

# TRACKING AND RECOGNITION OF MULTIPLE HUMAN AND NON-HUMAN ACTIVITIES

Deepika R Chimkode<sup>1</sup>, Vimuktha Evangeleen Sails <sup>2</sup>, Ranjana S Chakrasali<sup>3</sup>

<sup>1</sup>Student (M.Tech), Dept. of Computer Science & Engineering, BNMIT, Karnataka, India

<sup>2</sup>Associate Professor, Dept. of Information Science & Engineering, BNMIT, Karnataka, India

<sup>3</sup>Assistant Professor, Dept. of Computer Science & Engineering, BNMIT, Karnataka, India

\*\*\*

**Abstract:-** The worldwide monitoring system of today plays a significant role in the safety sector. Moving object detection was frequently used in the video surveillance scheme. Motion estimation is also an important part of monitoring video processing, such as video filtering and video frame compression. Video surveillance system is a strong instrument to monitor individuals and their activities in the field of public safety. The reason for having a surveillance system is not lonely to position cameras in the location of the human eye, but also to allow them to recognize actions automatically. This article produces a intelligent understanding of the recognition and detection of objects of human activity system. Using image processing methods at each point of the suggested scheme a system was constructed based on the Caltech database of human activity characteristics obtained from frame sequences. Relevance Vector classifier used to classify the activity model in the dataset. The results of classification show elevated efficiency throughout the phases of training, testing and validation.

**Key Words:** (Size 10 & Bold) Key word1, Key word2, Key word3, etc (Minimum 5 to 8 key words)...

## 1. INTRODUCTION

Identifying and analyzing everyday human activities is an attractive location for researchers due to its efficiency and wide-ranging implementation in image analysis, machine learning and interaction between humans and machines. Usually, the aim of video monitoring systems is to monitor human behavior and activities. Any shift in human movement for safety and administrative reasons should also be noted[1][2].

Analysis is performed by a individual in manual video monitoring systems. Human video evaluation and decision-making interventions are partial in semi-automated surveillance, while video input, evaluation, processing and treatment with the suspect situation is completely independent of any human interference in the fully automated system.

The basic monitoring system process starts with the removal of the background, called motion and object detection, and is a technique of extracting the area of interest from the built background[4]. The area data obtained was then used as an

input to detect objects and analyze behaviors[4]. In addition, the implementation of the surveillance scheme relies on a behavior evaluation involving activity recognition or descriptive pattern between the object[5].

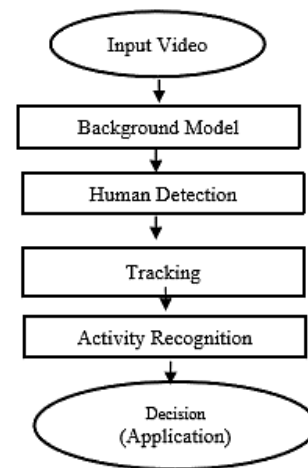


Fig 1. General Overview of Surveillance System

## 2. Related Work

Researchers have developed numerous algorithms and methods for the recognition of human activity in the past.

Suggested by Kaur, Rajvir and Sonit Singh[4], the Background Modeling, Detection and Tracking of Human in Video Surveillance System Video Surveillance System is a powerful tool used to monitor people and their security activities by government. The reason for having a surveillance system is not only to substitute cameras for human eyes, but also to allow them to automatically acknowledge operations. In this document, Human The Weizmann dataset is used to define and monitor different operations such as run, bend, hand wave, skip, etc. First background modeling is conducted with first n frames. Following this, human detection is conducted using the background subtraction algorithm and then tracking is done using Kalman filter.

Htike, Kyaw Kyaw, Othman O. Khalifa, Huda Adibah Mohd Ramli, and Mohammad AM Abushariah[2 ] proposed Human Activity Recognition for Video Surveillance using Posture

Sequences Human Activity Recognition has become an extremely interesting field of study as it has many potential applications, including automated monitoring, sign language interpretation and human-computer interfaces. Comprehensive research has been conducted in this industry in recent years. This paper presents a part of a novel using a static camera to acknowledge a video surveillance human posture. The training and testing stages were implemented using four separate K Means, Fuzzy C Means, Multilayer Perceptron Self-Organizing Maps, and Feedforward Neural networks. For used classifiers, the accuracy recognition is calculated.

Meng, Binghao, Lu Zhang, Fan Jin, Lu Yang, Hong Cheng, and QianWang.[5] proposed detection of abnormal events using profound video surveillance networks. A novel method is suggested to detect abnormal events. This method is based on deep spatio-temporal networks capable of sequentially representing video frames abnormal events in the real globe are rare and involve small samples as well as large quantities of common video data. It is hard to apply directly with deep networks that usually involve amounts of samples labeled. By pre-training video networks that are irrelevant to uncommon events and refining networks of fine tuning, our method solves this problem.

### 3. METHODOLOGY

Caltech dataset collects human activities video which is usually in a moving traffic in this job. The environment in which the video sequence obtained was an outdoor setting (which will be entered into the scheme). In addition, various individuals were regarded. The aim here is to test the system's ability to detect multiple human activity and object detection like car.

Figure 1 demonstrates the various recognition of human activity approach. The first stage in object detection and various human identification and recognition of human activity is to acquire the video datasets containing various human and distinct actions performed by each person. In this job datasets are regarded human activities such as walking, running. Recognition of activity begins with reading a video file. Once the next step is to transform this video into successive frames, 24 frames per second will be extracted.

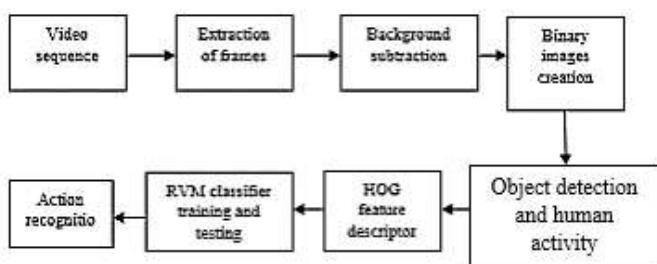


Fig 1. Block diagram of approach used in HAR

Frame extraction is essential as it is not possible to process videos straight. Later, the method of subtraction of the background is used to discover the moving people and cars. A background picture is regarded in this method, where each frame is removed from the background picture in order to acquire foreground pictures showing the place of the moving beings and cars. Convert the RGB picture acquired from the foreground to gray pictures.

The gray pictures will be transformed into 0s and 1s binary images, using binary 1 to depict a white colored human region, apart from the moving region., is used binary 0 representing human lack. Binary image development is therefore helpful to extract moving people and items in a video sequence.

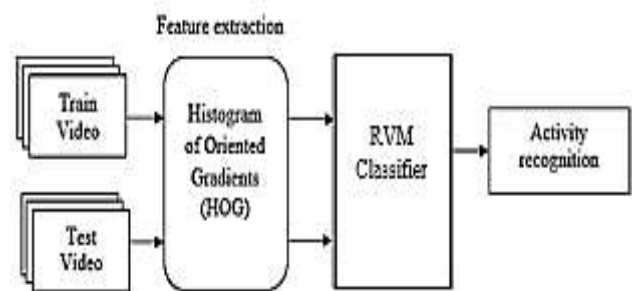


Fig. 2. Frame work of HAR

Figure 2 demonstrates recognition of human activity in which HOG and RVM classifier together with video train and test dataset. Once the next phase recognizes their operations after identifying each individual person. It is also necessary to determine the later number of people present, after detecting the moving people in a video, which activity recognition is performed. Activity recognition composed of two phases: phase 1 training and phase 2 testing.

As shown in Figure 3, the flow chart for the training and test stage. The dataset will first be loaded in the training stage, then frame extraction will be performed. It creates a training folder that includes the frames that belong to specific operations. For each activity being conducted, useful characteristics are extracted. The extraction method of the HOG function is used to extract characteristics. The RVM classifier is used to train this extracted feature and saves this trained RVM model to be used for further testing. The testing video will be loaded in the testing stage and the extraction of frames, background subtraction, binary picture development and extraction steps of the HOG function will be performed on the charged test video. Depending on the function match RVM classifier will acknowledge the specific operations conducted, the result obtained will be added to the earlier qualified RVM classifier.

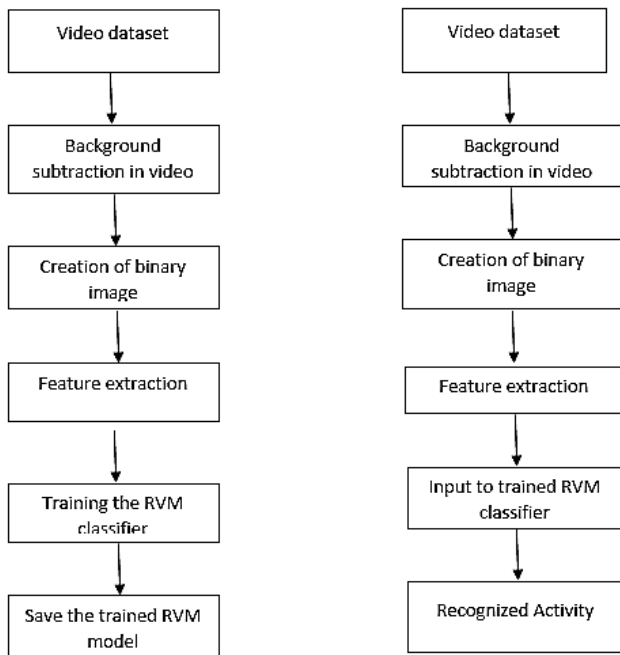


Fig. 3. Training and testing stage flow chart

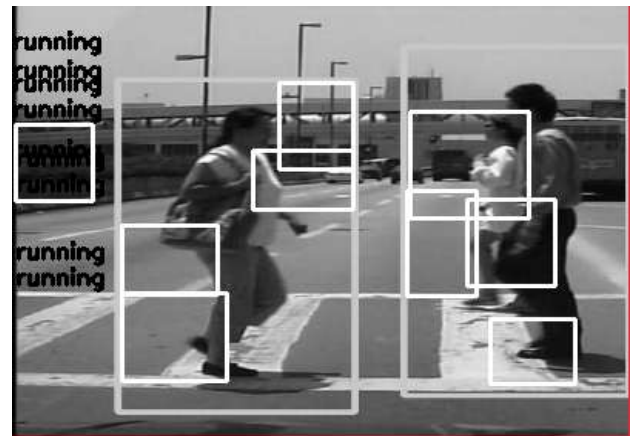
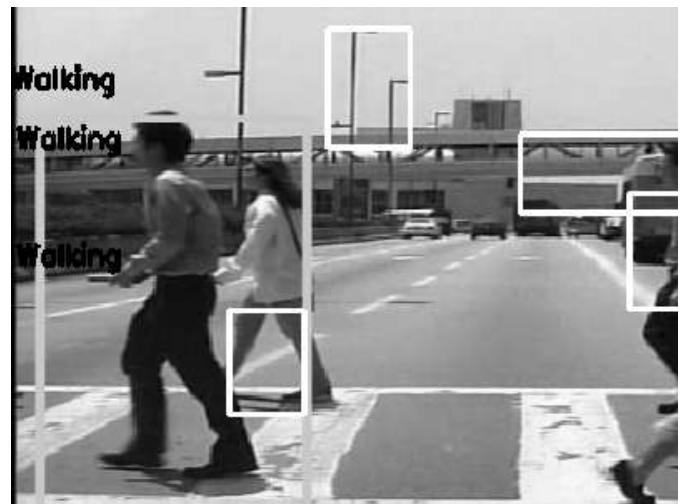
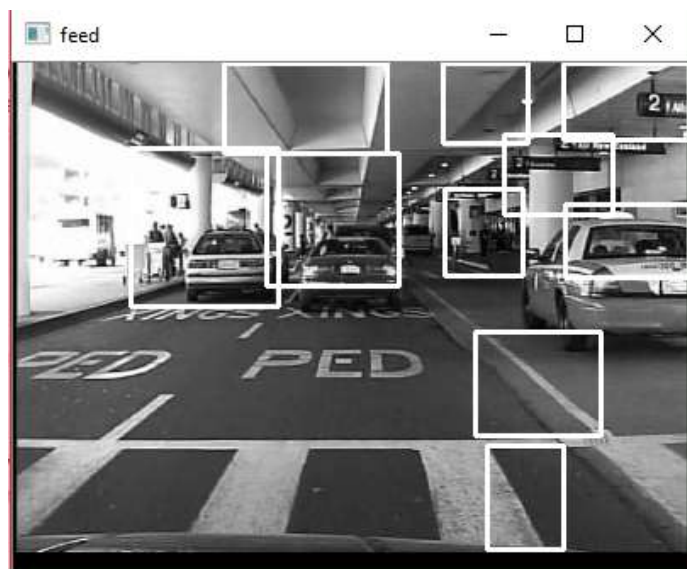


Fig. 4. Recogniton activities

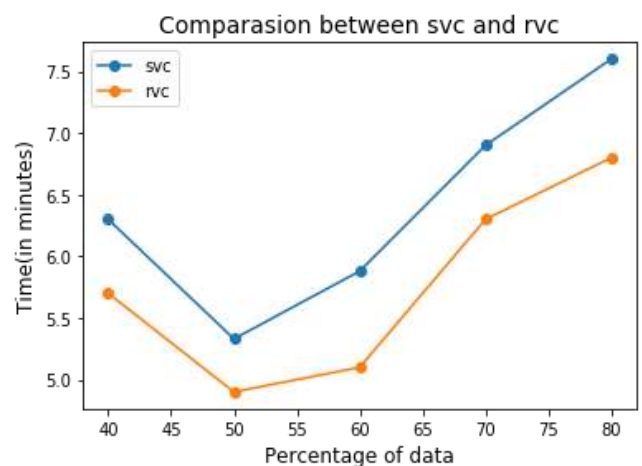
### 3. RESULTS

The results show the object detection in background and the activity recognition of multiple humans. Caltech dataset consists of 2-3 individuals who carry out various tasks such as walking, running and moving traffic in background. The video is 1:32 seconds long with 809 MB of memory size. Video pixel resolution is 640 X 480 with 20 frames per second frame rate. Human activity videos for walking, running and are collected from Caltech database and are then converted into frames. Around 50000 images are given for training the data.



#### 3.1 Comparison between SVM and RVM

The data is trained for 5 iterations using 40%, 50%, 60%, 70%, 80% of training data using support vector classifier and relevance vector classifier. Relevance vector classifier takes less time for training then support vector classifier. Implementation of this approach using RVC using better results compared to SVC



**REFERENCES**

- [1] Hossen, Muhammad Kamal, and Sabrina Hoque Tuli. "A surveillance system based on motion detection and motion estimation using optical flow." In Informatics, Electronics and Vision (ICIEV), 2016 5th International Conference on, pp. 646-651. IEEE, 2016.
- [2] Htike, Kyaw Kyaw, Othman O. Khalifa, Huda Adibah Mohd Ramli, and Mohammad AM Abushariah. "Human activity recognition for video surveillance using sequences of postures." In e-Technologies and Networks for Development (ICeND), 2014 Third International Conference on, pp. 79-82. IEEE, 2014.
- [3] M. Thida, Y. L. Yong, P. Climent-Pérez, H.-l. Eng, and P. Remagnino, "A literature review on video analytics of crowded scenes," in *Intelligent Multimedia Surveillance*. Springer, 2013, pp. 17–36..
- [4] Kaur, Rajvir, and Sonit Singh. "Background modelling, detection and tracking of human in video surveillance system." In *Computational Intelligence on Power, Energy and Controls with their impact on Humanity (CIPECH), 2014 Innovative Applications of*, pp. 54-58. IEEE, 2014.
- [5] Meng, Binghao, Lu Zhang, Fan Jin, Lu Yang, Hong Cheng, and Qian Wang. "Abnormal Events Detection Using Deep Networks for Video Surveillance." In *International Conference on Cognitive Systems and Signal Processing*, pp. 197-204. Springer, Singapore, 2016.
- [6] Peng, Qiwei, Gongyi Hong, Min Feng, Yuan Xia, Lei Yu, Xu Wang, and Yang Li. "Off-position detection based on convolutional neural network." (2016).
- [7] Sapana, Miss, K Mishra, and K S Bhagat. 2015. "A Survey on Human Motion Detection and Surveillance." *International Journal of Advanced Research in Electronics and Communication Engineering* 4 (4).
- [8] Automated Daily Human Activity Recognition for Video Surveillance Using Neural Network Mohanad Babiker<sup>1</sup>, Othman O. khalifa<sup>1</sup>, Kyaw Kyaw Htike<sup>2</sup>, Aisha Hassan<sup>1</sup>, Muhamed Zaharadeen<sup>1</sup> <sup>1</sup>Department of Electrical and Computer Engineering International Islamic University Malaysia <sup>2</sup>Department of Information Technology UCSI University, Malaysia
- [9] An Approach of Understanding Human Activity Recognition and Detection for Video Surveillance using HOG Descriptor and SVM Classifier Jagadeesh B Assistant Professor, Department of ECE VVCE, Mysuru jagadeesh.b@vvce.ac.in Meghana M N Department of ECE VVCE, Mysuru.