International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 07 | July 2020 www.irjet.net p-ISSN: 2395-0072

Development of an Algorithm for Action Detection in a Image

Priya Sunil Mehare[1], Snehil Jaiswal[2]

[1]M.Tech Student, GHRU, Amravati
[2]Professor. GHRU, Amravati

Abstract – *In this proposed model I proposed a simple and efficient method to detect human body recognition the proposed method is to rely the interest between frame and human action. Be calculated on the frames of a video with different different interest points using surf algorithm also be used support vector machine To detect multi-classification truth as a result is approaching towards more accuracy of other method such as SIFT , KTM. This kind of system requires every part of body reorganization with predefined models but SURF work with some untrained inputs also.*

Key Words: SURF, Open CV, Action Recognition. Python, Pattern Detection

1. INTRODUCTION

The Artificial intelligence is the simulation of human intelligence processes by machines, especially computer systems. These processes include learning (the acquisition of information and rules for using the information), reasoning (using rules to reach approximate or definite conclusions), and self-correction. Machine learning is a subfield of Artificial intelligence (AI). Machine learning is a field of computer systems the ability to "learn" i.e. progressively improve performance on a specific task with data, without being explicitly programmed. Machine Learning is a scientific discipline that is concerned with design and development of algorithms that allow computers to learn based on data. The focus of machine learning is to automatically learn to recognize complex patterns and make intelligent decisions based on data.

Human action recognition, as one of the most important topics in computer vision, has been extensively researched during last decades due to its potential diverse applications. Action recognition can be done by methods such as supervised and unsupervised. In supervised method, the computer is provided with example inputs that are labelled with their desired outputs. In the unsupervised method, the data is unlabelled. As the technology is rising up the use of internet and smart phone are increased. The action recognition in the personal videos has become an important research topic due to its wide applications such as automatic video tracking and video annotation. The videos that are uploaded on web produced by the users these videos will have certain camera shake and disturbance in videos so it is quite challenging task to recognize the action in such videos and it is very difficult to collect the labelled videos every time for real-world. In Order to enhance the performance and accuracy in videos we propose semi-supervised learning to leverage unlabeled data. The proposed adaptation method is different from the existing methods that can adapt the knowledge between domains that are in different feature spaces. In this method the performance and the accuracy is increased and flexibility is also increased

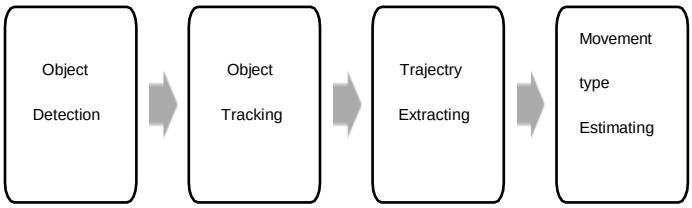

Fig -1: Main Steps of the Proposed Action Detection System

One of the ultimate goals of artificial intelligence research is to build a machine that can accurately understand humans' actions and intentions, so that it can better serve us. Imagine that a patient is undergoing a rehabilitation exercise at home, and his/her robot assistant is capable of recognizing the patient's actions, analyzing the correctness of the exercise, and preventing the patient from further injuries. Such an intelligent machine would be greatly beneficial as it saves the trips to visit the therapist , reduces the medical cost , and makes remote exercise into reality. Other important applications including visual surveillance, entertainment, and video retrieval also need to analyze human actions in videos. In the center of these applications is the computational algorithms that can understand human actions. Similar to human vision system, the algorithms ought to produce a label after observing the entire or part of a human action execution.

2. RELATED WORK

Mona M Moussa, planned "An increased methodology for human activity recognition". during this paper Associate in Nursing unsupervised methodology was used for the creation of dataset here the unsupervised means that making the information set with untagged videos and SIFT (sift invariant feature transform) was wont to extract the options from the pictures and for comparison. the most disadvantage is SIFT doesn't work well once there's lightning changes and once pictures area unit revolved, once the pictures area unit blurred.

Darshana Mistry and Asian Banerje planned "The Comparison between the SIFT and SURF".SIFT is employed for locating individualism options. during this paper it's same that the SURF is thrice higher than that of SIFT owing to victimization the integral image and box filters. Here the SIFT can take longer to extract the options when put next to SURF

Qing Lei has planned "Multi Surface Analysis for human activity Recognition in Video". during this paper the options area unit extracted by horizontal and surface the SVM (support vector machine) and NBNN (naive primarily based nearest neighbour) that area unit used for classification and here the feature extraction method takes longer.

Li Liu, planned Learning "Spatio-Temporal illustration for Action Recognition: A Genetic Program Approach". which suggests the programs area unit encoded as set of genes; here the options area unit extracted by changing the pictures to 3Dimages.so it's terribly time intense.

Exactly how this system works we basically used some model classifications of two layers in which each layer will detect action recognization purpose using self algorithm now when we work on such kind of algorithms need models here we are using only human body actions in binary form show all the models are work in a binary digits and whenever a object moves system will detect that its position with respect to body and the rate of change of frame with respect to body according to the predefined calculation this system will work more accurately

System is also follow methods of image processing and vector mechanism to recognise small changes of human body parts with respect to main frame but in this project while making such kind of math algorithms I need to work on sum of object with appropriate light and appropriate body portion.

3. METHODOLOGY

Detection of attention-grabbing points, feature description for the detected purpose, building the codebook and eventually the classification. inflated attention rabbing purpose detection starting among the system is interested purpose detection where SURF is used to do and do that methodology, exploitation algorithmic program fine standardization the brink parameter is performed to control the quantity of interest purpose automatically in step with the number of details in each frame. the brink value determines the number of details the detector returns, therefore once the brink value is high only the very important interest points unit detected, whereas the weak interest purpose unit neglected. that the useful information is not lost. the advance achieved by adjusting the brink. it's obvious that whereas not using a threshold the quantity of extracted purpose is unbelievably high which they unit insignificant where most of them song among the background. Utilizing a threshold, only important|the many|the numerous} purpose unit detected whereas not the requirement for an additional segmentation step that represent important method overhead.

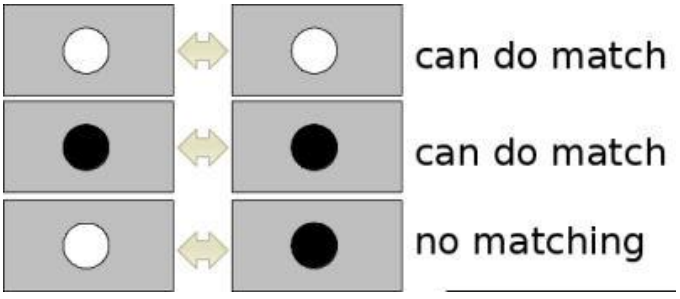

Fig -2: Template Matching

The projected frame work of image-to-video adaptation method is employed for acknowledge the action performed within the video that area unit taken from the user camera or personal videos. The action recognition in such videos is sort of difficult. Firstly, the user can transfer the video as input then that video is given to the videos splitter the video splitter can split the video into set of frames. The frames that area unit obtained area unit then given to comparison method here the frames area unit obtained within which the thing is gift which frames are compared with information|the info|the information} set that's gift within the data base. The informationset are created mistreatment SEMI-SUPERVISED approach and also the options are

hold on in data base for more method. we tend to use SURF rule for the feature extraction and comparison of the frames with the dataset. RANSAC rule is employed for the gap calculation of the frames and score base classification is finished supported most classification price. If the worth obtained is most then the action are recognized and displayed.

The system is very much beneficial to recognise suspected person using video from CCTV footage we found an issue and a real-time of problem that suspected criminals sometime change their face by plastic surgery but their body posture their walking style they are jogging this time they are running style is almost same so if we can work on sir we can able to recognise the body action and facial recognition to improve its productivity more

SURF mechanism is also advance of SBM and STIP which basically work on three Harish interest. Promo video in this method system detect three black combination which match exactly with parent image

SURF in OpenCV

Insaaf we need to use of incipient because image processing techniques can be e only but in in platform like open CV tensorflow now as we are more comfortable with open CP people for open Siri rather than tensorflow also sirf use 64 describe data transmission for each frame. That's why the whole the process is very much slow with respect to time and video duration. Hence we use short video for single action to take or detect it perfectly. Hands to set the threshold limit of a particular action we need to set pixel thresholding using Hessian thresholding. Where we keep thresholding of 404 single image of single action if the probability and prediction will more than 400 our approach 400 system indicate and able to recognise the predefined action such as clap wave run set walk squads.

4. RESULT

In this system video frames to find somebody's body actions with perennial space of interest that detected by facility with success trained our system with pre outline actions system can begin sleuthing frame comparison with new files and supply action recognised name. Action is solely recognised ahead read for higher result Vivo conjointly tried and prime view for action recognition victimization surf however its potency e is barely 2 hundredth thus within the viewing angle of zero degree we have a tendency to are becoming recognisation action potency regarding seventieth that is 100 percent additional then obtainable rule for action recognition

REFERENCES

[1] Liu, L. Shao, X. Li, and K. Lu," Learning spatio -temporal representations for action recognition: A genetic programming approach," IEEE trans. Cyber. vol. 46, no. 1, pp. 158-170, Jan. (2016)

[2] Hong-Bo Zhang*, Qing Lei, Bi- Neng Zhong , Ji-Xiang Du*, Jialin Peng, Tsung - Chih Hsiao and Duan-Sheng Chen" Multi-surface analysis for human action recognition in video" Zhang et al. Springer Plus vol. 65,pp.43 (2016)

[3] Darshana Minstry and Asim Banerjee, "Comparsion of Feature detection approaches: SIFT AND SURF". Cyber .vol.3,no.11.2017

[4] Mona M. Moussa, Elsayed Hamayed, Magda B. Fayek, Heba A. El Nemr "An enhanced method for human action recognition" Journal of Advanced Research .vol. 34 no.2,pp.234,(2015)

[5] Fathi A, Mori G. Action recognition by learning mid-level motion features. Comput Vision Pattern Recogn, CVPR IEEE 2008:1–8.

[6] Blank M, Gorelick L, Shechtman E, Irani M, Basri R. Actions as space-time shapes. Int Conf Comput Vision, ICCV IEEE 2005;2:1395–402.

[7] Ke Y, Sukthanka R, Hebert M. Efficient visual event detection using volumetric features. Int Conf Comput Vision, ICCV IEEE 2005;1:166–73.

[8] Sheikh Y, Sheikh M, Shah M. Exploring the space of a human action. Int Conf Comput Vision, ICCV IEEE 2005:144–9.

[9] Chen MY, Hauptmann AG. MoSIFT: recognizing human actions in surveillance videos. Technological report, CMU-CS-09-161, Carnegie Mellon University; 2009. p. 9–161.

[10] Schuldt C, Laptev I, Caputo B. Recognizing human actions: a local SVM approach. Int Conf Pattern Recogn, ICPR IEEE 2004;3:32–6.

[11] Csurka G, Dance C, Fan L, Willamowski J, Bray C. Visual categorization with bags of key points. ECCV International Workshop on Statistical Learning in Computer Vision 2004: 1–22.

[12] Gemert J, Geusebroe J, Veenman C, Smeulders A. Kernel codebooks for scene categorization. Proc Euro Conf Comput Vision, ECCV 2008:696–709.

[13] Lowe DG. Distinctive image features from scale-invariant keypoints. Int J Comput Vision 2004;60(2):91–110.

[14] Lin Z, Jiang Z, Davis LS. Recognizing actions by shapemotion prototype trees. Int Conf Comput Vision, ICCV IEEE. p. 1–8.

[15] Liu J, Shah M. Learning human actions via information maximization. Comput Vision Pattern Recogn, CVPR IEEE 2008:1–8.

[16] Bregonzio M, Xiang T, Gong S. Fusing appearance and distribution information of interest points for action recognition. Pattern Recogn 2012;45(3):1220–34.

[17] Niebles J, Wang H, Fei-Fei L. Unsupervised learning of human action categories using spatial-temporal words. Int J Comput Vision 2008;79(3):299–318.

[18] Sadanand S, Corso J. Action bank: a high-level representation of activity in video. Comput Vision Pattern Recogn, CVPR IEEE 2012:1234–41.

[19] Tran KN, Kakadiaris IA, Shah SK. Modeling motion of body parts for action recognition. British Mach Vision Conf, BMVC 2011.

[20] Kovashka A, Grauman K. Learning a hierarchy of discriminative space-time neighborhood features for human action recognition. Comput Vision Pattern Recogn, CVPR IEEE 2010:204