JOURNAL OF CRITICAL REVIEWS

ISSN- 2394-5125 VOL 07, ISSUE 02, 2020

HUMAN ACTION RECOGNITION BASED ON MOTION DESCRIPTOR

Yogish Naik G R,
Associate Professor, Department of Computer Science and MCA,
Kuvempu University, Shivamogga, Karnataka, India.
ynaik.ku@gmail.com

ABSTRACT

A difficult challenge in computer vision is deciphering human activity from videos. The primary capability of smart video surveillance systems is the automatic identification of human actions in the recorded sequence and the tagging of such actions. The aim of recognition of human behavior is to recognize the activities and purposes of any number of things through a series of investigations into those actions and those of the objects' surroundings. With the emergence of vast quantities of human-centric video data due to technological advancements, the efficient identification of human action from huge video data has become a bottleneck in video processing.

Occlusion and pixels bending moment effects make it difficult for object detection algorithms to work properly, which increases complexity and erroneous margin in addition. In this paper, characteristics are extracted using Apache Spark using in-memory computation and the distributed environment's goal of action by humans recognition. Additionally, the local movement descriptor is utilized to derive features that compute the intensity of the pixel information between two frames, taking both motion and appearance into account. To distinguish human actions, the spark Machine Learning Library random forest is then used. To investigate the model's capacity for activity recognition, experimental tests are carried out, and the outcomes are contrasted with widely used Decision Trees for supervised categorization. In the comparative analysis, it was found that the suggested approach performed better than the other categorization methods.

INTRODUCTION:

Modern computer vision research is increasingly focusing on the study of human action detection, which has led to the development of several systems that consistently outperform in various disciplines. Its many applications play a key role in the world of computer vision and media research. The widespread usage of human action recognition in the automatic retrieval of films of a specific activity utilizing features has increased its popularity. Human motion recognition can assist us in retrieving movies of interest in this technological age where automation is crucial. It is particularly helpful in the fields of surveillance footage, human-computer interaction, healthcare systems, etc. The process of identifying specific activities from thousands of recordings will take a very long time.

Therefore, recognizing human action in videos is a highly beneficial assignment. The monitoring system, networking sites, YouTube, etc. all contribute to the enormous amount of videos that are produced each day. The computer vision sector has recently paid a lot of attention to the difficult subject of human action recognition. Its uses range widely, from increasing human-computer interface to using it to comprehend behavior for intelligent monitoring systems. Due to the intricacy of human movements, both spatial and temporal variances seen as a result of differences in the lengths of the many actions done, and the shifting spatial properties of human form during each action, tackling this problem presents a number of obstacles.

The ideal action representation should be capable of generalizing across changes in point of view, human physical appearance, and spatiotemporal alterations. Additionally, for robust action recognition, the action representations' descriptions have to be sufficiently rich. Global and local representations are the two main types of human activity representation. Global representations are capable of encoding a large amount of data, although they are more susceptible to environmental factors (such as noise and viewpoint). Local depictions are less environment-sensitive, but their correctness depends on the local attributes or interest point detectors used. The majority of the explanatory and discriminatory data should be encoded via a robust activity representation model while being less environment-sensitive.

LITERATURE SURVEY:

A parallel architecture built on the Spark was developed in 2018 by Bin Wu, Jinna Lv, and others [4]. It is built on the Spark and Hadoop cloud computing platforms to handle massive amounts of video data in parallel across distributed storage. The CaffeOnSpark framework in particular supports parallel deep learning models. application of parallel face recognition & search algorithms in accordance with the video's hierarchical feature. To quickly analyze a big number of movies, their algorithms are built on Spark. The effectiveness of the method to analyze large amounts of video data is finally demonstrated by the performance assessment trials of face recognition and character searching using a cloud cluster.

The deployment of the system on a spark-based parallel compute platform. A system to extract discriminative features was developed in 2016 by Jinna Lv, Bin Wu, and colleagues [1] in order to increase precision and boost the effectiveness of large-scale video analysis. They developed a method that extracts Multi-Feature based Parallel method (MFPS) for large-scale Near-Duplicate Videos Detection (NDVD) in order to meet the demands of increased precision and efficiency while also overcoming a very difficult work. To effectively depict the crucial information in a video during feature extraction, they blend both global and local features. Local -Maximal Occurrence -(LOMO) and Scale-Invariant Feature Transform (SIFT) are presented as the local features, while Color Names (CN) is used as a global feature. instead of a standalone device, a spark cluster is used for parallel implementations.

The results of the experiments show that their method performs well in terms of accuracy and productivity. The usage of a cloud-based framework to handle the intelligent analysis and storage of video data was suggested in 2015 by Weishan Zhang et al to utilize the Apache Hadoop and Storm platforms, respectively, for batch and real-time processing. In this study, the only application area for big data techniques to intelligently handle large-scale video information with sequential processing integrated with a rapid processing architecture is large-scale video processing, not any specific application domain. An illustration of how a system performs, stores data, and tolerates errors.

PROPOSED SYSTEM:

Action recognition is accomplished by batch processing video datasets, and video data from the local file system is transferred to HDFS. UT-Interaction data set, KTH data set, UCF-50 data set, and UCF sports event dataset are a few examples of data sets. RDD is given a video data set from HDFS. It includes a number of concurrently operating items. Each RDD is split up into several partitions that can be calculated on various cluster nodes. RDD data is immutable and distributed in nature. After performing frame extract from visual data and frame resizing in the RDD partition, frames are converted from RGB color space to grayscale.

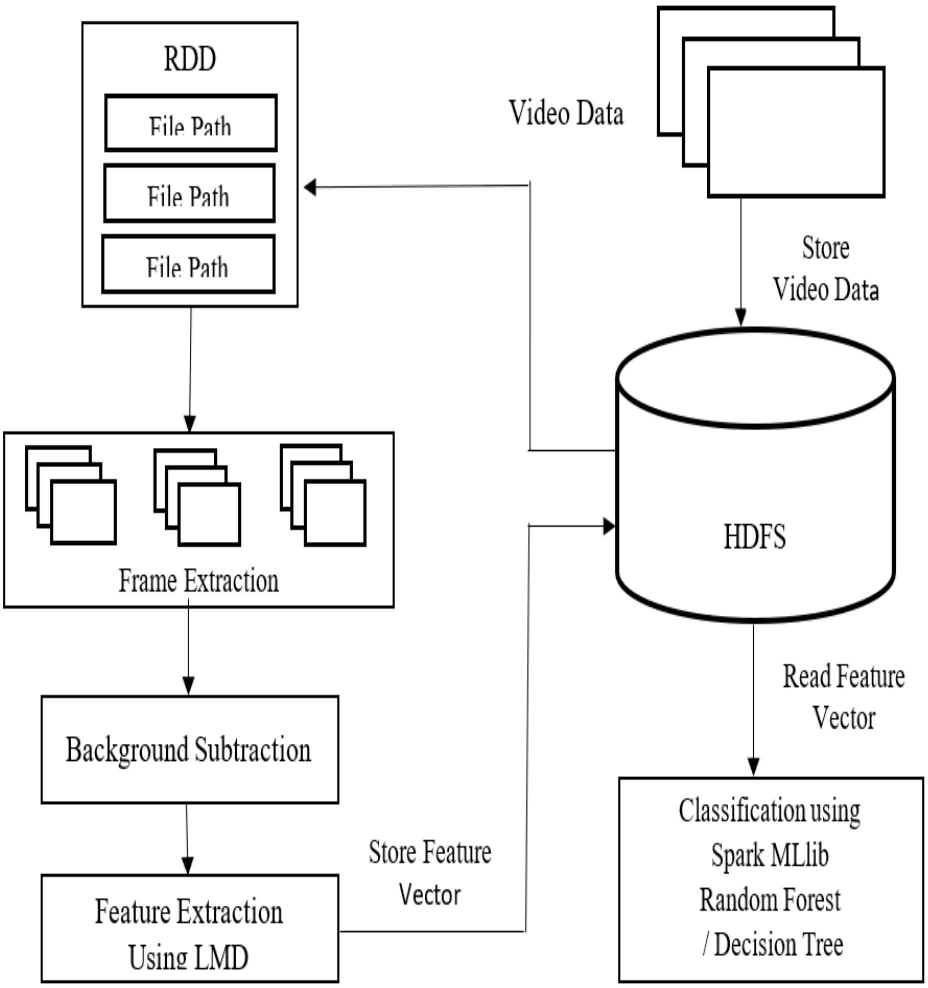

Fig. 1: Block diagram of the proposed framework

The "start-all.sh" shell script is used to launch the YARN daemons and HDFS daemons (Resource manager & Node manager), which are required to upload video data from the local file system to HDFS.

“start-YARNdaemon.sh” furthermore "start-dfsdaemon.sh"

And using the command "hadoop fs -put /video data path> / hdfs path>" or "hdfs dfs -put /video data path> / hdfs path>," data is transferred to the hadoop shared file system.

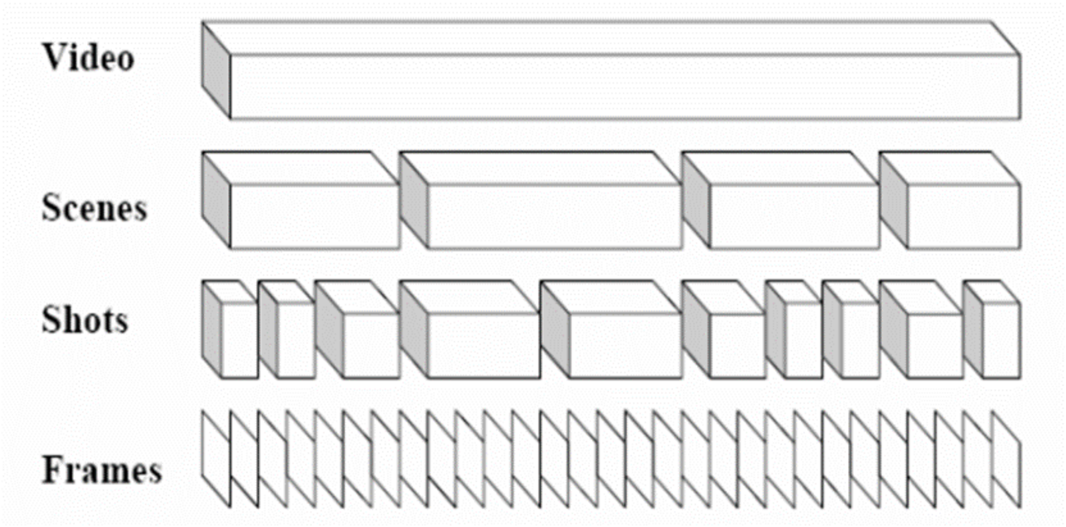

Fig. 2: Frame extraction from video

It is done to separate the backdrop from the foreground by subtracting the background from the foreground objects. Here, a person is the foreground object. In backdrop subtraction, the background portion of frames is removed. An application of the Gaussian probability distribution equation is made to remove backgrounds from an earlier frame.

When features are collected from frames using the best feature extraction techniques and placed back in HDFS, they include static texture features and dynamic data from frames. Spark MLlib then has these features. both random forest with a trained set of datasets, features are classified using gradient boosted trees. Fig. 1 contains the suggested system's block diagram.

Figure 2 illustrates how video frames are taken from the video and compared to one another to identify motion using the next and previous frames. The motion data from the prior and next frames is used to calculate the motion feature.

CONCLUSION

Diverse material was gathered and researched for the recognition of human action. To extract motion data from frames and perform action classification using Random Forest classifier & Decision Trees, an effective human action detection system called Local Movement Descriptor was developed and presented. Different benchmark datasets were obtained and saved in the distributed file system of Hadoop to test the functionality of the system. These films are divided into frames so that background subtraction can be used to distinguish between foreground objects using a combination of Gaussian methods. The implementation of motion description-based feature extraction and storage of the extracted features in the distributed file system of Hadoop.

These actions are classified using the Random Forest classifier. The decision tree classifier intended to be used instead of the Random Forest classifier. Additionally, the effectiveness of the Random Forest and Decision Trees classifiers are compared.

ISSN- 2394-5125 VOL 07, ISSUE 02, 2020

REFERENCES:

1. P Wang, W Li, C Li, Y Hou, "Action recognition based on joint trajectory maps with convolutional neural networks", Knowledge-Based Systems, 2018 – Elsevier

2. *Peng Zhou, Bingbing Ni, Cong Geng, Jianguo Hu, Yi Xu* ,"Scale-transferrable object detection" Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 528-537

3. Allah Bux Sargano; Xiaofeng Wang; Plamen Angelov; Zulfiqar Habib "Human action recognition using transfer learning with deep representations", 2017 International Joint Conference on Neural Networks (IJCNN)

4. C Chen, K Liu, N Kehtarnavaz , "Real-time human action recognition based on depth motion maps", Journal of real-time image processing, 2016 – Springer

5. Shasha Zhang, Weicun Zhang & Yunluo Li "Human action recognition based on multifeature fusion" ISC 2016: Proceedings of 2016 Chinese Intelligent Systems Conference pp 183–192

6. DK Vishwakarma, K Singh, "Human activity recognition based on spatial distribution of gradients at sublevels of average energy silhouette images" IEEE Transactions on Cognitive …, 2016 - ieeexplore.ieee.org

7. C. Chen, R. Jafari and N. Kehtamavaz, "Action recognition from depth sequences using depth motion maps-based local binary patterns", Proceedings of the IEEE Winter Conference on Applications of Computer Vision, pp. 1092-1099, 2015.

8. M. Mejdoub et al. Extending Laplacian sparse coding by the incorporation of the image spatial context, Neurocomputing (2015)

9. J. Cho et al. Robust action recognition using local motion and group sparsity JPR (2014)
10. H. Zhou, H. Hu, H. Liu, and J. Tang. Classification of upper limb motion trajectories using shape features. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 42(6):970–982, 2012.