| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

||Volume 8, Issue 5, May 2020||

A Stampede Detection using Motion Influence Vectors with OpenCV for Real Time Appraisal against Crime

Prakash Chandr Nandanwar[1], Ankur Taneja[2], Devendra Rewadikar [3]

PG Student, Department of Computer Science & Engineering, SAM College of Engineering & Technology, Bhopal, Madhya Pradesh, India[1]

Assistant Professor, Department of Computer Science & Engineering, SAM College of Engineering & Technology, Bhopal, Madhya Pradesh, India[2, 3]

ABSTRACT: Suspicious behavior is dangerous in public areas that may cause heavy causalities. There are various systems developed on the basis of video frame acquisition where motion or pedestrian detection occur but those systems are not intelligent enough to identify the unusual activities even at real time. It is required to recognized scamper situation at real time from video surveillance for quick and immediate management before any casualties. Proposed system focuses on recognizing suspicious activities and target to achieve a technique which is able to detect suspicious activity automatically using computer vision. Here system uses OpenCV library for classifying different kind of actions at real time. The motion influence map has been used to represent the motion analysis that frequently changes the position from one place to another. System uses pixel level presentation for making it easy to understand or identify the actual situation.

KEYWORDS: Stampede Detection, Unusual Activity Detection, Action Recognition, Motion Influence Map, OpenCV, Crowd based Activity Detection.

I. INTRODUCTION

Recognizing stampede is a challenging task that can be better achieved by computer vision. Sensors and wearable computing generally available, also called Internet of Things (IoT). It is at the core of assistive technologies to supply this data what's the activity when users attempt to understand their behaviour. Using unlisted data, researchers and an outsized number of users can enjoy more intelligent, with more knowledge of activity classification and understand the machines around them. There has been substantial research within the field of act Recognition, study to differentiate between normal activities in lifestyle (walking, running, sitting, standing etc.). The account has been suggested for a good sort of functions and activities with an algorithmically more complex structure. The aim of activity recognition is to spot the tasks and goals of 1 or more agents from a series of comments on the functions of agents and environmental conditions. Since the 1980s, this research area has attracted the eye of the many computing communities, because it provides personal assistance for several different applications and lots of different areas like medical, human-computer communication, or sociology thanks to its diverse nature, different area activity references are often mentioned as plan recognition, goal recognition, intention recognition, behaviour recognition, location estimation and location-based services [1].There are two practical limitations in using the real world surveillance footage, first of all, is the cost of upgrading the video capture equipment used in surveillance systems, and only cameras considered important are upgraded. Therefore it is common to create a video surveillance system by both modern and legacy hardware components. The quality of footage recorded from older cameras is usually poor, due to hardware limitations and it is common for footage having lower spatial and temporal resolution. Outdoor CCTV cameras are subject to natural light and generated bad contrast results when recording footage at night, which can effectively describe the content difficulty. Second, footage of densely populated urban environments depicts moving, self-aggrandizing. It can be difficult to generate a meaningful description of this as visual consistency of individuals and their actions. There are many fluctuations between frames due to recognizable shapes [2]. We propose a completely unsupervised method to detect unusual activity in crowded scenes. Neither normal nor unusual training examples are required before detection. Given that in crowded scenes, normal activities are behaviours performed by most people and abnormalities are behaviours that occur rarely and are different from most others, we use a motion influence method to solve the problem that scans a video with variable size and shape windows. The abnormality of each window is measured by a likelihood ratio test statistic, which compares two hypotheses about the characteristics of observations inside and outside the window. A semi-parametric density ratio method is used to model the observations, which is applicable to a wide variety of data. To reduce the search complexity of sliding window based scanning, a rapid two-round scanning algorithm is proposed. We have successfully applied our algorithm to detect activities that are incompatible in various ways, making performance competitive for other state-of-the-art methods that require supervision [3].

Fig. 1. Stampede [4]

Fig.1 represents group activities that can be traced either through physical sensor networks or computer vision. Sensors are not very much capable to be precise with the action recognized in group, rather than that computer vision is an effective approach for doing the same.Motion influence map is a motion representation technique through which energy is extractedthen motion influence map constructed from those energies. Motion influence map is able to classify the motion influence features and differences for detecting unusual human activity detection.

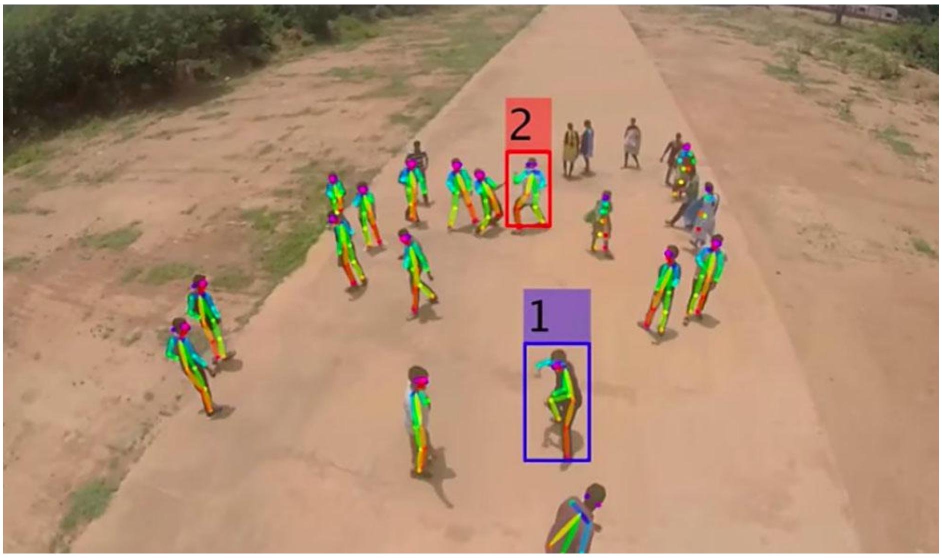


Fig. 2. Unusual Activity Detection using Skeleton Analysis [5]

While in practice such an airborne system would be one step above ground-based video cameras, it would still be hampered by the computational requirements of minimizing data received from large numbers of individuals. In addressing that issue, it can build a better system for autonomous drone surveillance that will work in real-time, by streamlining the process of collecting training data, using the cloud's computational power. But analysing action for a particular target in a crowd is a difficult task that might generate false alarm. It is useful for less crowded premises where less no. of subjects present.

II. Related work

Zakia Hammal et al. proposed a system that is based on traditional neural networks that provide training for human facial recognition. The system can be trained with various facial expressions and track activities w.r.t. Convicted sentiment. A similar modification was found in the findings in terms of infant quality between CNN-based Au detection tasks. The accuracy rate of recognition correct action or expression varies between 79 and 93% [6].He Xu et al. proposed a system based on RFID which is a physical sensor. The RFID system can be divided into the following three components: the reader, the tag, and the back-end computer system, which is shown in Fig.3 Readers and tags can communicate via antennas. The following are the steps of the RFID system work: (1) readers send radio frequency signals to the surrounding environment, and check if a tag is present; (2) When the tag in the reader's antenna reading range, the tag is activated by its own antenna to communicate with the reader and send its chip electronic code or other data; (3) The RFID reader receives the electronic product code (EPC) or data signal of the tag by the antenna; Then the data is decoded and processed, and will be routed to the back-end computer system. The system is complex due to complex installations and the cost of implementation increases.Action recognition module: Firstly, information is obtained from knowledge classification module. The knowledge formation of the method is then broken into different information segments with inverse temporal data. Through examination with all action models, if there is a collaborative action model, it tells the action response module to respond according to the action. If the action is not yet finished, then it does not react, and waits for a lot of knowledge. If no more information comes, it is assumed that the current action has ended and so the identification fails, then the formula clears all temporary knowledge and waits for new action knowledge. Fig 4 illustrates the action recognition method [7].Varsha Srirang Nanvare et al. performed a survey on various implemented systems on action recognition. Many researchers have worked to explore the methodology of multiple human pursuit and action detection in very real-time moving video, a thorough literature survey of recent work done by many authors in this exciting and application-Minded practical analysis area. In fact, the survey / review paper U.S. It is able to start for our analysis work on "multiple human pursuit and action recognition detection methods in very real-time moving video surveillance". The algorithms that have been proposed in earlier systems are effective for single targets, but are not convenient for multiple targets. Computational complexities are very high and do not work in crowded fields [8].Jiahao Li et al. proposed a system that is based on a pyramid energy map as a feature descriptor for a sequence of frames, it is able to save and present action histories that visually compare with recognized verbs. It is based on bidirectional neural networks that can track hidden layers and produce the most relevant results. It is also effective for single targets or skeletons, but confuses with multiple targets [9].Noor El Din Almadni et al. proposed a system that is based on bisite globality locality, preserving canonical correlation analysis, which aims to learn common feature subspaces between two sets. The second technique is Multiset Globality Preserving Canonical Correlation Analysis, which aims to deal with three or more sets. These form sequences of skeletons as a data set. The accuracy for the correct detection rate is 90.1% [10].

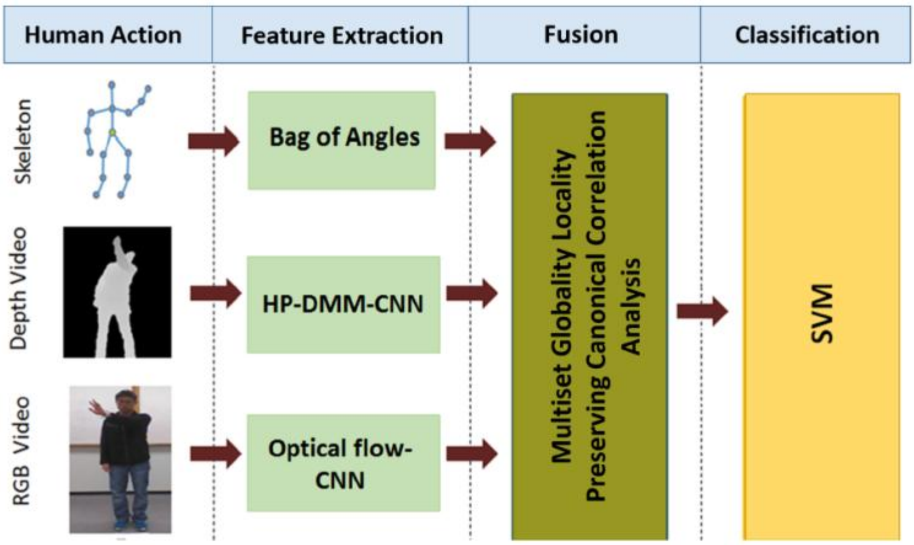
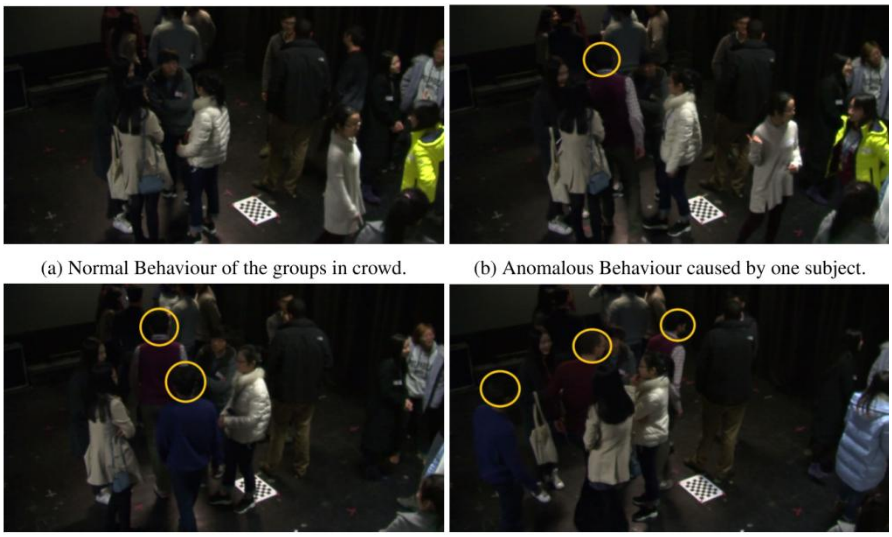


Fig. 3. The proposed human action recognition framework using MGLPCCA [10]

Aouaidjia Kam et al. proposed a system that is based on depth maps and posture recognition using deep neural networks. The system recognizes actions according to skeletal sequences. The system trained the network for various skeletal analyzes that later compared with recognized skeletons. The system achieved results based on score fusion operations that combined or fully analyzed results. But the congested region may contain different skeletal sequences that confuse the system to obtain the correct result [11].Saumalya Sen et al. proposed a system that is based on image parsing techniques. Image parsing deals with a wide variety of tasks that are performed by humans that can be identified in a sequence of frames. Classifies the verb - walking, running, clapping, jogging, cycling, surfing, etc. It is based on the correlation of foreground and background, through which the system extends the foreground object and stores these frames for future comparison. Image parsing integrates image segmentation, object recognition or recognition. The system is capable of recognizing human action with 88.70% precision [12].

III. Problem Identification

There are various researches have been implemented in the field of stampede detection but most of the systems are based on CNN or background or foreground subtraction that possess less accuracy rate with more false alarm rate. Muhammad Irfan et al. proposed a system which is based on random forest method. It does not work with large dataset. Random Forest creates a lot of trees (unlike only one tree in case of decision tree) and combines their outputs. By default, it creates 100 trees in Python sklearn library. To do so, this algorithm requires much more computational power and resources. On the other hand decision tree is simple and does not require so much computational resources.It is also not effective for noisy data or uneven data. Random Forest require much more time to train as compared to decision trees as it generates a lot of trees (instead of one tree in case of decision tree) and makes decision on the majority of votes.In learning phase the weight has been generated randomly that may affect the accuracy rate. The prediction rate is low because of crowded frames. The precision is 93.86 % that could be bit higher [13].



(a) Normal Behaviour of the groups in crowd. (b) Anomalous Behaviour caused by one subject.

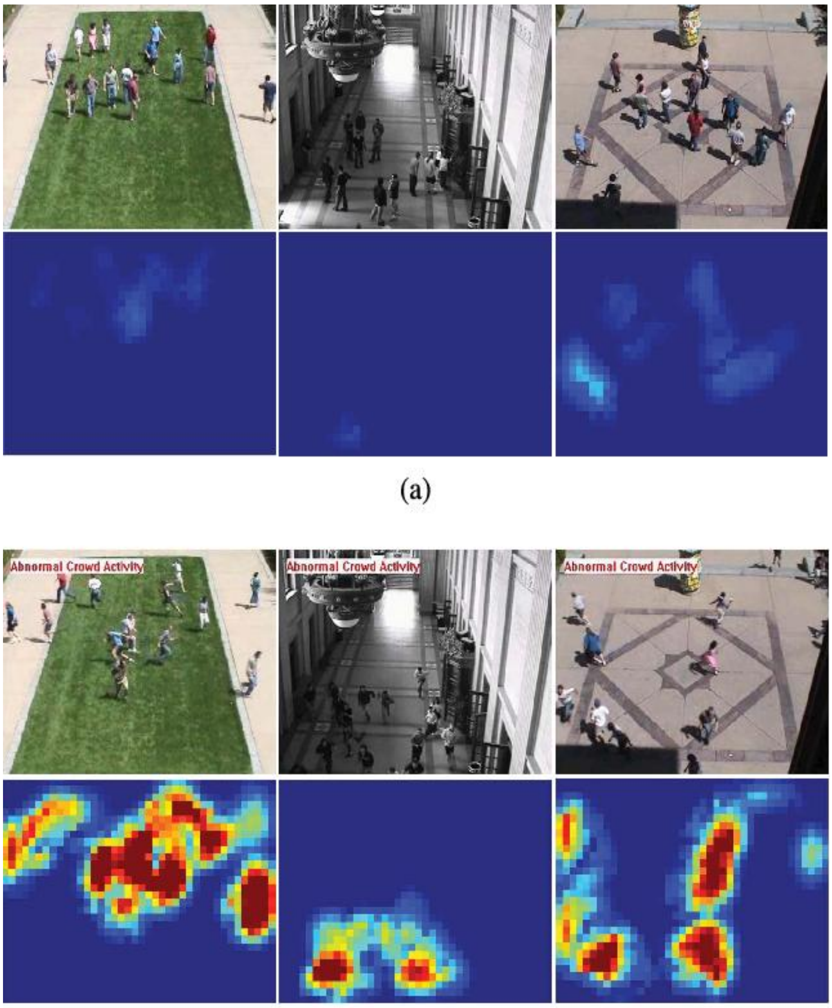
(c) Anomalous Behaviour caused by two subjects. (d) Anomalous Behaviour caused by three subjects.

Fig. 4. Unusual Activity Detection using Random Forest Method [13]

In this paper, authors proposed a mobile sensor based solution to anomaly detection in crowds. Owing to nonavailability of non-visual ground truth data, they collected, trainedandtested the systemon their owndatasetsintwodifferent experiments. To classify the two different behaviours of the groups we applied 'Random Forest' machine learningapproach [13].

IV. Proposed Work

Here the proposed work is able to recognize crowd activity and analyze whether the activity is usual or unusual. System purely debates with crowd based activities that ensure situations at real time. System uses OpenCV library along with python IDE including various packages that deals with best precision. System proposes motion influence map that comprises for better recognition rate. The proposed system is focused on the recognition of suspicious activity and is aimed at finding a method that can detect suspicious activity automatically by using computer vision methods. Proposed system classifies the differences among the frames using motion influence map that represents the frequent changes in the frames in a short interval of time. Recognizing unusual activity from crowd is difficult task especially for sensor networks; computer vision is an effective approach that can acquire real time human activities and later analyzes for uncommon frames.The main feature of the proposed motion influence map is that it effectively reflects the speed or speed, motion direction, and size of objects or subjects and their interaction speed characteristics in a frame sequence. Using the proposed motion effect map, we further developed a general framework in which we can detect both global and local paranormal activities. Furthermore, thanks to the representational power of the proposed motion effect map, we can localize unusual activities in a simple way. In our experiments on three public datasets, we compared the performance of the proposed method with other state-of-the-art methods and showed that the proposed method outperformed these competing methods.



(a)

(b)

Fig. 5. Unusual Activity Detection using Motion Influence Map

Let it be more precise through flow chart, first of all a crowd video has to be input that contains usual as well as unusual activities. Once the input made a frame selection getting started that also validates total number of frames. If current frame reached the last frame then process become end otherwise it will further proceed for unusual activity detection. Hj will be calculated that how much it influenced the map, that feature trace the feature vector. The feature which has been extracted is the influence density in motion influence map. It has been influenced the map as per the unusual density required to declare the uncommon activity only if it is greater than the threshold density. If it is greater than the threshold value then decision is declared as unusual activity otherwise no unusual activity has been detected. Once the unusual activity confirmed; it will cluster the influence area and represent it over the frames and finally shows on frame level that can easily identify by user whether the unusual activity has been performed or not. System is based on Motion Influence Map and OpenCV libraries. Motion Influence Map extracts the motion features and later clustered through K-Means Clustering. System clusters those frames which are having unusual activity that have been defined in motion influence map. Motion is either influenced or not; it can only be confirmed by observing influence direction and distance, if it is it far from the current position then influence has to be detected. Fig. 6 show the flow chart of the system.

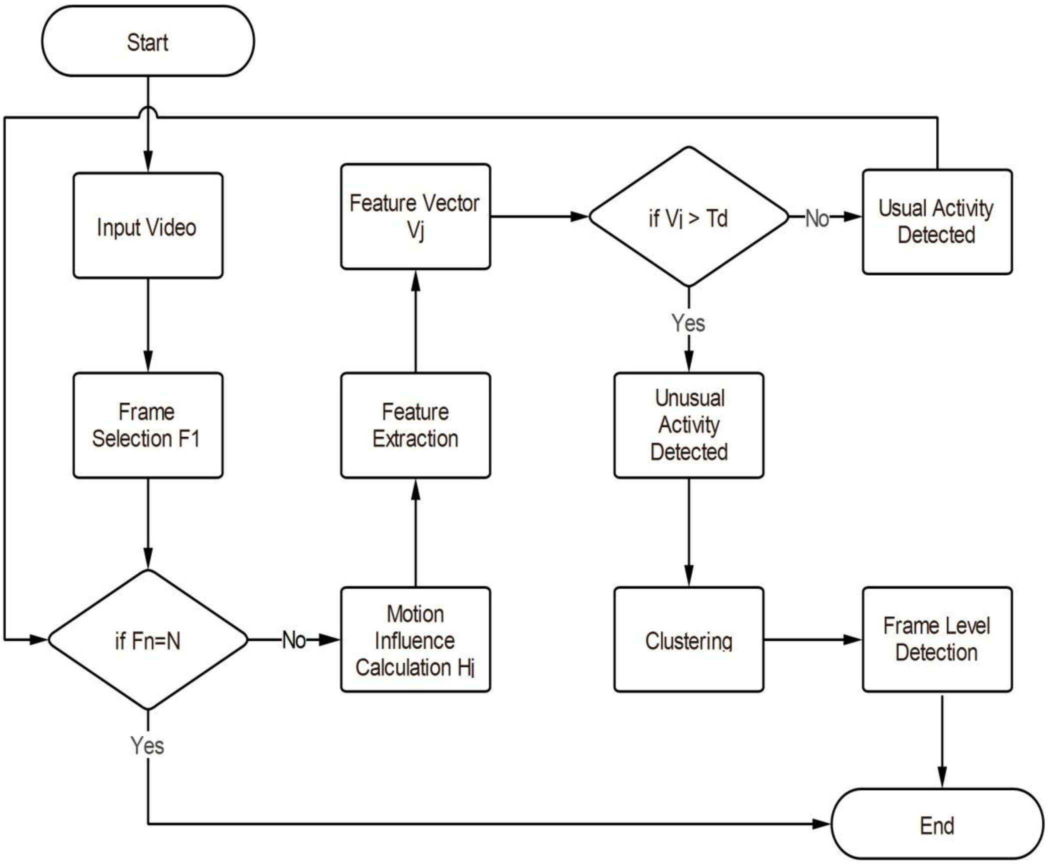


Fig. 6. Flow Chart

A. Motion Influence Vector Algorithm:

Require: S ← block size, K ← a set of blocks in a frame, B ← motion vector set, M ← motion influence map, I ← moving object, D(*i, j*) ← Euclidean distance between object i and block j, T_d is a threshold and φ_{ij} ← angle between a vector from object i to object j.

INPUT: B ← Motion Vector Set

OUTPUT: H ← Motion Influence Map

Step 1: Hj (j ∈ K) is set to zero at the beginning of each frame

Step 2: for all i ∈ K **do**

$T_d = \| b_i \| \times S$;

$\frac{Fi}{2} = \angle b_i + \frac{\pi}{2}$;

$-\frac{Fi}{2} = \angle b_i - \frac{\pi}{2}$;

for all j ∈ K **do**

if i ≠ j **then**

Calculate the Euclidean distance D(i, j) between b_i and b_j

if D(i, j) < Td **then**

Calculate the angle φ_{ij} between b_i and b_j

If $-\frac{Fi}{2} < \varphi_{ij} < \frac{Fi}{2}$ **then**

$H^j (\angle b_i) = H^j (\angle b_i) + \exp(-(D(i,j))/(\| bi \|))$

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

||Volume 8, Issue 5, May 2020||

end if
end if
end if
end for
end for
Step 3: H^j with respect to ∠ b_i is reflected motion influence map
Indicate motion influence vector V_j
Step 4:End

B is the input as motion vector set and H is output as motion influence map which is required to examine. In step 1, set H_j to zero at the beginning of each frame, H_j is motion influence at j block, where j belongs to K a set of blocks in a frame, In step 2, a 'for' condition has to be applied where i is position that belongs to K a set of blocks. Compute threshold value T_d equal to double mod of b_i an origin multiply with the block size S. There are two directions of a frame from origin-Fi/2 and Fi/2, so the angle is to be calculated as per the direction of motion vector. It is required to calculate the Euclidean distance between both positions- the origin and the motion vector. Once it has been calculated, then it will validate if it is less than the threshold value then calculate the angle between b_i and b_j, it is an angle between the origin and motion vector. Then it is required to find out in which direction it goes, if it is in satisfactory condition it will finally calculate the motion influence weight with vector position that later localize with pixel or frame level presentation. But there are certain steps towards localizations of motion influence. In the motion influence map, a block containing an unusual activity, with its neighboring blocks, has unique motion influence vectors.

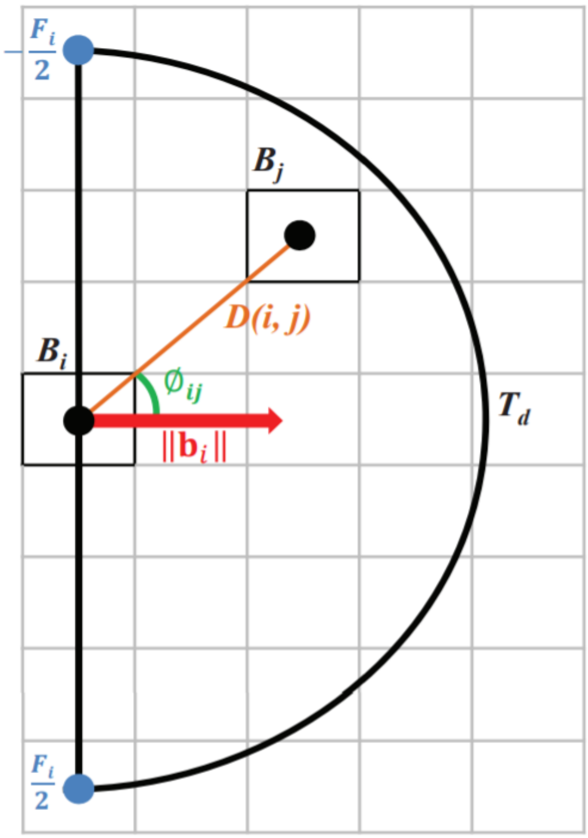


Fig. 7. Motion Vectors

Apart from this, since an activity is captured by several consecutive frames, we remove a feature vector from Cubeid as defined by the n × n blocks on the most recent T number of frames. Creating megablock frames are divided into non-overlapping mega blocks, each of which is a combination of multiple speed impact blocks. Motion Influence value of a megablock is the sum of the speed effect values of all small blocks that make up a large block. Extraction Features recently the number of 't' frames is divided into megablocks, for each megablock, an 8 × t-dimensional short feature vector is drawn in all frames. For example, we enclose the mega block ('T' number of 't' of frames) and their feature vectors, to create a distinct feature vector for the block (1,1). For each mega block clustering, we clustering using

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

||Volume 8, Issue 5, May 2020||

Spatio-Temporal features and setting the centers as a codeword. This is the reason, (i, j) for mega block, we have K codeword, {w (i, j) k} k k = 1. Here, we should note that in our training phase, Normal activities using clips. Therefore, the codeways of mega blocks create patterns of normal activities that can be in the respective area. In the case of minimum distance matrix testing, after removing spatio -typo feature vectors for all mega blocks, we make the minimum distance matrix e on the mega block, in which the value of one element between the attribute is less than Euclidean. The current test frame and related mega block are defined by the vector code. The frame level presentation of unusual activities in the minimum-distance matrix, the smaller the value of an element, the likelihood of having an unusual activity in related blocks is less.

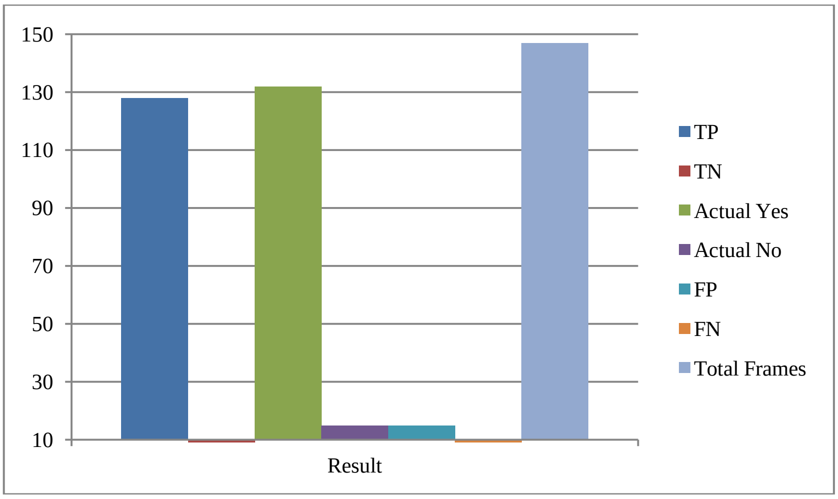
a) b)

Fig. 8. a) Usual Frame b)Unusual Frame

V. SIMULATION RESULTS

The simulation studies involve the deterministic various frames that may include usual or unusual activity. There are total number of 147 frames where actual unusual activity recorded in 132 frames and actual usual activity recorded as 15. There are four parameters- true positive, true negative, false positive and false negative. There are 128 true positive frames that mean there are 128 frames positively detected that returns correct recognition. There are 4 frames where unusual activity exists but system is not able to detect those frames that entertained in the category of true negative. There are 15 false positive frames where no unusual activity exists and system not detected it positively and there is no false frame considered as unusual, it means that there is no false negative. So, by observing all these datasets, the perceived accuracy is 9.27 %.

Graph 1 Result Analysis



Graph 1 shows the frame level results whether it contains unusual activity or not, result has been obtained on the basis of true positive, true negative, actual yes, actual no, false positive, false negative in the reference of total no of frames.

Table 1 Result Comparison

	Accuracy %
Nour El Din Elmadany [10]	94.14
Aouaidjia Kamel [11]	94.51
Soumalya Sen [12]	88.70
Muhammad Irfan [13]	93.86
Proposed	**97.27**

VI. CONCLUSION AND FUTURE WORK

A stampede detection using motion influence map is able to recognize unusual crowd activity with better precision rate. The precision rate is bit higher than other and less researches have been made over this concept. Proposed system is able to work for Prior Appraisal against Crime. The accuracy is 97.27 % which is good enough for recognizing unusual activity in complex backgrounds.The proposed system is capable enough to efficiently recognize the unusual human activity from crowd by using OpenCV and Motion Influence Map, which enhances the accuracy and proficiency of the system up to a great extent. The stampede detection can be implemented in various public places for prior and crime notification that enhances the casualty management. But accuracy is often important which requires enhancing for developing an ideal system that can be implemented practically.

REFERENCES

1. D. Lee, H. Suk, S. Park and S. Lee, "Motion Influence Map for Unusual Human Activity Detection and Localization in Crowded Scenes," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 25, no. 10, pp. 1612-1623, Oct. 2015.
2. https://arxiv.org/pdf/1605.05106.pdf
3. Y. Hu, Y. Zhang and L. S. Davis, "Unsupervised Abnormal Crowd Activity Detection Using Semiparametric Scan Statistic," 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops, Portland, OR, 2013, pp. 767-774, doi: 10.1109/CVPRW.2013.115.
4. https://www.voanews.com/science-health/coronavirus-outbreak/stampede-kenya-slum-residents-surge-food-aid
5. https://thenewstack.io/deep-learning-drone-detects-fights-bombs-shootings-in-crowds/
6. Zakia Hammal1, Wen-Sheng Chu1, Jeffrey F. Cohn1,2, Carrie Heike3, and Matthew L. Speltz4, "Automatic Action Unit Detection in Infants Using Convolutional Neural Network", 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII), IEEE.
7. He Xu1,2, Chengcheng Yuan1, Peng Li1,2, Yizhuo Wang3, "Design and Implementation of Action Recognition System Based on RFID Sensor", 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD 2017), IEEE.
8. Mrs. Varsha Shrirang Nanaware, Dr. Mohan Harihar Nerkar and Dr. C.M. Patil, "A Review of the Detection Methodologies of Multiple Human Tracking & Action Recognition in a Real Time Video Surveillance", IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI-2017).
9. Jiahao Li†, Hejun Wu∗ and Xinrui Zhou‡, "PeMapNet: Action Recognition from Depth Videos Using Pyramid Energy Maps on Neural Networks", 2017 International Conference on Tools with Artificial Intelligence, IEEE.
10. Nour El Din Elmadany, Yifeng He and Ling Guan, "Information Fusion for Human Action Recognition via Biset/Multiset Globality Locality Preserving Canonical Correlation Analysis", IEEE TRANSACTIONS ON IMAGE PROCESSING, 2018.
11. Aouaidjia Kamel, Bin Sheng , Po Yang, Ping Li , Ruimin Shen, and David Dagan Feng, "Deep Convolutional Neural Networks for Human Action Recognition Using Depth Maps and Postures", IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, 2018.
12. Soumalya Sen, Moloy Dhar and Susrut Banerjee, "Implementation of Human Action Recognition using Image Parsing Techniques", 2018 Emerging Trends in Electronic Devices and Computational Techniques (EDCT), IEEE.
13. Muhammad Irfan1,2, Laurissa Tokarchuk2, Lucio Marcenaro1, Carlo Regazzoni1, "Anomaly Detection In Crowds Using Multi Sensory Information", 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS).