International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 02 | Feb 2019 www.irjet.net p-ISSN: 2395-0072

Review on Human Action Detection in Stored Videos using Support Vector Machine

Dr. Rekha B Venkatapur[1], Sanket S Kulkarni[2]

[1]Professor and Head Department of Computer Science and Engineering, KSIT , Bangalore
[2] M. Tech Student, Department of Computer Science and Engineering, BNMIT, Banglore, Karnataka, India

---

Abstract - It is useful to detect human activities or events, counting people in crowd, identifying the vehicle number and many more. Analyzing a video for activity detection is tedious as there are many factors which contribute to the performance of the detection system. Major issue is the quality of the video. Today advances in technology has made video recording possible from a tiny smart phone to satellite captures.

The major applications of Human Activity Detection vary from Content-based Video Analytics, Robotics, Human-Computer Interaction, Human fall detection, Ambient Intelligence, Visual Surveillance, and Video Indexing etc. This paper collectively summarizes and deciphers the various methodologies, challenges and issues of Human Activity Detection systems. Variants of Human Activity Detection systems such as Human Object Interactions and Human-Human Interactions are also explored.

Key Words: **Support Vector Machine (SVM), Feature Detector, Classification, Feature descriptor**

1.INTRODUCTION

Video analytics is the analysis process in large volumes of real time or stored videos. It is useful to detect human activities or events, counting people in crowd, identifying the vehicle number and many more. Analyzing a video for activity detection is tedious as there are many factors which contribute to the performance of the detection system. Major issue is the quality of the video. Today advances in technology has made video recording possible from a tiny smart phone to satellite captures. Taking into consideration the quality of video as a major concern in this paper emphasis is on improving the quality of video and to detected the activity of human. The activity considered is limited to walking, jogging and running. Appropriate feature extraction and selection processes will be inculcated to implement the Support Vector Machine to detected the activity. It is suitable for creating set of hyper planes in an n dimensional space dividing the samples in to the regions with maximum margin. Video formats to be considered in the paper would be .avi, .mp4 and .mpg. The major Steps involved in activity detection are as follows.

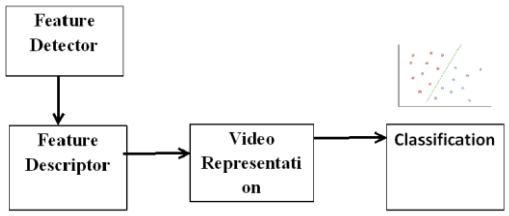

Fig -1 : Steps involved in activity detection

1. Feature Detector - Detect interest points(actions) in video .
2. Feature Descriptor - Encodes information in areas of interest points(SIFT-Scale Invariant Feature Transform)
3. Video Representation – Video format which defines sequence, structure and content of frame
4. Classification – Uses large number of training samples to train the classifier

Numerous attempts have been made in this field to automatize video surveillance but each and every approaches has its own pros and cons. Table 1 shows the spectrum of such approaches.On the basis of prior knowledge and human involvement in the learning process, the research in human activity detection can be categorized as supervised, unsupervised and semi supervised.

2.1 Supervised Learning

In this type of learning, a number of models of normal or abnormal behavior are built based on the labeled training samples. Video samples which do not fit any model are classified as abnormal. But this approach is limited to only events that are well defined and would require sufficient training data. However, real world video samples would mostly contain events that are not well defined and such events are rare and hence sufficient training samples are not available.

2.2 Unsupervised Learning

In this type of learning, a number of models of normal or abnormal behavior are built based on the labeled training samples. A video sample which does not fit any model is classified as abnormal. But this approach is limited to only events that are well defined and would require sufficient training data. However, real world video samples would mostly contain events that are not well defined and such events are rare and hence sufficient training samples are not available.

3. Methodology

Table 1 shows a brief summary of latest research works in the field of human behavior detection.

Title	Details	Constraint
Human Action Detection using SVM and Spatio-Temporal Features using 3D CNN	3D CNN approach is used to extract spatio-Temporal Features.	Number of CNN layers and resolution of input frames
Video Based Action Detection and Detection Human using Optical Flow and SVM Classifier	Video based human action detection is addressed on KTH dataset	Detection of human activities is a tedious task for real time videos
Human Action Detection by Learning Spatio-Temporal Features With Deep Neural Networks	To extract spatio-temporal features, Histogram of 3D Oriented Gradients and Histogram of Optical Flow (HOF) have been proposed	DNN-based method for action detection in video just using RGB data.
Human action detection using RGB data	To detected human activities. They use the Spatio-Temporal Interest Point (STIP) for detection of the important change in the image.	Specific action datasets as Weizmann, KTH, MSR action are considered.
Dense trajectories and motion boundary descriptors for action detection	Video representation based on dense trajectories and motion boundary descriptors.	Extract features aligned with the trajectories to characterize are restricted to few parameters
Human Action Segmentation and Detection Using Discriminative Semi-Markov Models	Human action understanding is to jointly segment and detected human actions from an unseen video sequence.	

4. Conclusion

In this study we used support vector machines approach for human action detection task. We propose to use a 3D CNN approach in order to extract spatial and temporal features from adjacent video frames with 80x60 resolution. The proposed architecture is trained and evaluated on KTH action detection dataset and achieved a good performance. As a future work, we are planning to use a weighted ensemble learning approach which integrates both of support vector machines and logistic regression in order to classify human actions from 3D CNN based extracted features. Moreover, a genetic algorithm based approach will be used to optimize the weights of the ensemble learner.

REFERENCES

[1]Human action detection using support vector machines and 3D convolutional neural networks , Majd Latah ,International Journal of Advances in Intelligent Informatics
ISSN: 2442-6571 Vol. 3, No 1, March 2017, pp. 47-55.

[2] Video Based Action Detection and Detection Human using Optical Flow and SVM Classifier Jagadeesh B, Chandrashekar M Patil , IEEE International Conference On Recent Trends In Electronics Information Communication Technology, May 20-21, 2016, India

[3] Human Action Detection by Learning Spatio-Temporal Features With Deep NeuralNetworks,Lei Wang , Yang Yang Xu, Jun Cheng et.,Al (Member, IEEE) Special Section On Cloud-Based Robotic System For Intelligent Systems IEEE Transaction Paper March 19, 2018

[4] Human action detection using RGB data ,Amel Ben Mahjoub and Mohamed Atri,2016 11th International Design & Test Symposium (IDT),IEEE.

[5] Dense trajectories and motion boundary descriptors for action detection Heng Wang, Alexander Kläser, Cordelia Schmid, Cheng-Lin Liu, International Journal of Computer Vision.

[6] Human Action Segmentation and Detection Using Discriminative Semi-Markov Models ,Q. Shi, L. Cheng, L. Wang, and A. Smola, International Journal Computer Vision (2011) .

[7] S. Ji, W. Xu, M. Yang, and K. Yu, ‘‘3D Convolutional Neural Networks for Human Action Detection’’, IEEE Transactions on pattern analysis and machine intelligence’’, vol. 35, no. 1, pp. 221-231, Jan. 2013.

[8] A. Klaser, M. Marszalek, and C. Schmid, ‘‘A spatio-temporal descriptor based on 3D gradients’’, in Proc. of the British Machine Vision Conference BMVC'08, 2008, pp. 1-10.

[9] I. Laptev, and T. Lindeberg, ‘‘Space-time interest points’’, in Proc. of the Ninth IEEE International Conference on Computer Vision ICCV'03, 2003, pp. 432-439.

[10] H. Wang, A. Klaser, C. Schmid, and C.L. Liu, ‘‘Dense trajectories and motion boundary descriptors for action detection’’, International Journal of Computer Vision. 103, pp. 60-79, May. 2013.

[11] C. Schuldt, I. Laptev, and B. Caputo, ‘‘Recognizing human actions: A local SVM approach’’, in Proc. of the 17th International Conference on Pattern Detection, 2004, pp. 32–36.

[12] A. Fathi and G. Mori, ‘‘Action detection by learning mid-level motion features’’, in Proc. of IEEE Conference on Computer Vision and Pattern Detection (CVPR), 2008, pp. 1-8.

[13] S. Nowozin, G. Bakir, and K. Tsuda, ‘‘Discriminative subsequence mining for action classification’’, in Proc. of the IEEE International Conference on Computer Vision (ICCV), 2007, pp. 1-8.