

# Innovative Approach to Detect Human Actions using Feature Extraction and Classification

Mrs. Bhavyashree D S

M.Tech Student

Electronics and Communication Engineering  
DECS, VTU PG Center, Mysuru

Mrs. Meghashree A C

Assistant Professor

Electronics and Communication Engineering  
DECS, VTU PG Centre, Mysuru Karnataka, India

**Abstract**— Image processing is one of the quickly developing area in the domain of software engineering and innovation. Human activity discovery is a PC innovation identified with PC vision and picture preparing that manages identifying the examples of human activity of certain class in recordings. An all around explored space of activity discovery incorporates person on foot recognition. Activity recognition has applications in numerous regions of PC vision including picture recovery and video reconnaissance. Activity acknowledgment is a procedure of finding and distinguishing activities in a video arrangement. You can recognize activity utilizing the appearance based techniques, for example, Edge coordinating, Divide-and-overcome search, Gray scale coordinating, Gradient coordinating, Histograms of open field reactions and Large model bases. We likewise utilize dark scale transformation strategy to know the estimation of every pixel in an example. In our application the foundation deduction additionally assumes a significant job so as to identify the frontal area activities. Foundation deduction is where in a picture's closer view is extricated for additional preparing. It is additionally used to distinguish moving activities in recordings from static cameras. Recognizing the moving activity from the contrast between the current casing and a reference outline is called as foundation deduction. The GLCM and multi SVM technique is utilized to acquire the portrayal of activity successions. From that point forward, strategy misuses the GLCM include extraction way to deal with separate highlights data from the edges. At long last, the multi bolster vector machine is applied to total the square highlights, which is then joined with the outrageous learning machine classifier.

**Keywords**—GLCM, SVM,

## I. INTRODUCTION

Image processing is one of the quickly developing areas in the domain of software engineering and innovation. Human detection and acknowledgment is one of the sub-spaces of Image processing. Human activity detection: It is a computer innovation identified with computer vision and image handling that manages distinguishing examples of semantic Humans of a certain class (such as people) in digital images and recordings. Human action detection: In PC vision, it is the undertaking of finding and recognizing Human activities in a picture or video stream. It manages properties or characteristics of the Human, example: human body shape, shading, surface data and so forth. Gray scale computerized picture is a picture wherein the estimation of every pixel is a solitary example, that is, it conveys only intensity information. Red, Green and Blue hues are separated from the image present. Pixel esteems are obtained for each point in the picture. Human activation recognition

has a wide scope of utilizations, for example, intelligent video observation and home monitoring, video stockpiling and recovery, smart human-machine interfaces and identity recognition. Human activity recognition covers many exploration subjects in computer vision, remembering human recognition for video, human posture estimation, human tracking and investigation and comprehension of time arrangement information. It is additionally a difficult issue in the field of computer vision and Artificial intelligence. At present, there are many key issues in human activity recognition that stay unsolved.

## II. LITERATURE SURVEY

Recently, propelled by the exceptional achievement of deep learning strategies in image categorization tasks, approaches dependent on the deep convolutional systems which intend to gain high level representations directly from training information have been received for action recognition. Space-time based strategies, for example, space-time volumes, spatio-transient highlights, and directions have been generally used for human activity acknowledgment from video groupings caught by customary RGB cameras. Cuboid descriptors were utilized for activity representation. Based on skeleton based feature extraction can be divided into joint-based and body based method. There are some shortcoming limits the action recognition using skeleton feature, unreliable joint estimation and fails in case of self-occlusion. Depth based approach is emerging and challenging technique in computer vision by considering spatial continuity and time constraints of human motion information. It consists of spatial depth shape features and temporal joints features to improve classification. This paper employ an activity chart to demonstrate unequivocally the elements of the activities and a sack of 3D focuses to describe a lot of notable stances that relate to the hubs in the activity diagram. Furthermore, they propose a basic, yet successful projection based testing plan to test the pack of 3D focuses from the profundity maps Compared to the 2D outline based acknowledgment, the acknowledgment blunders were split. Show the capability of the pack of focuses pose model to manage impediments through reproduction. The possibility of 2D portrayal of activity video succession by joining the picture groupings into a solitary picture, called double movement picture to perform human action acknowledgment. Tried this strategy to 3D profundity maps utilizing MSR activity 3D informational collection by extricating the 3BMI projections.

### III. PROPOSED METHODOLOGY

#### A. block diagram

A methodology for human action acknowledgment in a video stream by exploiting the key poses of the human, and developing another new classification model. The spatio-temporal shape variations of the human outlines are represented by partitioning the key stances of the outlines into a fixed number of matrices and cells, a noise free depiction. The calculation of parameter of grids and cells prompts displaying of feature vectors. This calculation of parameter of frameworks and cells is additionally organized in such a way in order to save the time succession of the silhouettes. To classify these feature vectors, a hybrid model is proposed dependent on Support Vector Machine (SVM) Classifier. The adequacy of the proposed approach of action interpretation and classification model is tried over real time video sequence with ordinary camera (Mobile camera). The pixel esteems in 2D picture arrange. The advantage of utilizing these pictures is that it takes out the time reliance limitations and easily segregating the comparable portrayal exercises of various classes based on power of the pixel esteems. The key edges, which have huge vitality esteem when contrasted with the most elevated vitality estimation of the casing, are picked for additional preparing, and vitality of the edge is determined utilizing Equation. These key edges are kept in a planned succession regarding the most noteworthy vitality outline and by this course of action of key edges the spatial difference in the shape as for time is kept up. Calculation of these key edges is very strong in segregating diverse human movement because of their capacity to separate both spatial and temporal data. For various human actions the spatial dissemination of gradient portrayal at level-2 of 8-orientation bins, it very well may be seen that the spatial transmission vector for various actions are unique. Consequently, it tends to be presumed that the GLCM portrayal is discriminative and have capacity to perceive the different shapes.

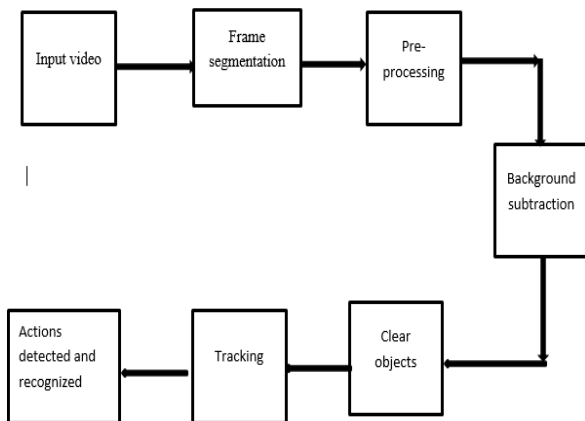


Fig. 1. Block diagram of proposed system

#### B. Flow chart

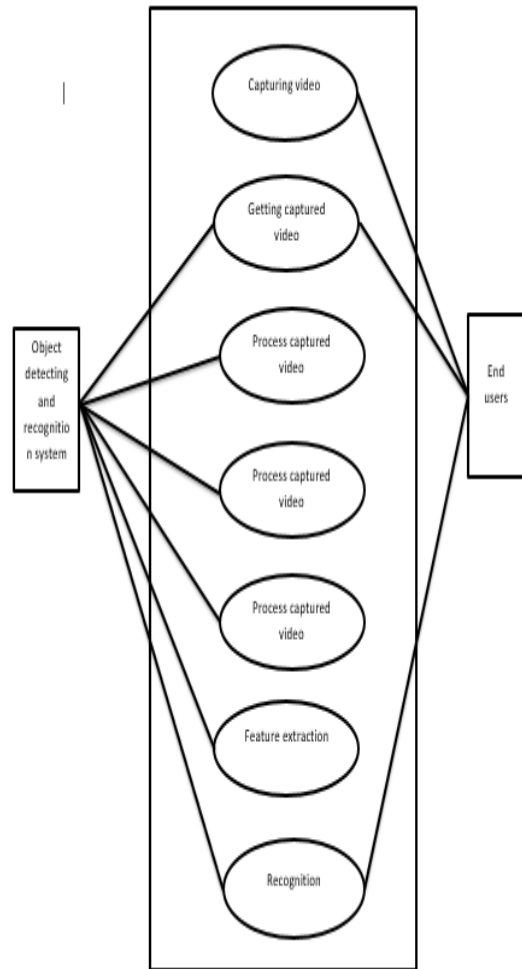


Fig.2 flow chart of proposed system

Speaking to the client association with the framework. At first the end client transfers the video from given set. At that point all the preparing in did inside the framework. The framework is liable for handling the video, movement division, position estimation, include extraction and perceiving the Humans in the video. Subsequent to handling all the previously mentioned advances the framework must have the option to show the recognized Human activities. Here the client is the on-screen character who communicates with the framework.

### IV IMPLEMENTATION

Frame Segmentation is the process of identifying components of the image. Segmentation involves operations such as boundary detection, difference and thresholding. Frame separation is a method for isolating a Human development in a picture from a static background. The background picture is aggregated over various pictures, and the invariant estimations of every pixel are taken to be the foundation either after certain number of edges or as middle an incentive for the pixel.

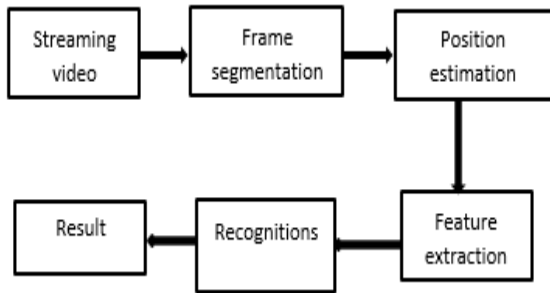


Fig. 3. overall data flow diagram

C. Architecture of human action detection

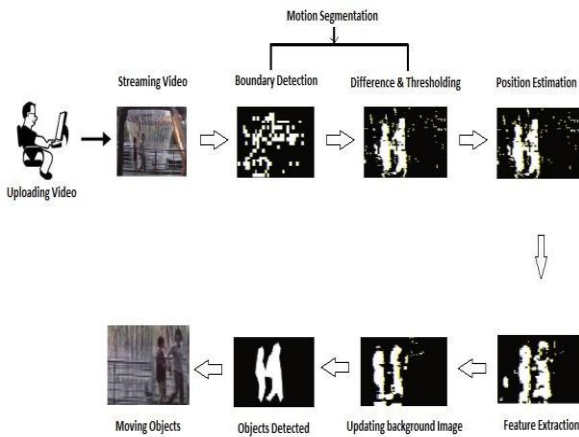


Fig. 4. Architecture of human action detection system

Engineering centres around thinking about a framework as a mix of a wide range of parts, and how they communicate with one another to create the ideal outcome. The emphasis is on distinguishing parts or subsystems and how they associate. At the end of the day, the attention is on what significant parts are required. The video stream recorded by the normal camera (mobile camera) is input to the system. When we provide input video segment which is of low resolution, the system does not accept the input. The image sequence must be clear and of minimum resolution of 240\*320 for the system to recognize. We are taken care about movement of the camera. The video sequence is divided into frames. Each frame represents the static images in the streaming video. The Human in each frame are identified later. A fixed set of frames are obtained depending on the length of the video. The input video's frame rate is reduced to 880frames to minimize the difficulty in processing large number of frames. It interns reduces the runtime or processing time. Pre-processing is the technique used to enhance the image / image sequence prior to computational processing. It improves the input image clarity, removes unwanted distortion, outliers and standardizes the data. We are using median filter to filter the noise and distortion by sorting all pixel values from the window into numerical order, then replaces the pixel with middle (median) pixel value. Background subtraction is a technique in image processing wherein an image's foreground is extracted

for further processing. Generally an image's region of interest is Humans in its foreground. Background is eliminated in order to highlight Human of interest. Foreground Human is filtered. Clear differentiation between foreground and background of the input video are obtained. The region and area of the images are identified using Blob analysis. The noise is removed in order to achieve a clear foreground Human. This enables us to identify the given Human more accurately. Here, the Human boundary is clearly defined by removing extra region around the Human. Video tracking is the process of locating movements in Human. It associates target Human in consecutive video frames. Human location is identified first. Based on the direction of human movement and changes in spatio temporal of the consequent frames, the actions are identified by applying GLCM feature extraction technique.

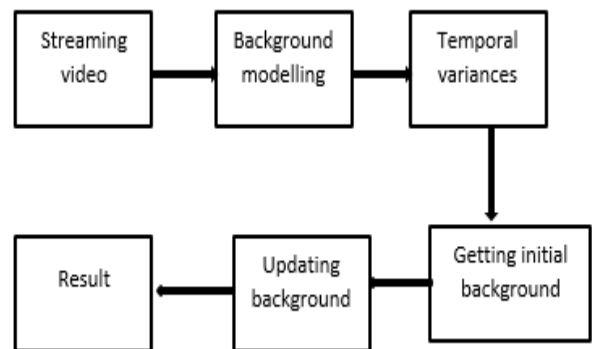


Fig.5. DFD of action recognition

Human activity acknowledgment assumes a significant job to perceive the moving Human body in succession of frames. A human body part has a particular component like shading, surface, edges, shape, and movement. Human acknowledgment should be possible by utilizing Feature-based and Appearance based methods.

V. RESULT AND DISCUSSION

Accuracy is to describe the system efficiency. It's a method that measures the percentage of error. It gives an idea on how a measured value is close to real value, is the degree of closeness of measured value. Accuracy can be calculated by using the relationship

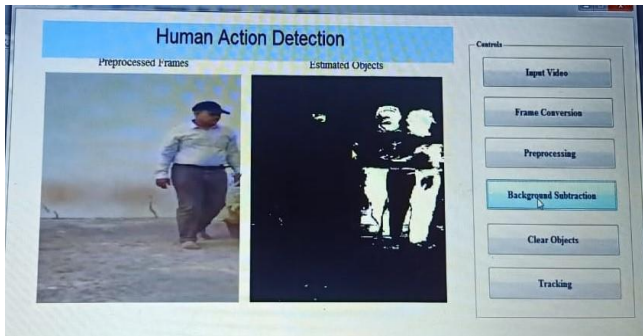
$$accuracy = \frac{\text{No. of frames detected correctly}}{\text{total number of frames}}$$

Sl. No	Action detected	Accuracy(%)
1	Waving hand	99%
2	Bending exercise	70%
3	Jogging	88%
4	Seeing phone	77%
5	Talking phone	88%
6	Throwing ball	56%
7	unknown	99%

The average recognition accuracy is calculated as

$$\text{Average Recognition Accuracy} = \frac{\text{sum of each actions accuracy}}{\text{total number of actions}} \times 100 \%$$

AVR= 82.42%



## VI. CONCLUSION

In this methodology the action recognition is finished by choosing the key poses of human body shape and further to represent these outline as a arrangement of grids and cells. The issue of less recognition rate under the variant natural conditions has been tended to by utilizing: (a) Accurate human body extraction through surface texture based background deduction method; (b) Simple and effective illustration of human figures by means of grids and cells. (c) An effective fusion classification model of "SVM".The "SVM" provides the least classification or taxonomy error. The overall performance of the proposed approach is seen as nearly increasingly powerful and more

effective. The parameters used for feature representation are simple and the calculation is easy.

## ACKNOWLEDGEMENT

We express our sincere thanks to teaching and Nonteaching staffs of Electronics and Communication Engineering department, VTU PG Center, Mysuru, Mysuru for suggestion and support.

## REFERENCES

- [1] Schudt, C., Laptev, I., Caputo, B.: Recognition human activities: a nearby SVM approach. Procedures of IEEE International Conference on Pattern Recognition, vol. 3, pp. 32–36, Cambridge, UK (2004)
- [2] Dollar, P., Rabaud, V., Cottrell, G., Belongie, S.: Behavior acknowledgment by means of inadequate spatio-worldly highlights. IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, pp. 65–72, Beijing, China (2005)
- [3] B. Liang and L. Zheng, "3D movement trail model based pyramid histograms of arranged angle for activity acknowledgment," in Proc. 22nd Int. Conf. Example Recognit., Aug. 2014, pp. 1952–1957.
- [4] . Chen, M. Liu, H. Liu, B. Zhang, J. Han, and N. Kehtarnavaz, "Multitemporal profundity movement maps-based neighborhood twofold examples for 3-D human activity acknowledgment," IEEE Access, vol. 5, pp. 22590–22604, 2017.
- [5] X. Ji, J. Cheng, and W. Feng, "Spatio-worldly cuboid pyramid for activity acknowledgment utilizing profundity movement arrangements," in Proc. eighth Int. Conf. Adv. Comput. Intell., Feb. 2016, pp. 208–213.
- [6] ] N. Dalal and B. Triggs, "Histograms of situated slopes for human discovery," in Proc. Int. Conf. Comput. Vis. Example Recognit. (CVPR), vol. 1, Jun. 2005, pp. 886–893.