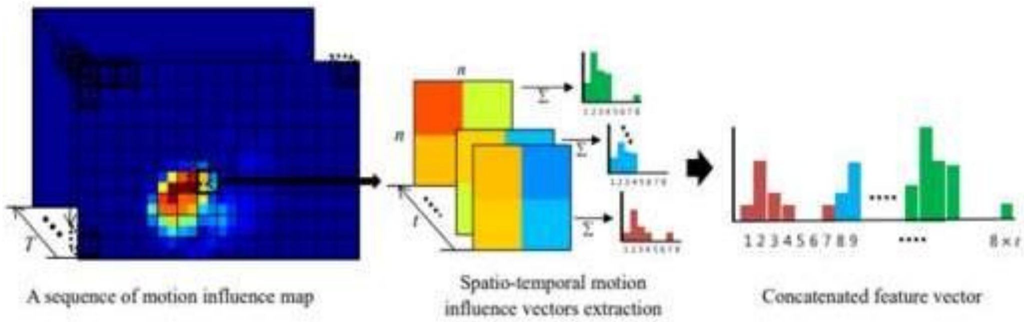
In a minimum-distance matrix, the smaller the value of an element, the less likely an abnormal activity is to occur in the respective block. Conversely, we can state that there are abnormal activities in t consecutive frames if a higher value exists in the minimum-distance matrix. Therefore, we identify the highest value in the minimum-distance matrix as the frame representative feature value. If the highest value of the minimum distance matrix is greater than the threshold, we classify the current frame as abnormal.

***Pixel-level Detection of Abnormal Activities***

Once a frame is detected as abnormal, we compare the value of the minimum distance matrix of each mega block with the threshold value. If the value is greater than the threshold, we classify that block as abnormal.

## *VII.* RESULTS AND ANALYSIS:-

To validate the efficacy of the proposed approach, we conducted experiments on publicly available datasets, namely UMN and UCSD datasets, which comprise of global and local unexpected activities, respectively. We evaluated the performance of the method on two different benchmarks for each dataset: 1) frame-level detection of global unusual activities in the UMN dataset, and 2) both frame-level and pixel-level detection of local unusual activities in the UCSD dataset. We evaluated the accuracy of pixel level detection based on the criterion that the overlapping area between the detected unusual area and the ground truth is greater than 40%, following the approach of Mahadevan et al.'s [5]. We utilized Receiver Operating Characteristic (ROC) curve, the area under the ROC curve (AUC), and Equal Error Rate (EER) metrics, as employed in prior work.




## *VIII.* CONCLUSIONS:-

We have introduced a method for detecting unusual incidents through the use of a motion influence map. This approach learns the features of motion influences directly, leading to a substantial increase in testing speed without sacrificing effectiveness. Our technique achieves the latest advancements in various datasets and is distinct from conventional subspace clustering.

In a world where security cameras are prevalent, efficient monitoring of these cameras is crucial. This requires a smart and speedy method of checking the footage. This would greatly reduce the workload of individuals who are tasked with monitoring the cameras, and allow for quicker responses to unusual events by integrating alarm systems and taking actions such as contacting the police or emergency services.

At a rate of 100 frames per minute, our approach can examine the recordings at an adequate pace and could be incorporated into surveillance cameras to automatically recognize anomalous occurrences. Response systems and alarms could then be activated based on the notifications produced by our approach.

## *IX.* FUTURE WORK:-

- ➢ Our future endeavors will be to expand the framework of the motion influence map to other video applications.

This can be expanded to detect various other types of uncommon occurrences commonly experienced.

By doing so, a system that can detect any abnormality will be created, making it adaptable to different environments. The algorithm should be thoroughly evaluated for opportunities for parallel processing and, if necessary, modified to improve accuracy.

- ➢ The model can be expanded by incorporating various other features that will aid in the detection of other uncommon events. For instance: Detection of knives, guns, or any suspicious objects.
- ➢ The model can also be used to address the current open challenge by incorporating it with an audio interface. For example: Detecting a woman who is running and shouting for help in an area.
- ➢ The model can be integrated with facial recognition technology to assist in criminal identification.

## *X.* ACKNOWLEDGEMENT: -

We wish to express our sincere gratitude to Prof. Vandana Kate and Prof. Manoj Kumar Gupta for mentoring our final year major project. We are grateful for their valuable guidance and encouragement in carrying out this project efficiently within the given time constraints.

Signature of the Student ……………………..

Name of Student:- Kunal Baraniya, Meetali Jain, Shreya Sharma

Enrollment Number:- 0827CO191029, 0827CO191033, 0827CO191055

Date 07/12/2020

### *XI.* REFERENCES: -

1. G. Xiong, J. Cheng, X. Wu, Y. Chen, Y. Ou, and Y. Xu, "An Energy Model Approach to People Counting for Abnormal Crowd Behavior Detection," Neurocomputing, Vol. 83, 2012, pp. 121-135.

2. X. Cui, Q. Liu, M. Gao, and D. N. Metaxas, "Abnormal Detection Using Interaction Energy Potentials," Proc. IEEE Conference on Computer Vision and Pattern Recognition, Colorado, USA, June 20-25, 2011, pp. 3161-3167.

3. J. Kim and K. Grauman. "Observe Locally, Infer Globally: A SpaceTime MRF for Detecting Abnormal Activities with Incremental Updates," Proc. IEEE Conference on Computer Vision and Pattern Recognition, Miami, USA, June 20-25, 2009, pp. 2921-2928.

4. L. Kratz and K. Nishino, "Anomaly Detection in Extremely Crowded Scenes Using Spatio-Temporal Motion Pattern Models," Proc. IEEE Conference on Computer Vision and Pattern Recognition, Miami, USA, June 20-25, 2009, pp. 1446-1453

5. B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in Proc. 7th Int. Joint Conf. Artif. Intell., San Francisco, CA, USA, Aug. 1981, pp. 674–679.

6. Dong-Gyu Lee, Heung-Il Suk, Sung-Kee Park and Seong-Whan Lee "Motion Influence Map for Unusual Human Activity Detection and Localization in Crowded Scenes" IEEE transactions on circuits and systems for video technology, vol. 25, no. 10, October 2015

7. Anderson D, Keller JM, Skubic M, Chen X, He Z (2006) Recognizing falls from silhouettes. In: 28th annual international conference of the IEEE engineering in medicine and biology society, 2006. EMBS 06. IEEE, pp 6388–6391

8. Auvinet E, Milton F, Saint-Arnaud A, Rousseau J, Meunier J (2011) Fall detection with multiple cameras: an occlusion-resistant method based on 3-d silhouette vertical distribution. IEEE Trans Inf Technol Biomed 15(2):290–300

9. Beleznai C, Gemeiner P, Zinner C (2013) Reliable left luggage detection using stereo depth and intensity cues. In: IEEE international conference on computer vision workshops (ICCVW), 2013, IEEE, pp 59–66

10. Candamo J, Shreve M, Goldgof DB, Sapper DB, Kasturi R (2010) Understanding transit scenes: a survey on human behavior-recognition algorithms. IEEE Trans Intell Transp Syst 11(1):206–224