JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 499-506
https://publishoa.com
ISSN: 1309-3452

Human Activity Recognition

Dr. SHAIK ABDUL NABI
Professor, Head of the Department, CSE
Sreyas Institute of Engineering and Technology,
Telangana, India
cse.hod@sreyas.ac.in

BATHULA NARENDRA SAI
Department of CSE,
Sreyas Institute of Engineering and Technology, Telangana, India
Narendrasai.BATHULA75@gmail.com

BAPU SHARANYA
Department of CSE,
Sreyas Institute of Engineering and Technology, Telangana, India
sharanyabapu041@gmail.com

NOMULA SHRAVYA SHREE
Department of CSE,
Sreyas Institute of Engineering and Technology, Telangana, India
shravyashree33@gmail.com

KOMMU PRACHEEN KUMAR
Department of CSE,
Sreyas Institute of Engineering and Technology, Telangana, India
pracheen.kommu@gmail.com

Received 2022 April 02**; Revised** 2022 May 20**; Accepted** 2022 June 18**.**

ABSTRACT

Human activity recognition, or HAR, is a challenging time series classification task. It involves predicting the movement of a person based on sensor data and traditionally involves deep domain expertise and methods from signal processing to correctly engineer features from the raw data in order to fit a machine learning model. Activity recognition is the problem of predicting the movement of a person, often indoors, based on sensor data, such as an accelerometer in a smartphone. Streams of sensor data are often split into subs-sequences called windows, and each window is associated with a broader activity, called a sliding window approach. Recognition of human activities aims a wide diversity of applications. Recognizing human activities by means of sensors attached on the body has been widely studied. common activity and functional performance level of a person can be determined by the capability to record and identify distinctive daily activities. Security and entertainment are others fields impacted by investigation of human behaviour through mobile phone data. Motion capture video systems have become an important research topic in the monitor human activity. The proposed human activity detection system recognizes human activities including walking, running, and sitting. While walking and running can be recorded as daily fitness activities, falling will also be detected as anomalous situations and alerting messages can be sent as needed.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 499-506
https://publishoa.com
ISSN: 1309-3452

Keywords: *Human Activity Recognition, OpenCV, Human computer Interface, Surveillance and Monitoring, camera.*

INTRODUCTION

Human activity recognition (HAR) plays an important role in people's daily lives because it has the ability to learn profound advanced knowledge about human activities from raw sensor data. With the development of human-computer interaction applications, the technology of HAR has become a popular research direction at home and abroad. People could automatically classify the type of human motion and obtain the information that the human body needs to convey by extracting features from daily activities, which in turn provides a basis for other intelligent applications. Hitherto, this technology has been widely used in the fields of home behavior analysis, video surveillance, gait analysis, and gesture recognition, etc. The associate editor coordinating the review of this manuscript and approving it for publication was Yongping Pan . Due to the rapid development of sensor technology and ubiquitous computing technology, sensor-based HAR has become more and more popular, and it is widely used with privacy being well protected. Researchers have explored the role of different types of sensing technology in activity recognition to improve recognition accuracy. According to the manner in which sensors are employed in an environment, the technologies of human activity recognition could be widely divided into two categories: approaches based on fixed sensors and approaches based on mobile sensors. The methods based on fixed sensors mean that the information is obtained from sensors mounted at a fixed position, involving acoustic sensors, radars, static cameras, and other ambient-based sensors. Among them, camera-based methods are the most popular methods, among which background subtraction method, optical flow method and energy-based segmentation method are usually applied to extract features. Representative is an image processing method based on Kinect sensors which could acquire the depth image features of moving targets whereby Jun Liu et al.proposed a space-time short-term memory (ST-LSTM) network to recognize activities. Kitani et al. presented a sparse optical flow algorithm to acquire the histogram of human motion features and proposed an unsupervised Dirichley hybrid model to classify 11 human activities. Although these activity monitoring methods can provide better recognition accuracy, they are not suitable in many indoor environments, especially where privacy is a concern. Furthermore, the results of vision-based approaches are easily affected by illumination variations, ambient occlusion, and background change. This greatly limits their practical use.

Machine learning methods may rely heavily on heuristic manual feature extraction in most daily human activity recognition tasks. It is usually limited by human domain knowledge. To address this problem, researchers have turned to deep learning methods that could automatically extract appropriate features from raw sensor data during the training phase and present the low-level original temporal features with high-level abstract sequences. In view of the successful application of deep learning models in image classification, voice recognition, natural language processing, and other fields, it is a new research direction in pattern recognition to transfer it to the field of human activity recognition. In, authors proposed to convert the data acquired by three-axis accelerometers into an ‘‘image’’ format, and then they used CNN with three convolutional layers and one fully-connected layer to identify human activities. Ordóñez and Roggen. proposed an activity recognition classifier, which combined deep CNN and LSTM to classify 27 hand gestures and five movements. Finally, simulation results showed that the F1 score on the two classifiers were 0.93 and 0.958, respectively. Lin et al.presented a novel iterative CNN strategy with autocorrelation pre-processing capability, instead of traditional micro-Doppler image pre-processing, which can accurately classify seven activities or five subjects. And this strategy used an iterative deep learning framework to automatically define and extract features. Finally, traditional supervised learning classifiers were used to mark different activities based on the captured radar signals. Although the above models could generally recognize human activities, the overall network structure is relatively complex. In addition, these models have a large number of parameters, which results in high computational cost. It is difficult to be used in occasions that require high real-time performance. Many researchers have made great efforts in this regard.

In recent years, an enormous amount of researches has been conducted by researchers in exploring different sensing technologies and a number of methods have been proposed for modeling and recognizing human activities. Early researches mainly used decision tree, support vector machine (SVM), naïve Bayes and other traditional machine learning methods to classify the data collected by sensors. In, gradient histogram and Fourier descriptor based on centroid feature were used to extract the features of acceleration and angular velocity data. Then Jain et al. used two classifiers, support vector machine and k-nearest neighbor (KNN), to recognize the activities of two public datasets. Jalloul et al. used six inertial measurement units to construct a monitoring system. After performing network analysis, a number of network measures that satisfy the statistical test were selected to form a feature set, and then the authors used the random forest (RF) classifier to classify the activities. Finally, an overall accuracy of 84.6% was achieved. The paper presented a wearable wireless accelerometer-based activity recognition system and its application in medical detection. Relief-F and sequential forward floating search (SFFS) were combined for feature selection. Finally, Naïve Bayesian and k-nearest neighbor (KNN) were used for activity classification and comparative analysis.

The other methods of activity recognition are to use mobile sensors. In these methods, the information from different kinds of behaviors is usually collected from a set of dedicated body-worn motion sensors, such as accelerometers, gyroscopes, and magnetometers. Acceleration and angular velocity data would change according to human motion. Therefore, they could be used to infer human activities. The miniaturization and flexibility of sensors allow individuals to wear or carry mobile devices embedded with various sensing units. This is different from fixed sensor-based approaches. Moreover, these sensors have the characteristics of low cost, low power consumption, high capacity, miniaturization, and less dependence on surroundings. Therefore, activity recognition based on mobile sensors has received widespread attention because of its portability and high acceptance in daily life. Correspondingly, a large number of researches have been carried out to explore the potential of mobile sensors for activity recognition in a ubiquitous and pervasive way. Margarito et al. put accelerometers on the wrist of subjects to collect acceleration data and then used template matching algorithm to classify 8 common sports activities. In, a smart life assistant system (SAIL) for the elderly and disabled was proposed. Zhu et al. collected the features by the way of multi-sensor fusion strategy and achieved the target of recognizing 13 kinds of daily activities.

Human activity recognition (HAR) aims to classify a person's actions from a series of measurements captured by sensors. Nowadays, collecting this type of data is not an arduous task. With the growth of the Internet of Things, almost everyone has some gadget that monitors their movements. It can be a smartwatch, a pulsometer, or even a smartphone. Usually, this is performed by following a fixed-length sliding window approach for the features extraction. Here two parameters need to be fixed: the size of the window and the shift.

These are some of the data you could use:

Body acceleration.

Gravity acceleration.

Body angular speed.

Body angular acceleration.

Etc.

The machine learning model used for activity recognition relies on top of the devices' available sensors. Neural networks are the perfect algorithms to determine a person's physical activity. This is due to their ability to recognize the patterns behind the data. The following graph illustrates a neural network that classifies among different activities using CNN data. Human activity recognition is the problem of classifying sequences of accelerometer data recorded by specialized harnesses or smart phones into known well-defined movements. Classical approaches to the problem involve hand crafting features from the time series data based on fixed-sized windows and training machine learning models, such as ensembles of decision trees. The difficulty is that this feature engineering requires deep expertise in the field. Recently, deep learning methods such as recurrent neural networks and one-dimensional convolutional neural networks, or CNNs, have been shown to provide state-of-the-art results on challenging activity recognition tasks with little or no data feature engineering, instead using feature learning on raw data.

LITERATURE SURVEY

For many years human action recognition has been studied well. Most of the action recognition methods require to manually annotate the relevant portion of the action of interest in the video. In recent years it has been studied that the relevant portion of action of interest can be found out automatically and recognize the action. We can review the action recognition methods .

- **Action Recognition**

For representing video, feature trajectories have shown efficiency. But the quality and quantity of these trajectories were not sufficient. As the use of dense sampling came popular for image classification Wang et al.proposed to use dense trajectories for representing videos. Dense points from each frame are sampled and traced them based on displacement information. For improving the performance Wang et al. takes into account the camera motion. The camera motion is estimated by matching feature points between the frames by using SURF descriptors and dense optical flow. Another approach aimed at modeling the motion relationship. The approach operates on top of visual codewords derived from local patch trajectories, and therefore does not require accurate foreground-background separation .Dorr et al. proposed another method for finding the informative regions. They used saliency mapping algorithms. As a new method this paper proposes using a joint learning framework for learning spatial and temporal extents of action of interest

- **Action Detection**

Recognition was performed using the Mahalanobis distance between the moment description of the input and each of the known actions. Recent popular methods which employ machine learning techniques such as SVMs and AdaBoost, provide one possibility for incorporating the information contained in a set of training examples. introduces the Action MACH filter, a template-based method for action recognition which is capable of capturing intraclass variability by synthesizing a single Action. Another method is proposed in, multiple-instance learning framework, named SMILE-SVM (Simulated annealing Multiple Instance Learning Support Vector Machines), is presented for learning human action detector based on imprecise action location. Wang et al. [6]used a figure-centric visual word representation. In that localization is treated as latent variable so as to recognize the action. A spatio-temporal model is learned .During the training[7] model parameters is estimated and the relevant portion is identified.[8] proposed an independent motion evidence feature for distinguishing human actions from background motion. Most of the methods require that the relevant portion of the video has to be annotated with bounding boxes. Human intervention was tedious. So to overcome the bounding box Brendel et al.[9] divides the video into a number of subgroups and then a model was generated that identify the relevant subgroup. This paper introduces a method that learns both spatial and temporal extents for detection improvement. Dense trajectory is used here as local features to represent the human action.

EXISTING SYSTEM

- Gestures are considered as primitive movements of the body parts of a person that may correspond to a particular action of this person
- Atomic actions are movements of a person describing a certain motion that may be part of more complex activities
- 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.

Disadvantages

- Most of the action recognition methods require to manually annotate the relevant portion of the action of interest in the video.

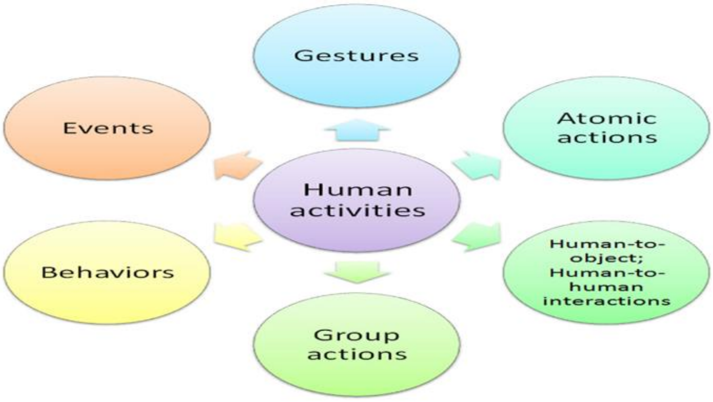


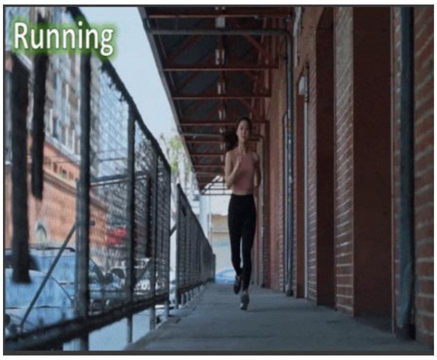
Figure 1 Human Activity

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. 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.

Advantages

- The feature points are such that the parts between the points are rigid. We finally form the skeleton structure of the human body.
- Then these points can used to form rectangles resembling human structure now the final structure formed is model on which the computer works on.
- Once we had divided the human into rectangular segments. We can track them in following frames .And hence we can track their motion.
- This can be done by searching for the rectangular region which matches the original rectangular region that was in first frame and tracking it.
- Thus we can at any point of time keep the track of rectangular frames which help us to track the human motion as a whole.



Screenshot 1 Human Activity(Running)

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 499-506
https://publishoa.com
ISSN: 1309-3452

Screenshot 2 Human Activity(Jumping)

Screenshot 3 Human Activity(Man Is Setting)

CONCLUSION

A novel deep neural network that combines convolutional layers with LSTM for human activity recognition was proposed in this paper. The weight parameters of CNN mainly concentrate on the fully-connected layer. In response to this characteristic, a GAP layer is used to replace the fully-connected layer behind the convolutional layer, which greatly reduces the model parameters while maintaining a high recognition rate. Moreover, a BN layer is added after the GAP layer to speed up the convergence of the model and obvious effect was obtained. In the proposed architecture, the raw data collected by mobile sensors is fed into a two-layer LSTMs followed by convolutional layers, which makes it capable of learning the temporal dynamics on various time scales according to the learned parameters of LSTMs so as to obtain better accuracy. In order to prove the generalization ability and effectiveness of the proposed model, the three public datasets, UC-HAR, WISDM, and OPPORTUNITY, were used for the experiment. Considering that the accuracy is not an appropriate and comprehensive measure of performance, the F1 score was used to evaluate the model performance. Eventually, the F1 score reached 95.78%, 95.85% and 92.63% on the UCI-HAR, WISDM and OPPORTUNITY datasets, respectively. Furthermore, we also explored the impact of some hyper-parameters on model performance such as the number of filters, the type of optimizers and batch size. Finally, the optimal hyper-parameters for the final design were selected to train the model. To sum up, compared with the methods proposed in other literatures, the LSTM-CNN model shows consistent superior performance and has good generalization. It can not only avoid complex feature extraction but also has high recognition accuracy under the premise of a few model parameters.

FUTURE SCOPE

- We can identify activities, classify or group participants to activities, get additional insights of activity durations and patterns.
- Need of Activity recognition for monitoring and surveillance, video segmentation etc is of growing demand in which this system can greatly help.
- It can not only adaptively extract activity features, but also has fewer parameters and higher accuracy.
- The results show that the proposed model has higher robustness and better activity detection capability than some of the reported results.

REFERENCE

1. Google Scholar Salakhutdinov, R., Torralba, A., and Tenenbaum, J. B. (2011). “Learning to share visual appearance for multiclass object
   CrossRef Full Text | Google Scholar
2. Sanchez-Riera, J., Cech, J., and Horaud, R. (2012). “Action recognition robust to background clutter by using stereo vision,” in Proc. European Conference on Computer Vision (Firenze), 332–341.
   Google Scholar
3. Sapienza, M., Cuzzolin, F., and Torr, P. H. S. (2014). Learning discriminative space-time action parts from weakly labelled videos. Int. J. Comput. Vis. 110, 30–47. doi:10.1007/s11263-013-0662-8
   CrossRef Full Text | Google Scholar
4. Satkin, S., and Hebert, M. (2010). “Modeling the temporal extent of actions,” in Proc. European Conference on Computer Vision (Heraklion), 536–548.
5. CrossRef Full Text | Google Scholar
   Zhang, Z., Wang, C., Xiao, B., Zhou, W., and Liu, S. (2015). Robust relative attributes for human action recognition. Pattern Anal. Appl. 18, 157–171. doi:10.1007/s10044-013-03493.
6. Mamidi Harshitha Reddy et al.,, Backorder Prognostication, YMER Open Access Journal, I SSN 0044 -0477, Vol 21,No 4., April,2022, Page no 460-472.
7. Dr. Shaik Abdul Nabi, An Efficient and Secure Deduplication Scheme for Cloud Assisted Health Systems,The International Journal of Analytical and Experimental modal Analysis (IJAEMA).
   ISSN 0886-9367, Vol 13, No 6., June,2021, Page no 1005-1012.
8. Dr. V.Goutham et al, “Block chain Technology: An Overview”,International Journal of Advanced Science and Technology (IJAST), ISSN 2207-6360, Vol 29, No 6. May,2020, Page no 3459-3463
9. Dr. Allam Balaram et al, Hybrid approach of segmentation in Brain Tumor Data using Fuzzy C-means and Level Set, Journal of Advanced Research in Dynamical and Control Systems – JARDCS, 1943-023X, Vol 12, No 6, May,2020, Page no 47-56.
10. Kadari Akshitha et al, Attendance Marking System Using Facial Recognition, **Advanced Science Letters Journal,** E-ISSN:1936-7317, Vol 26, No 4, April 2020, Page no 25-30.
11. S. Vaishnavi Sneha et al,”, Intelligent Boarder Security Intrusion Detection auto Destroy System using IOT”,**Juni Khyat International Journal,** ISSN 2278-4632, Vol 10, No 14, May 2020, Page no 86-91.
12. V. Sridhar Reddy et al, A Fast Supportive and Trustworthy Cloud Services, **International Journal on Future Research in Computer Science & Communication Engineering,** ISSN 2278-4632, Vol 10, No 14, May, 2020, Page no 86-91.
13. G.Dayakar et al, Securing Data and Communications Key Management in Dynamic Wireless Sensor Network, International Journal & Magazine of Engineering ,Technology,IJMETMR, ISSN 2348-4845, Vol 3,No 9, Sep,2016, Page no 181-187.
14. Dr. Shaik Abdul Nabi, Video Streaming Adaptivity and Efficiency in Social Networking Sites, International Journal of Advanced Technology and Innovative Research ,IJATIR, ISSN 2348–2370, Vol 8, No 8, July, 2016, Page no 1650-1655.

15. Vandana Janjirala et al, Social Network Analysis by Using Graphs to Represent Social Relations, International Journal & Magazine of Engineering, Technology, Management and Research, IJMETMR, ISSN 2348-4845, Vol 2, No 3, March, 2015, Page no 341 – 347.
16. Dr. Shaik Abdul Nabi, Detection and Extraction of Videos using Decision Trees, International Journal of Advanced Computer Science and Applications(IJACSA), ISSN 2156 – 5570, Vol 2 , No 12, Dec, 2011, Page no 147 -151.
17. SRILATHA PULI, A MACHINE LEARNING MODEL FOR AIR QUALITY PREDICTION FOR SMART CITIES, DESIGN ENGINEERING || ISSN: 0011-9342 | YEAR 2021 - ISSUE: 9 | PAGES: 18090 – 18104
18. SRILATHA PULI, N-GRAMS ASSISTED YOUTUBE SPAM COMMENT DETECTION, YMER || ISSN: 0044-0477, APRIL 2022
19. SRILATHA PULI, ANALYSIS OF BRAND POPULARITY USING BIG DATA AND TWITTER, YMER|| ISSN: 0044-0477, APRIL 2022
20. SRILATHA PULI, CYBER THREAT DETECTION BASED ON ARTIFICIAL NEURAL NETWORKS USING EVENT PROFILES,THE INTERNATIONAL JOURNAL OF ANALYTICAL AND EXPERIMENTAL MODAL ANALYSIS, ISSN NO:0886-9367
21. SRILATHA PULI, FACE MASK MONITORING SYSTEM, THE INTERNATIONAL JOURNAL OF ANALYTICAL AND EXPERIMENTAL MODAL ANALYSIS, ISSN NO:0886-9367
22. SRILATHA PULI, IOT BASED SMART DOOR LOCK SURVEILLANCE SYSTEM USING SECURITY SENSORS, ADVANCED SCIENCE LETTERS E-ISSN:1936-7317
23. SRILATHA PULI, SAFETY ALERTING SYSTEM FOR DROWSY DRIVER, 9TH INTERNATIONAL CONFERENCE ON INNOVATIONS IN ELECTRONICS & COMMUNICATION ENGINEERING (ICIECE-2021), Page – 40
24. N. SWAPNA SUHASINI, SRILATHA PULI, BIG DATA ANALYTICS FOR MALWARE DETECTION IN A VIRTUALIZED FRAMEWORK, JOURNAL OF CRITICAL REVIEWS, ISSN:2394-5125 VOL.7, ISSUE 14, JULY – 2020
25. Dr. Shaik Abdul Nabi, Srilatha Puli, Framework For Social Distance Monitoring With Computer Vision And Yolov3, 2nd International Conference On Iot Based Control Networks And Intelligent Systems (Icicnis 2021), Ssrn-Id3883372