International Journal of Innovative Research in Computerand Communication Engineering

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

||Volume 8, Issue 6, June 2020||

Suspicious Human Activity Detection Using Image Processing

Harshith T R[1], Gurushekara K[2], Kumar Shivam[3], S Vidhya[4]

Final year B.E Students (UG),Department of Information Science & Engineering, The Oxford College of Engineering, Bangalore, India[1,2,3]

Assistant Professor, Department of Information Science & Engineering, The Oxford College of Engineering, Bangalore, India[4]

ABSTRACT: In today's insecure world the video surveillance plays an important role for the security of the indoor as well as outdoor places. The components of video surveillance system such as behavior recognition, understanding and classifying the activity as normal or suspicious can be used for real time applications. In this paper the hierarchical approach is used to detect the different suspicious activities such as loitering, fainting, unauthorized entry etc. This approach is based on the motion features between the different objects. First of all, the different suspicious activities are defined using semantic approach. Then the object detection is done using background subtraction. The detected objects are then classified as living (human) or nonliving (bag). These objects are required to be tracked which is done using correlation technique. Finally using the motion features & temporal information the events are classified as normal or suspicious. As the semantic based approach is used computational complexity is less and the efficiency of the approach is more.

KEYWORDS: Video Surveillance, Suspicious activity, Object detection.

I. INTRODUCTION

Computer vision is mainly used to study how to interpret, reconstruct and understand 3D scenes from its 2D images in terms of the properties of the structures present in the scene. Computer vision mainly includes methods for acquiring, analyzing, processing and understanding digital images. Video processing is a prominent research area in the field of computer vision. The detection and tracking of moving objects and activity recognition of these objects in video surveillance is one of the important tasks.

Human detection and tracking are a major component in many of the intelligent video management and monitoring applications in recent times. This finds application in surveillance video analysis for security, sports video analysis, detection of abnormal activities, patient monitoring, traffic monitoring and many more. Human activity recognition is mainly used for human-to-human interaction as it provides information about a person's identity, their personality and many more. As a result, it has many applications in video surveillance systems, human-computer interaction and robotics for characterization of human behavior all these require multiple activity recognition systems.

The two main approaches of detecting and tracking humans are frame difference method and background modelling method. Frame difference method is most suitable for no change in background and when there is a relatively static situation. Background modelling method is based on the Gaussian mixture model (GMM), Graph cut method. GMM and Graph cut methods are more complex and a larger amount of calculation is involved. The activity recognition approaches can be termed as local or global approaches. Local approach of video analysis mainly uses local interest points wherein each interest point contains a local descriptor which describes the characteristics of a point. By the analysis of these descriptors motion analysis is done. Scale Invariant Feature Transform (SIFT) and Space Time Interest Points (STIP) are some of the most commonly used local descriptors for videos. Global approach mainly uses the overall movement characteristics of the video. Most of the methods make use of optical flow to represent motion in a frame of video.

The classifier used for human activity recognition is mainly categorized into three basic types. Conditional Random Field (CRF), Hidden Markov Model (HMM) and Support vector machine (SVM). Where the CRF and HMM belong to the state model method, wherein due to the continuous change in activity sequence, activity recognition can be manipulated by modelling. Whereas, SVM uses nonlinear classification function which is established by known

samples to classify activities and hence overcomes the difficulties of parameter estimation in state model method. There is no need for considering the probability distribution. Hence it has its own wide variety of applications.

II. PROBLEM STATEMENT

The need for automated surveillance systems have become urgent as reliance on the human factor gives inaccurate results in the recognition of suspicious activities. Public places like subway stations, airports, and government buildings require detection of abnormal and suspicious activities to prevent crime before an occurrence such as automatic reporting of a person with Suspicious activity at the airport and to overcome crime. In this video surveillance moving object detection and recognition is the important research area of computer vision. Detection and recognition of moving is not an easy task as continuous deformation of objects takes place during movement. Any moving object has several attributes in temporal and spatial spaces. In spatial space objects vary in size whereas in temporal space it varies in moving speed.

III. PROPOSED SYSTEM

The proposed system is to automatically detect and estimate the 2D pose of humans in images recorded under uncontrolled environments. This project emphasis on detecting suspicious activities in public places like subway stations, airports, and government buildings require detection of abnormal and suspicious activities to prevent crime before an occurrence. In this video surveillance moving object detection and recognition is the important research area of computer vision. Detection and recognition of moving is not an easy task as continuous deformation of objects takes place during movement. Any moving object has several attributes in temporal and spatial spaces. In spatial space objects vary in size whereas in temporal space it varies in moving speed. This work mainly focuses on multiple human detection and activity recognition. Multiple human video datasets are considered and in order to detect and track multiple humans. Background subtraction technique is used for detecting moving multiple humans. Histogram of Oriented Gradient feature descriptor is used to extract features. For human activity recognition Support Vector Machine classifier is used.

Advantages:

- No need of human intervention as automatic Functioning performs the proper operation without any supervision.
- Complete automated operation.
- One-time installation.
- Low maintenance cost

IV. METHODOLOGY

1. **Input Data:**
   The input for the system is video stream. As the system is to be implemented to detect the suspicious activity its input is to be taken from the CCTV.
2. **Background Image Acquisition:**
   The background image is dynamically updated so that any new object entered the scene can be captured.
3. **Image Pre-Processing:**
   The changing light conditions, movement of reference background cause some noise introduced in the image.
4. **Object detection:**
   The foreground image is obtained by the subtraction of the input image from the background image. From this foreground image the required object is detected.
5. **Object Tracking:**
   The detected object is tracked in the scene so that we can determine if any new object is entered in the scene or if any object left the scene i.e. the person walk off the scene.

V. SYSTEM ARCHITECTURE

The architecture of a system describes its major components, their relationships (structures), and how they interact with each other. Software architecture and design includes several contributory factors such as Business strategy, quality attributes, human dynamics, design, and IT environment. We can segregate Software Architecture and Design into two distinct phases: Software Architecture and Software Design. In Architecture, non-functional decisions are cast and separated by the functional requirements. In Design, functional requirements are accomplished.

BLOCK DIAGRAM AND ITS COMPONENTS

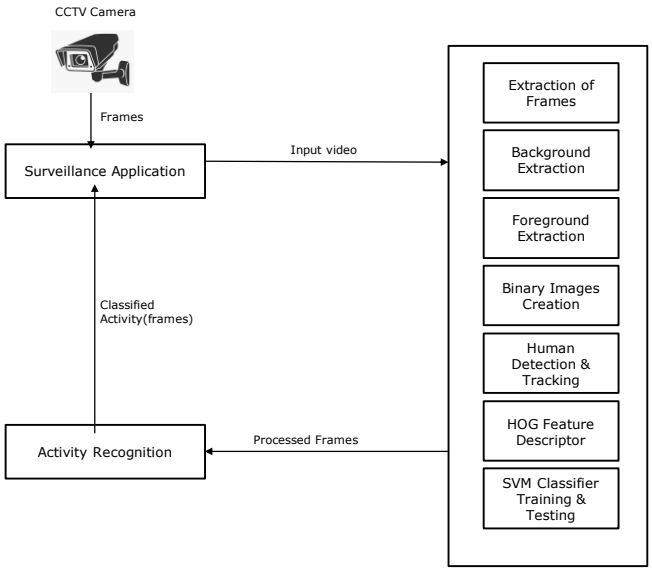

Fig. 1 Block diagram of approach used in the proposed system

A block diagram is a graphical representation of a system – it provides a functional view of a system. Block diagrams give us a better understanding of a system's functions and help create interconnections within it. Block diagrams derive their name from the rectangular elements found in this type of diagram. They are used to describe hardware and software systems as well as to represent processes. Block diagrams are described and defined according to their function and structure as well as their relationship with other blocks.

Extraction of frames is necessary as videos cannot be processed directly. Later, background subtraction technique is used to find the moving humans. In this technique a background image is considered, where each frame is subtracted by background image to obtain foreground images which show the moving human location. Convert the obtained foreground RGB image to grayscale images. Onto this result 2-D median filtering is used to remove noise components present in the video.

Once noise removal is done the grey scale images will be converted to binary images of 0s and 1s, where binary 1 is used for representing human region which is filled with white color and apart from moving human region binary 0 is used which represents absence of humans. Hence, binary image creation is useful for extracting any moving humans and objects in a video sequence. Followed by which dilation process is carried out which is typically applied on binary images obtained.

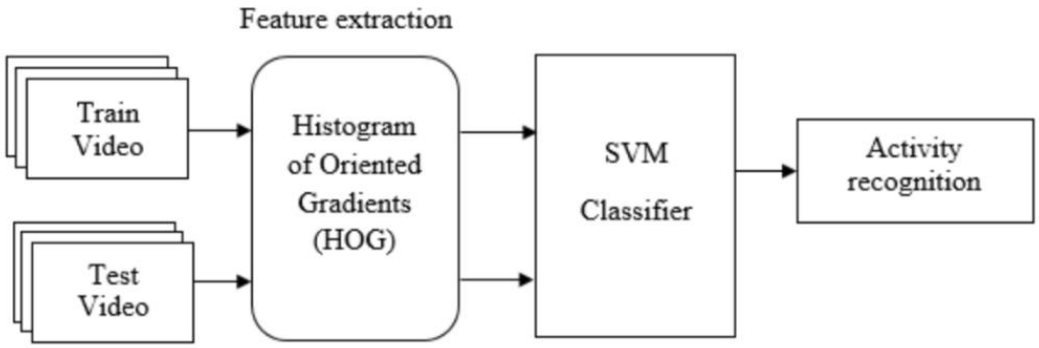

Fig. 2 Framework of Suspicious Activity Recognition

Figure 2 shows Human activity recognition in which HOG feature descriptor and SVM classifier along with train and test video dataset.

Once after detecting each individual human the next stage is recognizing their activities. After detecting the moving humans in a video, it is also necessary to determine the number of humans present later which activity recognition is

carried out. HAR consists of two phases: training and testing.

The flow chart of training and testing phase is as shown Figure 4.3.3. In the training phase the video dataset will be loaded first, then frames extraction is carried out. Training folder is created which contains the frames belonging to particular activities. Useful features are extracted for each activity being performed. HOG feature extraction technique is used for extraction of features. SVM classifier is used to train this extracted feature and it saves this trained SVM model which will be further used for testing. In the testing phase, the testing video is loaded and frames extraction, back- ground subtraction, binary image creation and HOG feature extraction steps will be done on the loaded test video. The obtained result will be inputted to the earlier trained SVM classifier, depending upon the feature match SVM classifier will recognize the particular activities performed.

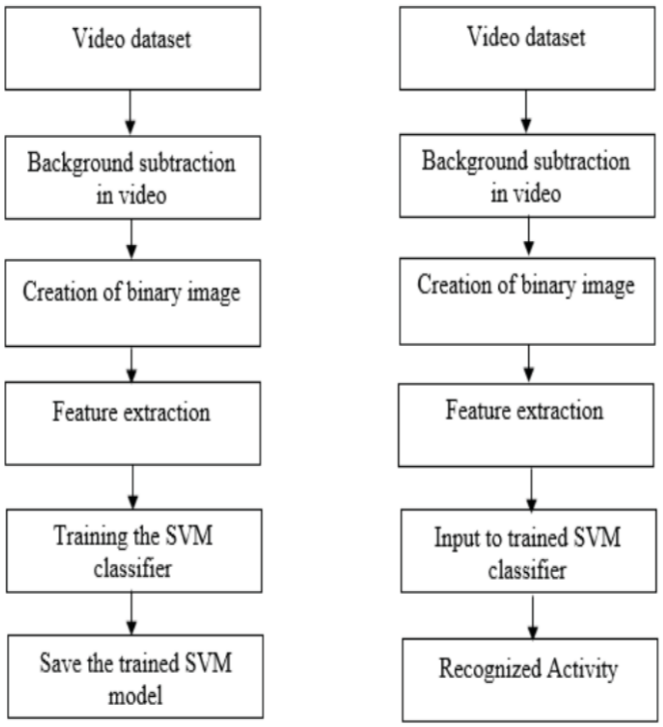

Fig. 3 Flow Chart of Training and Testing Phase

VI. IMPEMENTATION

Object Detection:

The template matching is used for object detection. In this method the cross correlation between a template image & the new image is performed.

Object Tracking:

For tracking the objects detected in the scene we use the correlation-based tracking method. In this method a small tracking window is centered on the object in the first frame.

Object Features Extraction:

Once the object in the frame to be tracked is fixed then its features are required to be extracted.

Defining the Suspicious Activities:

There are lots of activities which come under suspicious activity. But for the project work we have selected the following activities-

- Abandoned luggage
- Unauthorized Access
- Stabbing
- Shooting
- Fighting

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

||Volume 8, Issue 6, June 2020||

VII. RESULT

The suspicious activity detection in the video data is challenging task. It has number of difficulties such as complexity of scene, illumination of light, camera angle etc. Also, the definition of the suspect activity is scene/place dependant. Another problem is that the standard and challenging data sets are not easily available
for testing.

Fig. 4 Detected the suspicious objects in frame

VIII. CONCLUSION AND FUTURE WORK

The system was tested on the video recorded from a CCTV that captured scenes where a group of volunteers were exhibiting suspicious actions. The system has been tested on different real scenarios offering promising results.
The future work of this paper will include the detection of suspicious objects such as unmoved bags to make the system more effective. Reducing the computational time is also an important factor that can be improved.
Abandoned object detection and theft detection Majority of the works have been done for the abandoned object detection from surveillance videos captured by static cameras. Few works detected the static human as an abandoned object. To resolve such problems, human detection methods should be very effective and the system should check the presence of the owner in the scene, if the owner is invisible in the scene for a long duration then alarm should be raised. To resolve the problem of theft or object removal, the face of the person who is picking up the static object, should match with the owner otherwise an alarm must be raised to alert the security. Future work may also resolve the low contrast situation i.e. similar color problems such as black bag and black background which lead to miss detections. Future improvements may be integration of intensity and depth cues in the form of 3D aggregation of evidence and occlusion analysis in detail. Spatial-temporal features can be extended to 3-dimensional space for the improvement of abandoned object detection methods for various complex environments. Thresholding based future works can improve the performance of the surveillance system by using adaptive or hysteresis thresholding approaches. Few works have been also proposed for abandoned object detection from the multiple views captured by multiple cameras. To incorporate these multiple views to infer the information about abandoned objects can also be improved. There is a large scope to detect abandoned objects from videos captured by moving cameras.

Falling detection Most of the works have been done for fall detection of single people in indoor videos based on human shape analysis, posture estimation analysis and motion-based analysis. Future works can include the integration of multiple elderly monitoring which is able to monitor more than one person in the indoor scene. Many older people go for morning walks every day in public areas such as parks; to monitor these elder people, a future work can include one or more than one human fall detection from outdoor surveillance videos.
Accidents, illegal parking, and rule breaking traffic detection Several researchers have presented accidents detection, illegal parking detection and illegal U-turn detections from static video surveillance. These systems become incapable to detect these abnormal activities in more crowded traffic on roads. Future works should be based on unsupervised learning of transportation systems because no standard dataset is available for the training.
Violence detection Several research works have been done for the prevention of violence activities such as vandalism, fighting, shooting, punching, and hitting. To detect such violence activities, a single view static video camera has been used but sometimes this system fails in occlusion handling. Therefore, a multi-view system has been proposed by few researchers to resolve this problem but it requires important cooperation between all views at the low-level steps for abnormal activity detection. Future work may be an automatic surveillance system for moving videos. Improvements are required in accuracy, false alarm reduction, and frame rate to develop an intelligent surveillance system for the road traffic monitoring.

Fire detection Future work can include more improvement in accuracy, frame rate, false alarms reduction and also it

can be improved to detect far distant small fire covered by dense smoke.

REFRENCES

[1] Hanan Samir, "Suspicious Human Activity Recognition using Statistical Features." *2018 13th International Conference on Computer Engineering and Systems (ICCES)* (2018): 589-594.
[2] Mohannad Elhamod and Martin D. Levine, "Automated Real-Time Detection of Potentially Suspicious Behavior in Public Transport Areas", IEEE Transactions on Intelligent Transportation Systems, vol. 14, no. 2, June 2013.
[3] Sandesh Patil and Kiran Talele, "Suspicious Movement Detection and Tracking based on Color Histogram", 2015 International Conference on Communication Information & Computing Technology (ICCICT), Jan. 16–17
[4] P. Birch, W. Hassan, N. Bangalore, R. Young and C. Chatwin, "Stationary traffic monitor", Proc. 4th Internat. Conf. on Imaging for Crime Detection and Prevention (ICDP-11), pp. 1-6, 2011.
[5] Y. Tian, R. Feris, H. Liu, A. Humpapur and M.-T. Sun, "Robust detection of abandoned and removed objects in complex surveillance videos", Proc. IEEE, 2010.
[6] F. Porikli, Y. Ivanov and T. Haga, "Robust abandoned object detection using dual foregrounds", EURASIP Journal on Advances in Signal Processing 2008, 2008.
[7]James Ferryman et al., "Robust abandoned object detection integrating wide area visual surveillance and social context" in Pattern Recognition Letters, Elsevier, vol. 34, pp. 789-798, 2013.
[8] Mohannad Elhamod and Martin D. Levine, "Real-Time Semantics-Based Detection of Suspicious Activities in Public Spaces", *Proc. 9th Conf. CRV*, 2012.
[9]Achkar F, Amer A (2007) Hysteresis-based selective gaussian mixture models for real- time background maintenance. SPIE Vis Commun Image Process 6508: J1–J11.
[10] Adam A, Rivlin E, Shimshoni I, Reinitz D (2008) Robust real-time unusual event detection using multiple fixed-location monitors. IEEE Trans Pattern Anal Mach Intell 30(3):555–560