International Journal of Technical Innovation in Modern Engineering & Science (IJTIMES)
Impact Factor: 5.22 (SJIF-2017), e-ISSN: 2455-2585
Volume 5, Issue 07, July-2019

Unusual Crowd Activity Detection using Euclidean Distance in OpenCV for Creating Motion Influence Map

[1]Lucky, [2]Bhupendra Panchal

[1]*Computer Science & Engineering, Oriental Institute of Science & Technology, Bhopal, Madhya Pradesh, India*
[2]*Computer Science & Engineering, Oriental Institute of Science & Technology, Bhopal, Madhya Pradesh, India*

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 Open CV 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*— Unusual Activity Detection, Action Recognition, Motion Influence Map, Open CV, Crowd based Activity Detection.**

I. INTRODUCTION

Recognizing human activities is a great work that is moving forward in the connected era. Sensors and wearable computing generally available, also called Internet of Things (IoT). It is at the core of assistive technologies to provide this knowledge what is the activity when users try to understand their behavior. Using unlisted data, researchers and a large number of users can benefit from more intelligent, with more knowledge of activity classification and understand the machines around them. There has been substantial research in the field of Human Activity Recognition (HAR), study to distinguish between normal activities in daily life (walking, running, sitting, standing etc.). The account has been suggested for a wide variety of functions and activities with an algorithmically more complex structure. It has been analyzed, which occurs in irregular intervals and potentially happens while doing other activities. The purpose of activity recognition is to identify the tasks and goals of one 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 attention of many computer science communities, because it provides personal assistance for many different applications and many different areas like medical, human-computer communication, or sociology due to its diverse nature, different area activity references can be referred to as plan recognition, goal recognition, intention recognition, behavior recognition, location estimation and location-based services [1].

Fig. 1. Group Activity Detection [2]

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 extracted then 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.

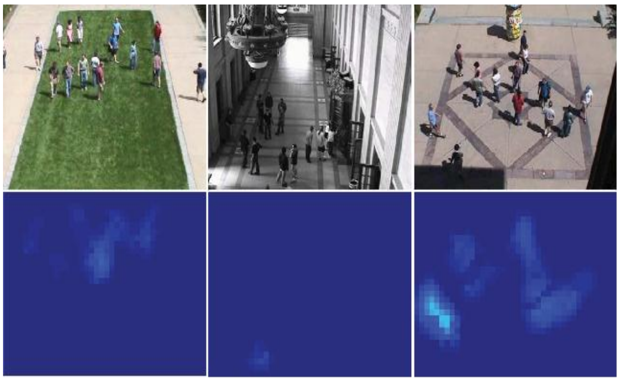


Fig. 2. Usual Motion Influence Map [3]

As in the Fig. 2, motion is normal or usual and there are no frequent changes over the frames in the particular interval of time. It can be observed that there is no influence over the map, if frequent changes made then it will represent the influence on map as shown in Fig. 3 given below.

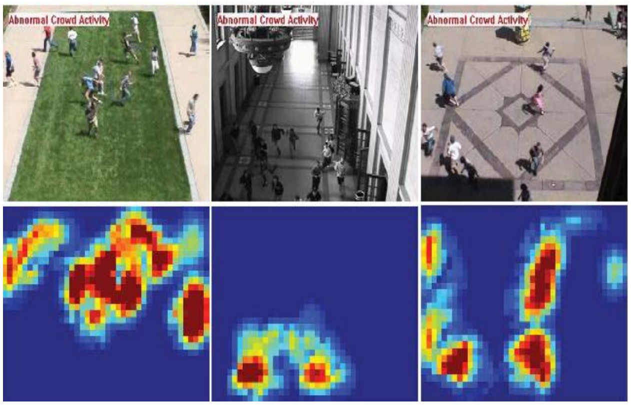


Fig. 3. Unusual Motion Influence Map [3]

II. Related Works

A. Literature Survey

Suspicious activities on public areas and personal safety are in serious danger. In public areas, millions of video surveillance systems are used, such as roads, prisons, holy sites, airports and supermarkets. Video surveillance cameras are not intelligent enough to recognize unusual activities even at real time. It is necessary to monitor the detection of suspicious activities and to verify the validity of surveillance video. It is required to recognized scamper situation at real time from video surveillance for quick and immediate management. Zakia Hammal et al. [4] proposed a system which is Based on Conventional Neural Network that trains for human facial recognition. System can be trained with different facial expressions and track activities w.r.t. to convicted expressions. The CNN-based AU detection revealed a similar modification in findings with reference to infant quality between tasks. The accuracy rate for recognition correct action or expression ranges between 79 to 93 %. He Xu et al. [5] proposed a system which is based on RFID which is a physical sensor. The RFID system can be divided into the following three components: Reader, Tag and Back-end computer system. Can communicate through the reader and tag antennas. The steps of the work of RFID system are as follows: (1) The readers send radio frequency signals in the surrounding environment, and check whether there is any tag; (2) When the tag in the reader's antenna reading range is activated by its own antenna to communicate with the reader and send its chip electronic code or other data; (3) RFID Reader receives an electronic product code (EPC) or data signal of the tag by antenna; Then the data is decoded and processed, and it will be sent to the back-end computer system. Varsha Shrirang Nanaware et al. [6] made a survey over various implemented system over action recognition. A number of researchers have worked on detection methodologies of multiple human chase & action recognition in a very real time moving video, thorough literature survey of the recent works done by numerous authors is being conferred during this exciting & application minded practical analysis field. In fact, the survey / review paper is done by U.S. as this is able to be the place to begin for our analysis work on “detection methodologies of multiple human chase & action recognition in a very real time moving video surveillance”. Jiahao Li et al. [7] proposed a system which is based on pyramid energy map as feature descriptor for a sequence of frames, it is able to save and present the action history that spatially compares with the actions recognized. It is based on bidirectional neural network which can back track the hidden layers and present the most relevant results. It is also effective for single target or skeleton but confuses with multiple targets. Nour El Din Elmadany et al. [8] proposed a system which is based on Biset Globality Locality Preserving Canonical Correlation Analysis, which aims to learn the common feature subspace between two sets. The second technique is Multiset Globality Locality Preserving Canonical Correlation Analysis, which aims to deal with three or more sets. It create sequences of skeletons as data sets. The accuracy for correct recognition rate is 90.1%. Soumalya Sen et al. [9] proposed a system which is based on image parsing technique. Image parsing relates different types of actions which are performed by human that can be recognized in sequence of frames. Action classifies as – walking, running, clapping, jogging, cycling, surfing, etc. It is based on foreground and background correlation through which system enhances the foreground object and stores these frames for future comparison. Image parsing unifies image segmentation, object detection or recognition.

III. PROBLEM IDENTIFICATION

There are various researches have been done in the field of human activity detection but fewer researches found for unusual human activities detection at real time using camera. Computer vision is a challenging approach that can acquire real time human activity. Earlier proposed systems are capable to detect the actions especially in single user human activity category using skeleton recognition. This kind of action can only be recognized in plain backgrounds not worth for non-plain backgrounds or outdoor scenes. Multi user action cannot be recognized by this system, moreover it is difficult for the system to recognize group or crowd based activities. System is intended to recognize human activities from crows to detect the unusual activities for getting prior notifications that helps to cure from heavy casualties. Skeleton tracking works with certain angle or distance, not useful for far elevations that results inaccurate precision.

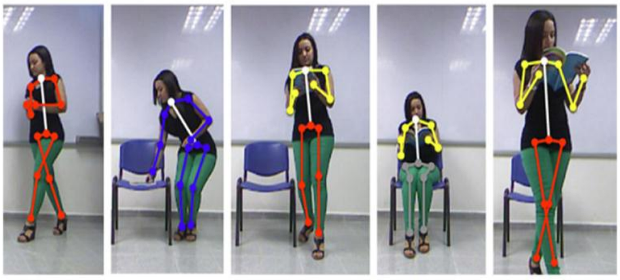


Fig. 4. Skeleton Recognition for Single User Activity Detection [10]

Suspicious activities on public areas and personal safety are in serious danger. In public areas, millions of video surveillance systems are used, such as roads, prisons, holy sites, airports and supermarkets. Video surveillance cameras are not intelligent enough to recognize unusual activities even at real time. The objective of the system is to recognize unusual human activity from crowd using motion influence map and Open CV for prior appraisal against crime in public places using camera. It will help to build or install a system that can work 24x7 for real time surveillance and decrease the crowd based criminal activities and saves us from fatal results.

IV. PROPOSED WORK & IMPLEMENTATION

Proposed work is able to recognize human activity in crowd and analyze whether the action is usual or unusual. System purely debates with crowd based activities that ensure situations. System uses Open CV library along with python IDE that deals with best precision. System proposes motion influence map that comprises for correct 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.

Fig. 5. Crowd Stampede Scene [11]

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 pixel 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. Figure 6 show the flow chart of the system.

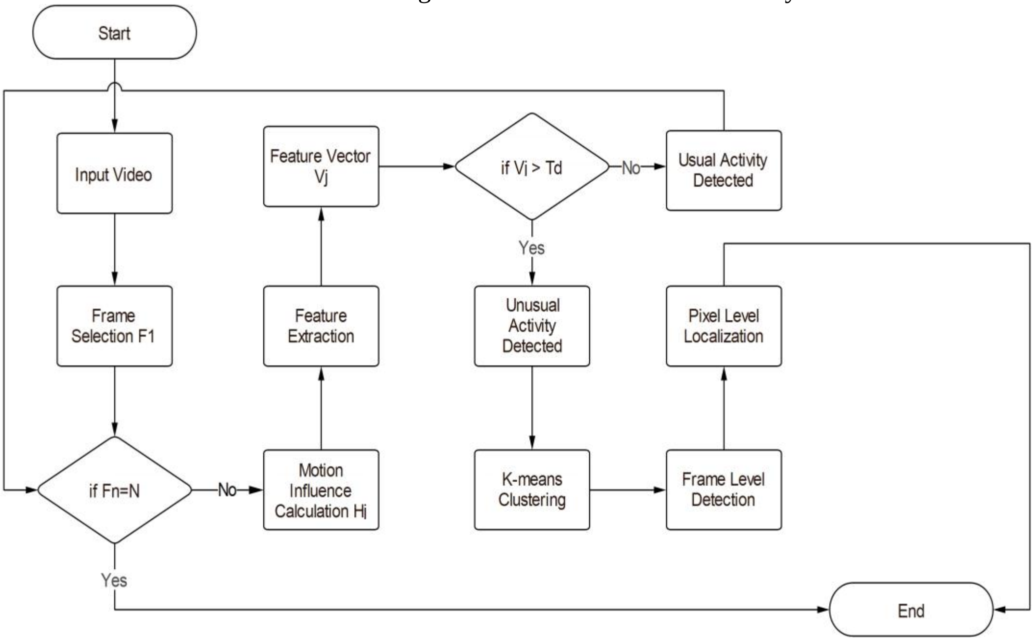


Fig. 6. Flow Chart

The implementation phase involves the actual materialization of the ideas expressed in the analysis document and it has been developed in the design phase. Implementation should be an ideal mapping of design documentation in a suitable programming language to obtain the required final product. This section discusses key decisions about the selection of the platform, the language used etc. These decisions are often influenced by many factors such as the actual environment in which the system works, requires speed, security concerns, other implementations-specific requirements, etc. and also we discuss briefly about important modules and methods that are present in the project. The code is divided into 5 modules, optoflobblocks, motion fluegenjener, creativemoblock, training and testing. In this section, a method has been described in the crowded scene to identify unusual activities and represent speed characteristics for localization.

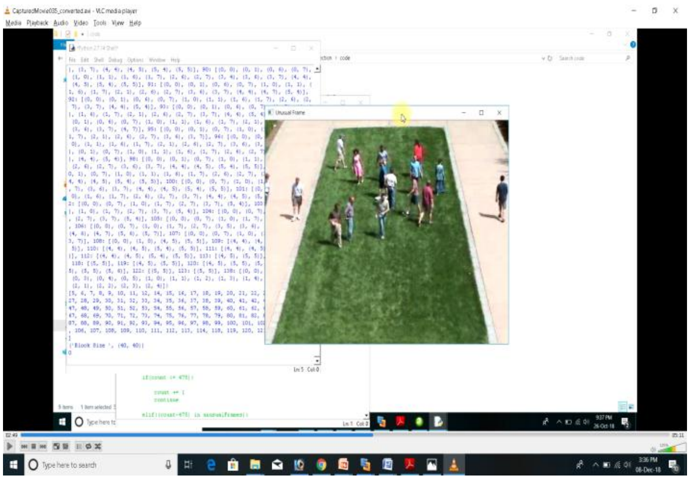


Fig. 7. Usual Activity Detection

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**
$T_d = \| b_i \| \times S$;
$= \angle b_i +$;
$= \angle b_i -$;
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 $< \varphi_{ij} <$ **then**
$H^j (\angle b_i) = H^j (\angle b_i) + \exp(-(D(i,j))/(\| bi \|))$
end if
end if
end if
end for
end for
Step 3: H^j with respect to $\angle b_i$ 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. 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 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.

Computer vision is now getting advanced using Open CV along with python. Python IDE is a platform where several image processing concepts computed with high level of precision. The configuration is quite easy to install in python where several packages available and can be used from any part of the country. Open CV is an advanced computer vision libraries that helps to compute effectiveness in the field of image processing. The motion influence map can be drawn easily in python and the pixel level presentation is often easier. Motion influence requires some python packages to run effectively with high level of accuracy of true recognition. Python also makes it easy to coding and testing simultaneously by adopting the test driven development (TDD) approach. There is a huge library for python as well as for Open CV which allow users to implement a system with fewer codes. Most of the smart phone applications develop in python that interact intellectually with human.

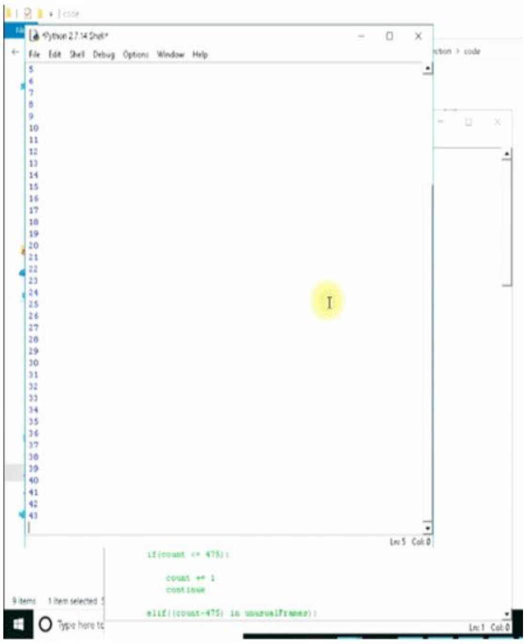
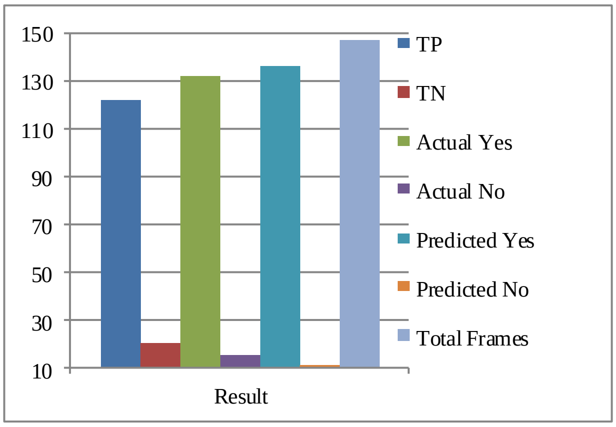


Fig. 8. Frame Analysis in Python IDE

V. RESULT ANALYSIS

The result has been analyzed on the basis of true positive, true negative, actual yes, actual no, predicted yes and predicted no. There are 147 as total no. of frames where 122 as true positive, it means that 122 frames have unusual activities that have been detected positively and 20 true negative that represents true rejection, system contains either usual or unusual but system rejected the real existence. But actual yes is 132 and actual no is 15. As per the prediction 136 is yes and 11 is no. So, on the basis of that over all accuracy has been calculated as 96.59 %, precision as 89.70 % and recall value as 92.42 %.

GRAPH 1. RESULT ANALYSIS

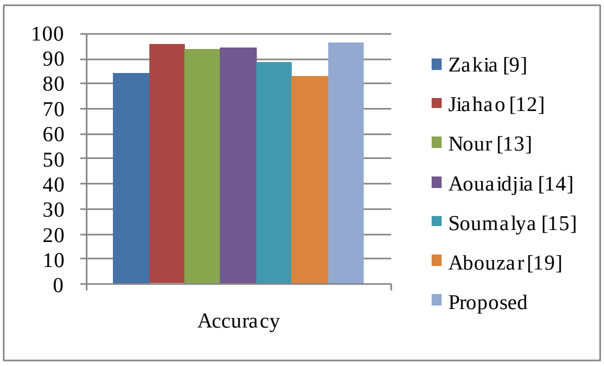


Graph 1 shows the frame level results whether it contains unusual activity or not, result has been obtained on the basis of true positive, true negative, actual yes, actual no, predicted yes, predicted no in the reference of total no of frames. On the basis of these simulations, precision, recall and overall accuracy have been calculated.

Table I. RESULT COMPARISON

	Accuracy %
Zakia Hammal [1]	84
Jiahao Li [4]	95.97
Nour El Din Elmadany [5]	94.14
Aouaidjia Kamel [6]	94.51
Soumalya Sen [7]	88.70
Proposed	**96.59 %**

GRAPH II. RESULT COMPARISON



VI. CONCLUSION & FUTURE SCOPE

The systems which have been proposed till now are intended to recognize simple human action such as walking, running and many more but not suitable for crowded area. System which has been proposed is able to recognize unusual human action from crowd and action accordingly using motion influence map and Open CV. 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 96.42 % 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 Open CV and Motion Influence Map, which enhances the accuracy and proficiency of the system up to a great extent. The Unusual Crowd Activity 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] Donghui Song, "https://www.youtube.com/watch?v=y7Jo_9Z9z0U", Published on Jun 13, 2018.
[3] 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.
[4] 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.
[5] 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.
[6] 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).
[7] 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.
[8] 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.
[9] 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.
[10] Frontiersin,"https://www.frontiersin.org/files/Articles/160288/frobt-02-00028-HTML/image_m/frobt-02-00028-g008.jpg".
[11] Hqdefault, "https://hvss.eventsmart.com/events/intro-concealed-carry/hqdefault/".
[12] D. Aranki, G. Kurillo, P. Yan, D.M. Liebovitz, and R. Bajcsy, "Continuous, real-time, tele-monitoring of patients with chronic heart-failure: lessons learned from a pilot study," in International Conference on Body Area Networks (Bodynets), 2014.
[13] Abouzar Ghasemi and C.N. Ravi Kumar, "A Novel Algorithm To Predict And Detect Suspicious Behaviors Of People At Public Areas For Surveillave Cameras" Proceedings of the International Conference on Intelligent Sustainable Systems (ICISS 2017)
[14] Y. Nam, S. Rho, and J. Park, "G3d: A gaming action dataset and real time action recognition evaluation framework," IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2012.

[15] Y. Nam, S. Rho, and J. Park, "Intelligent video surveillance system: 3-tier context-aware surveillance system with metadata," Multimedia Tools and Applications, vol. 57, pp. 315–334, 2012.
[16] A. F. Bobick and J. W. Davis, "The recognition of human movement using temporal templates," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, pp. 257–267, 2001.
[17] M. Blank, L. Gorelick, E. Shechtman, M. Irani, and R. Basri, "Actions as space-time shapes," 2005, pp. 1395–1402.
[18] I. Laptev and T. Lindeberg, "Local descriptors for spatio-temporal recognition," in Spatial Coherence for Visual Motion Analysis, 2006.
[19] X. Yang, C. Zhang, and Y. Tian, "Recognizing actions using depth motion maps-based histograms of oriented gradients," in ACM International Conference on Multimedia (MM), 2012.
[20] J. Wang, Z. Liu, J. Chorowski, Z. Chen, and Y.Wu, "Robust 3d action recognition with random occupancy patterns," in European Conference on Computer Vision(ECCV), 2012.
[21] X. Yang and Y. Tian, "Eigenjoints-based action recognition using navebayes- nearest-neighbor," in IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2012.
[22] M. Gowayyed, M. Torki, M. Hussein, and M. El-Saban, "Histogram of oriented displacements (hod): Describing trajectories of human joints for action recognition," in International Joint Conference on Artificial Intelligence (IJCAI), 2013.

