International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

HUMAN ACTIVITY DETECTION SYSTEM

Kola Nithisha[1], Komal Singh Thakur[2], Komal Vitthal Kamthikar[3], M Sowmya[4]

Department of Computer Science and Engineering, Stanley College of Engineering and Technology for Women, Telangana, India

Abstract.

The purpose of this project is to develop a model that would take the image input from different sources and detect actions that are being performed in them by the people also commonly referred to as Human actions recognition. The project starts by analyzing the procedures and models that already exist to perform this activity along with the assessment of their advantages and disadvantages respectively. There were many approaches to build this project that mainly used sensors and various modelling procedures such as Convolutional Neural Networks (CNN), Recurrent Neural Networks(RNN), Deep Neural Networks(DNN) and few Image Processing techniques. The goal is to come up with a model that is able to detect a selection of activities on which the model was trained during the development phase. The designed model helps to overcome the disadvantage of the previously built models that relied on the use of sensors. With the elimination of sensors, we eliminate the cost associated with respect to the sensors and the errors that can be caused due to the use of sensors. The vision-based HAR(human activity recognition) has produced fruitful findings with the advantage of high optical sensor resolution and the rapidly evolving computer vision techniques.

Keywords: human action recognition, CNN, elimination of sensors, vision-based HAR

1. Introduction

1.1 About Paper

Human activity recognition (HAR) aims to recognize activities from a series of observations on the actions of subjects and the environmental conditions. The vision-based HAR research is the basis of many applications including video surveillance, health care, and human-computer interaction (HCI).

The purpose of this project is to develop a model that would take the image input from different sources and detect actions that are being performed in them by the people also commonly referred to as Human actions recognition. The project starts by analyzing the procedures and models that already exist to perform this activity along with the assessment of their advantages and disadvantages respectively. There were many approaches to build this project that mainly used sensors and various modelling procedures such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Deep Neural Networks (DNN)

and few Image Processing techniques.

Research on recognition of human activity (HAR) has progressed considerably over the last decade. Applications that are successful include surveillance, smart home, video analytics, autopilot. HAR's objective is to identify the behavior of a user so that computer systems can proactively support the user.

HAR consists of two main classes: vision and sensor. The vision-based HAR has produced fruitful findings with the advantage of high optical sensor resolution and the rapidly evolving computer vision techniques. Different studies have been conducted in the intelligent environment to observe the activities of the human.

The detection of the physical activities by different such sensors and recognition process is a key topic of research in wireless, smartphones and mobile computing. Human Activity recognition Systems is able to perform different tasks and recognize the multi day to day actions performed by humans which can be either simple activities like sleeping or the complex activities like running and eating etc

1.2 Objectives of the Paper

The main objectives identified which illustrate the relevance of the topic are listed out below..

1. To prepare a dataset(here:UCF50) consisting of images for 50 activities namely running, pushups, playing piano, tai-chi, swimming, basketball, jump rope, horse race etc.
2. The dataset is segmented and the relevant data is annotated with labels of the above-mentioned activities.
3. Data is pre-processed and redundant and irrelevant data is edited out.
4. The development of an optimized model is carried out that uses implement the LRCN Approach by combining Convolution and LSTM layers in a single model to segment the dataset images into their respective action labels.
5. This is done following a supervised learning procedure for training the model. The threshold is also set to determine the probability of correctness in the determination of the action being performed.
6. Demonstrating the results.

1.3 Scope of the Paper

Systems like **Human Activity Recognition systems** is quite useful and effective for solving a variety of application, whether it would be surveillance or monitoring, or aiding the elderly and blind people etc. This not only provide additional comfort to the end-users but also can be deployed into different Organizations in order to reduce the employ workload. The model shows good results on video streams while performing

decently on image data.

Activity Recognition system are of great importance in modern days due to the convenience and problems which the system offers and solves. Need of Activity recognition for monitoring and surveillance, video segmentation etc is of growing demand in which this system can greatly help.

2. Literature Survey

2.1 Existing System

Research on recognition of human activity (HAR) has progressed considerably over the last decade. Applications that are successful include surveillance, smart home, video analytics, autopilot. HAR's objective is to identify the behaviour of a user so that computer systems can proactively support the user.

Since HAR is a significant subject of research in recent years, numerous investigations have been published.

The existing system was manual where a person had to sit in front of a monitor to monitor and guide human activities, it was hectic, time consuming and costlier system and was prone to human errors and negligence. Further some systems started using sensor data to recognize human activities but they were needed to be worn by the user which limited the scope of activity recognition in open environment in general.

2.2 Proposed System

Unlike the existing system, the proposed system takes input in the form of video and image to recognize the activity being performed in it using computer vision techniques. It is much faster and a cost-effective solution. It uses deep learning to recognize the activities. This system can be used, incorporated or expanded further to cover a wide range of applications. Hence, it acts as a base system for various applications and tasks. Moreover, it can reduce the need of additional staff for entering data. Thereby, reducing the cost of the companies considerably.

The proposed system processing is done at fixed intervals continuously for the entire video and the trained model determines which action is being performed in the video as well as an image.

The output of human activity is displayed when the activity is detected and determined by the system. It keeps on changing based on the activity being performed at that instance of the video being played. The existing system determine the following activities:
Hula hoop, horse race, pushups, basketball, swimming, jumping, tai-chi, piano playing, and many other actions performed by humans.

Unique Features of the Proposed System:

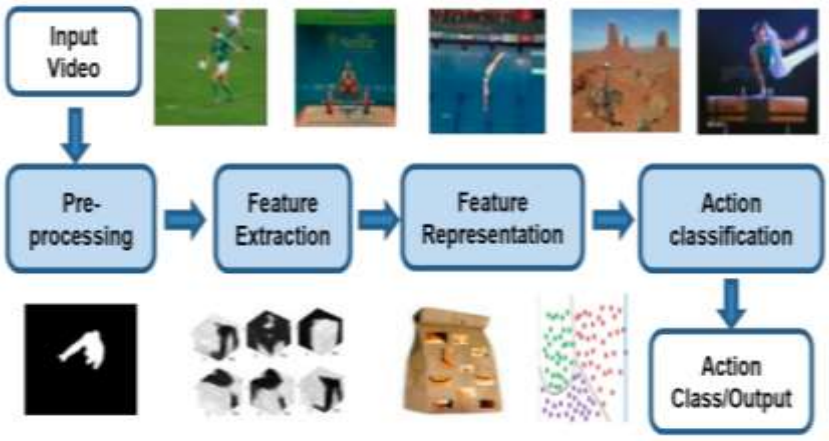
The unique features of the Human Activity Recognition system are listed as follows:

- The system analyses activity from both image and video data. It recognizes activities from youtube videos. It is very easy to use and understand.
- It is sort of a base system which can be used in various number of applications.
- It uses CNN-LSTM layers in a single model to solve the solution.
- It works with a great accuracy on videos dataset to recognize activity in the video. It generates words describing the activity on the image dataset with decent accuracy.

Firstly a set of activities to detect are determined. This is followed by the process of collection of the data collection for these activities. The collected data is then cleaned and pre-processed while making sure to remove redundant or unnecessary data. Further steps include the extraction of key features from the data to help build the model to detect human activities in real-time. This data is labelled and used to train the model training in the supervised learning process. The model is then deployed on the camera feeds and the activities performed are determined by the trained model.

So, we have developed a human activity recognition model using **UCF50 dataset** and modelling procedure **Convolutional Neural Networks (CNN) and Long Short-Term Memory** (**LSTM).**

The goal of proposed system is to come up with a model that is able to detect a selection of activities on which the model was trained during the development phase. The designed model helps to overcome the disadvantage of the previously built models that relied on the use of sensors. With the elimination of sensors, we eliminate the cost associated with respect to the sensors and the errors that can be caused due to the use of sensors.



3. Proposed Architecture

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

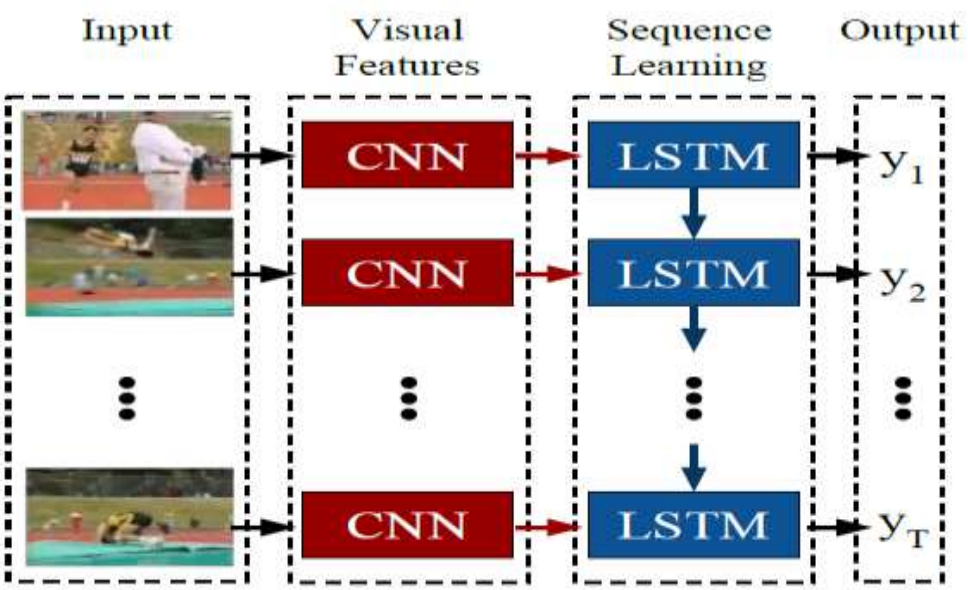


Fig 3.1: Block diagram of Human Activity Detection System

Fig 3.2: CNN-LSTM Model used for Human Activity Detection System

4.

Implementation

4.1 Algorithm

Step 1: Download and Visualize the Data with its Labels

Step 2: Pre-process the Dataset:

Step 2.1: Create a Function to Extract, Resize & Normalize Frames

Step 2.2: Create a Function for Dataset Creation

Step 2.3: Now we will utilize the function create_dataset() created above to extract the data of the selected classes and create the required dataset.

Step 2.4: Now we will convert labels (class indexes) into one-hot encoded vectors.

Step 3: Split the Data into Train and Test Set

Step 5: implement the LRCN Approach

Step 5.1: Construct the Model

Step 5.2: Compile & Train the Model

Step 5.3: Plot Model's Loss & Accuracy Curves

Step 6: Test the Best Performing Model on YouTube videos.

4.2 Code Implementation

Tensorflow. TensorFlow is an amazing information stream in machine learning library made by the Brain Team of Google and made open source in 2015.It is intended to ease the use and broadly relevant to both numeric and neural system issues just as different spaces. Fundamentally, TensorFlow is a low level tool for doing entangled math and it targets specialists who recognize what they're doing to construct exploratory learning structures, to play around with them and to transform them into running programs.

Python 3.7. Python is broadly utilized universally and is a high-level programming language. It was primarily introduced for prominence on code, and its language structure enables software engineers to express ideas in fewer lines of code. Python is a programming language that gives you a chance to work rapidly and coordinate frameworks more effectively.

Google Colab. Colaboratory, or "Colab" for short, is a product from Google Research. Colab allows anybody to write and execute arbitrary python code through the browser, and is especially well suited to machine learning, data analysis and education. More technically, Colab is a hosted Jupyter notebook service that requires no setup to use, while providing access free of charge to computing resources including GPUs.

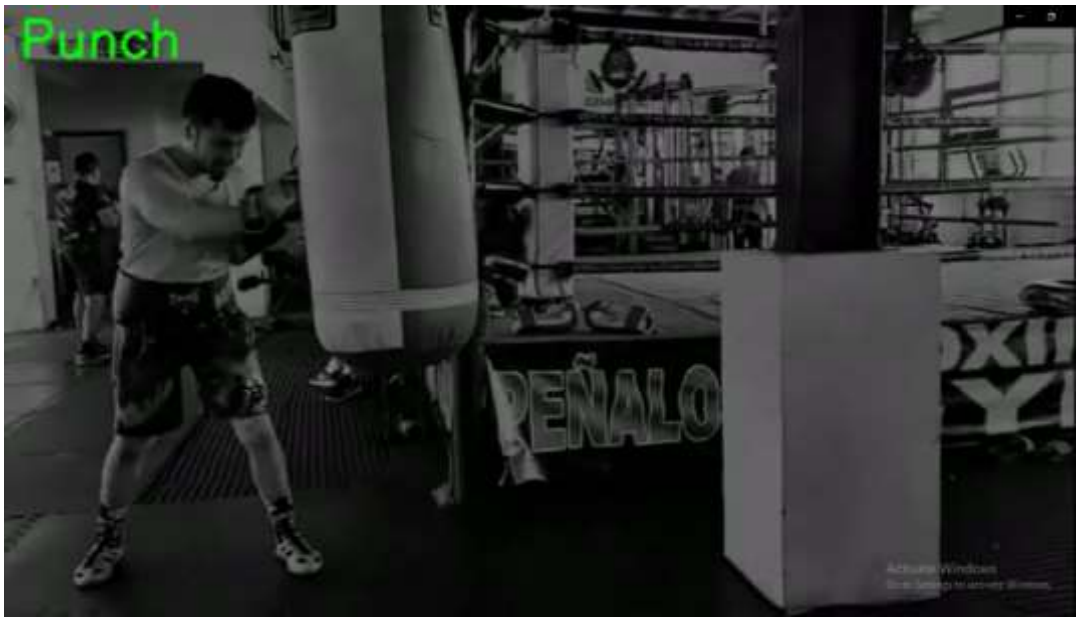
5. Result

After implementing the algorithm we have determined the activities being performed in the video inputs:

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

6.



Conclusion

Although progress in recent video-based human activity recognition has been encouraging, there are still some apparent performance issues that make it challenging for real-world deployment. The viewpoint issue remains the main challenge for human activity recognition. In real world activity recognition systems, the video sequences are usually observed from arbitrary camera viewpoints; therefore, the performance of systems needs to be invariant from different camera viewpoints.

However, most recent algorithms are based on constrained viewpoints, such as the person needs to be in front-view (i.e., face a camera) or side-view. Some effective ways to solve this problem have been proposed, such as using

multiple cameras to capture different view sequences then combining them as training data or a self-adaptive calibration and viewpoint determination algorithm can be used in advance.

Apart from the classification of primary and individual actions, the recognition of interactions of humans with objects and the detection of actions has become the new leading research topics.

.

7. Future Scope

Since computer vision is a trending topic in these days, systems like Human Activity Recognition systems is quite useful and effective for solving a variety of application, whether it would be surveillance or monitoring, or aiding the elderly and blind people etc. This not only provide additional comfort to the end-users but also can be deployed into different Organizations in order to reduce the employ workload.

The model shows good results on video streams while performing decently on image data. Activity Recognition system are of great importance in modern days due to the convenience and problems which the system offers and solves. Need of Activity recognition for monitoring and surveillance, video segmentation etc is of growing demand in which this system can greatly help. This system can be incorporated in mobiles apps to further aid the elderly and blind people.

It is cost-effective and an immense time saving system which is also prone to human errors. This system acts as a base solution for many other applications involving activity recognition. Hence, this system is very beneficial for both individual and organizations for general or specialize purposes.

8. References

1. Blank, M.; Gorelick, L.; Shechtman, E.; Irani, M.; Basri, R. Actions as Space-time Shapes. In Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV), Beijing, China, 17–21 October 2005; Volume 2, pp. 1395–1402.

2. Ke, Y.; Sukthankar, R.; Hebert, M. Spatio-temporal Shape and Flow Correlation for Action Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Minneapolis, MN, USA, 17–22 June 2007; pp. 1–8.

3. Yamato, J.; Ohya, J.; Ishii, K. Recognizing Human Action in Time-sequential Images using Hidden Markov Model. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Champaign, IL, USA, 15–18 June 1992; pp. 379–385.

4. Lu, W.; Little, J.J. Simultaneous tracking and action recognition using the PCA-HOG descriptor. In Proceedings of the 3rd Canadian Conference on Computer and Robot Vision, Quebec, PQ, Canada, 7–9 June 2006; p.

5. R. Poppe, "A survey on vision-based human action recognition," *Image and Vision Computing*, vol. 28, pp. 976–990, 2010.View at: Google Scholar

6. J. K. Aggarwal and M. S. Ryoo, "Human activity analysis: a review," *ACM Computing Surveys*, vol. 43, p. 16, 2011.View at: Google Scholar

7. T. B. Moeslund, A. Hilton, and V. Krüger, "A survey of advances in vision-based human motion capture and analysis," *Computer Vision and Image Understanding*, vol. 104, pp. 90–126, 2006.View at: Google Scholar

8. J. M. Chaquet, E. J. Carmona, and A. Fernández-Caballero, "A survey of video datasets for human action and activity recognition," *Computer Vision and Image Understanding*, vol. 117, pp. 633–659, 2013.View at: Google Scholar

9. M. Blank, L. Gorelick, E. Shechtman, M. Irani, and R. Basri, "Actions as space-time shapes," in *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, pp. 1395–1402, Beijing, China, 2005.View at: Google Scholar

10. Raju G V S, Kishor Kumar Reddy C and Narasimha Prasad L V, "Revealing of Earth Quake Magnitude using Seismic Signals and Wavelet Transforms", Elsevier International Conference on Soft Computing and Software Engineering, California, USA, March 2015.
11. Kishor Kumar Reddy C, Anisha P R, Narasimha Prasad L V, "A Pragmatic approach for Detecting Liver Cancer using Image Processing and Data Mining Techniques", IEEE International Conference on Signal Processing And Communication Engineering Systems, Guntur, India, January 2015.
12. Azmathulla Shaik, Kishor Kumar Reddy C, Anisha P R and Siddarth K, "AMTS: Advanced Movie Ticketing System", ACM International Conference on Information and Communication Technology for Competitive Strategies, Rajasthan, India, November 2014.
13. Azmathulla Shaik, Kishor Kumar Reddy C, Anisha P R and Ravi Shekar Reddy A, "MRTS: A Robust and Scalable Architecture for Metro Rail Ticketing System", IEEE International Conference on Computational Intelligence and Communication Networks, Bhopal, India, November 2014.
14. Narasimha Prasad, Kishor Kumar Reddy and Ramya Tulasi Nirjogi, "A Novel Approach for Seismic Signal Magnitude Detection Using Haar Wavelet", IEEE International Conference on Intelligent Systems, Modelling and Simulation, October 2014.
15. Narasimha Prasad Lakkakula, Kishor Kumar Reddy, and Murali Prasad Raja, "Remote Sensing Of Snow Wrap Using Clustering And Wavelet Transform," IEEE International Conference on Mathematical Modelling and Computer Simulation, Malaysia, September 2014.

16. Narasimha Prasad Lakkakula, Mannava Munirathnam Naidu, and Kishor Kumar Reddy, “An Entropy Based Elegant Decision Tree Classifier to Predict Precipitation”, IEEE European Modelling Symposium, Italy, October, 2015.
17. Kishor Kumar Reddy, Anisha P R and Narasimha Prasad L V, “A Novel Approach for Detecting the Bone Cancer and its Stage based on Mean Intensity and Tumor Size”, International Conference on Recent Researchers in Applied Computer Science, 2015.