

Semi-Supervised Image-To-Video Adaptation for Video Action Recognition Using Speeded-Up-Robust Feature (Surf)

Y.Vamsi Krishna Pallavi S², Saagarika P S³

Associate Professor, Department of ISE, SKIT Engineering College, Chikkabanavara, Bangalore, India¹

U.G. Student, Department of ISE, SKIT Engineering College, Chikkabanavara, Bangalore, India²

U.G. Student, Department of ISE, SKIT Engineering College, Chikkabanavara, Bangalore, India³

ABSTRACT: Computer vision deals with how computers gain high-level understanding from videos or images. It seeks to automate tasks that human visual system can do. In computer vision, human actions are well explored due to its wide applications. It is concerned with automatic extraction, analysis and understanding of useful information from a single or sequence of images. The existing video action recognition method use supervised method for action recognition so that it is very difficult to collect the labelled videos that cover different types of actions. In such cases the performance of action is restrained and becomes excessively complicated. Fewer efforts were made to improve the performance of action recognition in videos. In this paper we propose an adaptation method which enhances the action recognition in the videos by adapting knowledge from images. Then we extend the adaptation method to the semi-supervised method that can control both the labelled and unlabelled videos and by using this adaptation method the performance of action recognition is improved and over-fitting can be prevented.

KEYWORDS: Adaptation, computer vision, action recognition, semi-supervised, labelled and unlabelled videos.

I. INTRODUCTION

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.

II. RELATED WORK

Mona M Moussa, proposed "An enhanced method for human action recognition". In this paper an unsupervised method was used for the creation of dataset here the unsupervised means creating the data set with unlabelled videos and SIFT (sift invariant feature transform) was used to extract the features from the images and for comparison. The main drawback is SIFT does not work well when there is lightning changes and when images are rotated, when the images are blurred.

Darshana Mistry and Asian Banerje proposed "The Comparison between the SIFT and SURF". SIFT is used for finding uniqueness features. In this paper it is said that the SURF is three times better than that of SIFT because of using the integral image and box filters. Here the SIFT will take more time to extract the features when compared to SURF

Qing Lei has proposed "Multi Surface Analysis for Human Action Recognition in Video". In this paper the features are extracted by horizontal and vertical surface the SVM (support vector machine) and NBNN (naive based nearest neighbour) which are used for classification and here the feature extraction process takes more time.

Li Liu, proposed Learning "Spatio-Temporal Representation for Action Recognition: A Genetic Program Approach". Which means the programs are encoded as set of genes; here the features are extracted by converting the images to 3D-images. so it is very time consuming.

III. METHODOLOGY

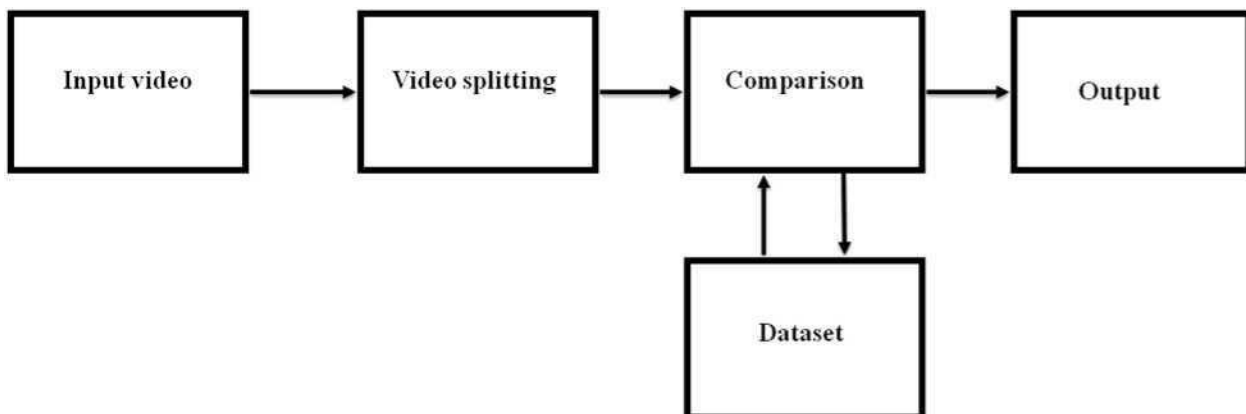


Figure: The frame work for image-to-video adaptation.

The proposed frame work of image-to-video adaptation process is used for recognize the action performed in the video which are taken from the user camera or personal videos. The action recognition in such videos is quite challenging. Firstly, the user will upload the video as input then that video is given to the videos splitter the video splitter will split the video into set of frames. The frames which are obtained are then given to comparison process here the frames are obtained in which the object is present and that frames will be compared with the data set that is present in the data base. The dataset will be created using SEMI-SUPERVISED approach and the features will be stored in data base for further process. We use SURF algorithm for the feature extraction and comparison of the frames with the dataset. RANSAC algorithm is used for the distance calculation of the frames and score base classification is done based on maximum classification value. If the value obtained is maximum then the action will be recognized and displayed.



Figure : Data sets for video action recognition.

IV. CONCLUSION

The main concentration is to develop a system that detects various human actions in the given video. To achieve good performance of video action recognition, we propose SURF as classifier and for extraction of features from the images and for comparison process there is no need for a separate classifier when SURF is used. In our experiments, we extract the knowledge learnt from the images can influence the recognition accuracy of videos. Experimental results show that semi-supervised adaptation method and SURF has better performance of video action recognition compared to the state-of-art methods and the performance of the SURF will be better even when there are few labelled training videos are available.

V. FEATURE ENHANCEMENT

The proposed system can only recognise the single action performed in the video but the multiple actions cannot be recognised. The future enhancement can be made to recognise the multiple actions in the video and improvements in accuracy can also be done

REFERENCES

- [1] Mona M. Moussa, Elsayed Hamayed, Magda B. Fayek, Heba A. El Nembr "An enhanced method for human action recognition" Journal of Advanced Research .vol. 34 no.2,pp.234,(2015)
- [2] L. 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)
- [3] 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)
- [4] Darshana Ministry and Asim Banerjee, "Comparson of Feature detection approaches: SIFT AND SURF". Cyber .vol.3,no.11.2017