HUMAN ACTIVITY DETECTION SYSTEM

CH.S.D.N.DURGA[1], SHAIK SHIRALI AHMED[2]

[1] Assistant Professor **MCA, DEPT, Dantuluri Narayana Raju College , Bhimavaram, Andharapradesh**
Email:- nemovarma38@gmail.com
[2]**PG Student of MCA, , Dantuluri Narayana Raju College , Bhimavaram, Andharapradesh**
Email:- shiraliahmed3@gmail.com

Abstract Human activity recognition, or HAR for short, is a broad field of study concerned with identifying the specific movement or action of a person based on sensor data.
The sensor data may be remotely recorded, such as video, radar, or other wireless methods. It contains data generated from accelerometer , gyroscope and other sensors of Smart phone to train supervised predictive models using machine learning techniques like SVM , Random forest and decision tree to generate a model. Which can be used to predict the kind of movement being carried out by the person which is divided into six categories walking, walking upstairs, walking down-stairs, sitting, standing and laying .
MLM and SVM achieved accuracy of more than 99.2% in the original data set and 98.1% using new feature selection method. Results show that the proposed feature selection approach is a promising alternative to activity recognition on smartphones.

Index Term: - The Online Book Store (OBS), . Recommendation systems,

I Introduction

Physical activity is well-known by the general public to be crucial for leading a healthy life. Thus, researchers are seeking a better understanding of the relationship between physical activity and health. Precise recording of the conducted activities is an essential requirement of their research. (Bauman et al., 2006) This data can be used to design and construct activity recognition systems. These systems allow physicians to check the recovery development of their patients automatically and constantly (da Costa Cachucho et al., 2011). Another rising concern is the sedentary life many people live, due to the shift in lifestyle occurring in the modern world, where work and leisure tend to be less physically demanding (Gyllensten, 2010). Several reports have already found links between common diseases and physical inactivity (Preece et al., 2009). Thus, activity recognition can be used by recommender systems to help the users track their daily physical activity and promote them to increase their activity level. With the recent progress in wearable technology, unobtrusive and mobile activity recognition has become reasonable. With this technology, devices like smartphones and smartwatches are widely available, hosting a wide range of built-in sensors, at the same time, providing a large amount of computation power.
Overall, the technological tools exist to develop a mobile, unobtrusive and accurate physical activity recognition system. Therefore, the realization of recognizing theindividuals' physical activities while performing their daily routine has become feasible. So far, no-one has investigated the usage of light-weight devices for recognizing human activities. An activity recognition system poses several main requirements. First, it should recognize activities in real-time. This demands that the features used for classification are computable in real-time. Moreover, short window durations must be employed to avoid delayed response. Finally, the classification schemes should be simple, light-weight and computationally inexpensive to be able to run on hand-held devices.

2 Literature survey

Human activity recognition has been studied for years and researchers have proposed different solutions to attack the problem. Existing approaches typically use vision sensor, inertial sensor and the mixture of both. Machine learning and threshold-base algorithms are often applied. Machine learning usually produces more accurate and reliable results, while threshold-based algorithms are faster and simpler. One or multiple cameras have been used to capture and identify body posture [8, 9]. Multiple accelerometers and gyroscopes attached to different body positions are the most common solutions [10-13]. Approaches that combine both vision andinertial sensors have also been purposed [14]. Another essential part of all these algorithms is data processing. The quality of the input features has a great impact on the performance. Some previous works are focused on generating the most useful features from the time series data

set [15]. The common approach is to analyze the signal in both time and frequency domain. Active learning technique has been applied on many machine learning problems that are time-consuming and labor-expensive to label samples. Some applications include speech recognition, information extraction, and handwritten character recognition [18,19,20]. This technique, however, has yet been applied on the human activity problem before.

3.Implementation Study

Several investigations have considered the use of widely available mobile devices. Ravi et. al. collected data from only two users wearing a single accelerometer-based device and then transmitted this data to the phone carried by the user (Ravi et al.,2005). Lester et. al. used accelerometer data from a small set of users along with audio and barometric sensor data to recognize eight daily activities (Lesteret al., 2006). However, the data was generated using distinct accelerometer-based devices worn by the user and then sent to the phone for storage.

Some studies took advantage of the sensors incorporated into the phones themselves. Yang developed an activity recognition system using a smart-phone to distinguish between various activities (Yang, 2009). However, stair climbing was not considered and their system was trained and tested using data from only four users. Brezmes et. al. developed a real-time system for recognizing six user activities (Brezmeset al., 2009). In their system, an activity recognition model is trained for each user, i.e., there is no universal model that can be applied to new users for whom no training data exists. Bayat et al. gathered acceleration data from only four participants, performing six activities. (Bayat et al., 2014) Shoaib et al. evaluated different classifiers by collecting data of smart-phone accelerometer, gyroscope, and magnetometer for four subjects, performing six activities. (Shoaib et al., 2013).

3.1proposed methodology

PROPOSED SYSTEM

The purpose of being able to classify what activity a person is undergoing at a given time is to allow computers to provide assistance and guidance to a person prior to or while undertaking a task.

The difficulty lies in how diverse our movements are as we perform our day-to-day tasks.

There have been many attempts to use the various machine learning algorithms to accurately classify a person's activity, so much so that Google have created an Activity Recognition API for developers to embed into their creation of mobile applications

The activity recognition process is described, containing four main stages.

1. Data Collection: The first step is to collect multivariate time series data from the phone's and the watch's sensors. The sensors are sampled with a constant frequency of 30 Hz. After that, the sliding window approach is utilized for segmentation, where the time series is divided into subsequent windows of fixed duration without interwindow gaps (Banos et al., 2014). The sliding window approach does not require preprocessing of the time series, and is therefore ideally suited to real-time applications.
2. Preprocessing: Filtering is performed afterwards to remove noisy values and outliers from the accelerometer time series data, so that it will be appropriate for the feature extraction stage. There are two basic types of filters that are usually used in this step: average filter (Sharma et al., 2008) or median filter (Thiemjarus, 2010). Since the type of noise dealt with here is similar to the salt and pepper noise found in images, that is, extreme acceleration values that occur in single snapshots scattered throughout the time series. Therefore, a median filter of order 3 (window size) is applied to remove this kind of noise.
3. Feature Extraction: Here, each resulting segment will be summarized by a fixed number of features, i.e., one feature vector per segment. The used features are extracted from both time and frequency domains. Since, many activities have a repetitive nature, i.e., they consist of a set of movements that are done periodically like walking and running. This frequency of repetition, also known as dominant frequency, is a descriptive feature and thus, it has been taken into consideration.
4. Standardization: Since, the time domain features are measured in (m/s 2), while the frequency ones in (Hz), therefore, all features should have the same scale for a fair comparison between them, as some classification algorithms use distance metrics. In this step, Z-Score standardization is used, which will transform the attributes to have zero mean and unit variance, and is defined as

$$x_{new} = (x-\mu)/\sigma$$

where μ and σ are the attribute's mean and standard deviation respectively (Gyllensten, 2010).

identified functions must be exercised.

Output : identified classes of application outputs must be exercised.

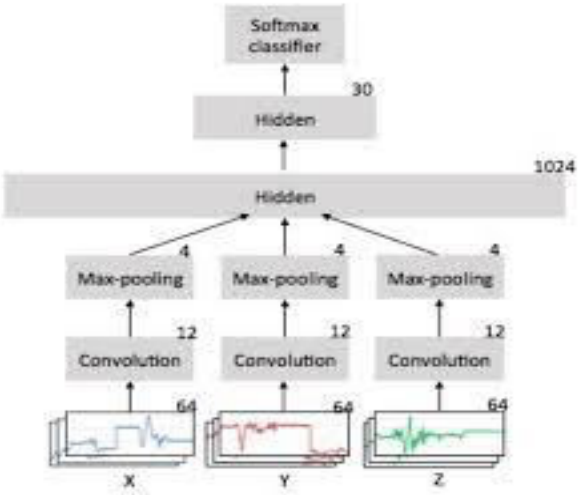

Fig:- System Architecture

4 Results and Evolution Metrics

```
PERFORMANCE ON TEST SET: Batch Loss = 1.879166603088379, Accuracy = 0.8944689035415649
Training iter #270000:   Batch Loss = 1.582758, Accuracy = 0.9386667013168335
PERFORMANCE ON TEST SET: Batch Loss = 2.0341007709503174, Accuracy = 0.8361043930053711
Training iter #300000:   Batch Loss = 1.620352, Accuracy = 0.9306666851043701
PERFORMANCE ON TEST SET: Batch Loss = 1.8185184001922607, Accuracy = 0.8639293313026428
Training iter #330000:   Batch Loss = 1.474394, Accuracy = 0.9693333506584167
PERFORMANCE ON TEST SET: Batch Loss = 1.7638503313064575, Accuracy = 0.8747878670692444
Training iter #360000:   Batch Loss = 1.406998, Accuracy = 0.9420000314712524
PERFORMANCE ON TEST SET: Batch Loss = 1.5946787595748901, Accuracy = 0.902273416519165
Training iter #390000:   Batch Loss = 1.362515, Accuracy = 0.940000057220459
PERFORMANCE ON TEST SET: Batch Loss = 1.5285792350769043, Accuracy = 0.9046487212181091
Training iter #420000:   Batch Loss = 1.252860, Accuracy = 0.9566667079925537
PERFORMANCE ON TEST SET: Batch Loss = 1.4635565280914307, Accuracy = 0.9107565879821777
Training iter #450000:   Batch Loss = 1.190078, Accuracy = 0.9553333520889282
...
PERFORMANCE ON TEST SET: Batch Loss = 0.42567864060401917, Accuracy = 0.9324736595153809
Training iter #2070000:   Batch Loss = 0.342763, Accuracy = 0.9326667189598083
PERFORMANCE ON TEST SET: Batch Loss = 0.4292983412742615, Accuracy = 0.9273836612701416
Training iter #2100000:   Batch Loss = 0.259442, Accuracy = 0.9873334169387817
PERFORMANCE ON TEST SET: Batch Loss = 0.44131210446357727, Accuracy = 0.9273836612701416
Training iter #2130000:   Batch Loss = 0.284630, Accuracy = 0.9593333601951599
PERFORMANCE ON TEST SET: Batch Loss = 0.46982717514038086, Accuracy = 0.9093992710113525
Training iter #2160000:   Batch Loss = 0.299012, Accuracy = 0.9686667323112488
PERFORMANCE ON TEST SET: Batch Loss = 0.48389002680778503, Accuracy = 0.9138105511665344
Training iter #2190000:   Batch Loss = 0.287106, Accuracy = 0.9700000286102295
PERFORMANCE ON TEST SET: Batch Loss = 0.4670214056968689, Accuracy = 0.9216151237487793
Optimization Finished!
FINAL RESULT: Batch Loss = 0.45611169934272766, Accuracy = 0.9165252447128296
```

Fig1

```
Testing Accuracy: 91.65252447128296%

Precision: 91.76286479743305%
Recall: 91.65252799457076%
f1_score: 91.6437546304815%

Confusion Matrix:
[[466   2  26   0   2   0]
 [  5 441  25   0   0   0]
 [  1   0 419   0   0   0]
 [  1   1   0 396  87   6]
 [  2   1   0  87 442   0]
 [  0   0   0   0   0 537]]

Confusion matrix (normalised to % of total test data):
[[ 15.81269073   0.06786563   0.88225317   0.           0.06786563   0.        ]
 [  0.16966406  14.96437073   0.84832031   0.           0.           0.        ]
 [  0.03393281   0.          14.21784878   0.           0.           0.        ]
 [  0.03393281   0.03393281   0.          13.43739319   2.95215464
    0.20359688]
 [  0.06786563   0.03393281   0.           2.95215464  14.99830341   0.        ]
 [  0.           0.           0.           0.           0.          18.22192001]]
Note: training and testing data is not equally distributed amongst classes,
so it is normal that more than a 6th of the data is correctly classifier in the last category.
```

Fig 2

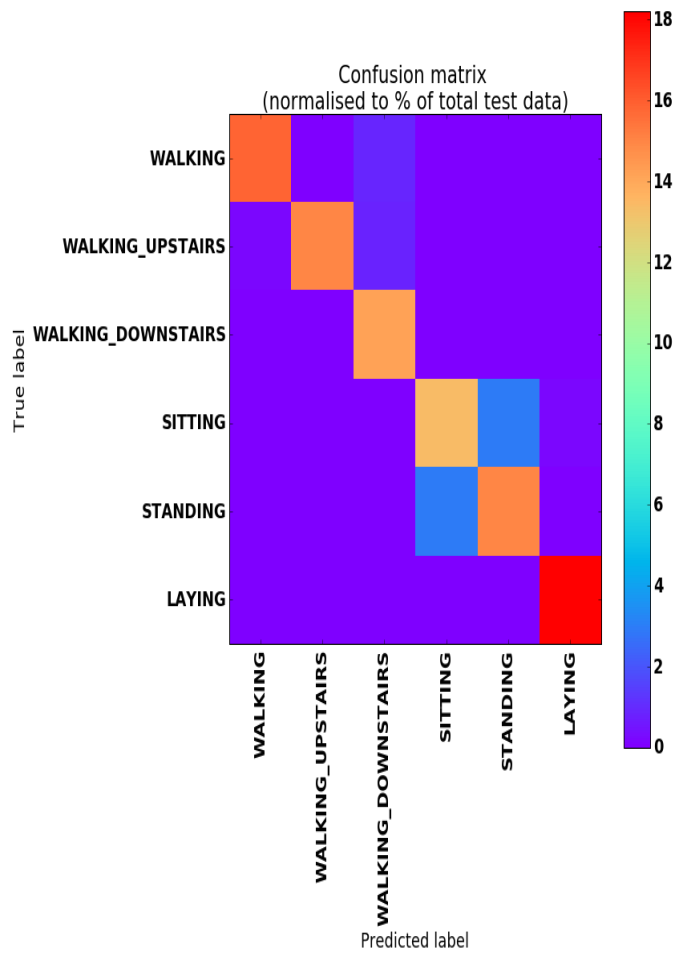

Fig 3: Predicted label

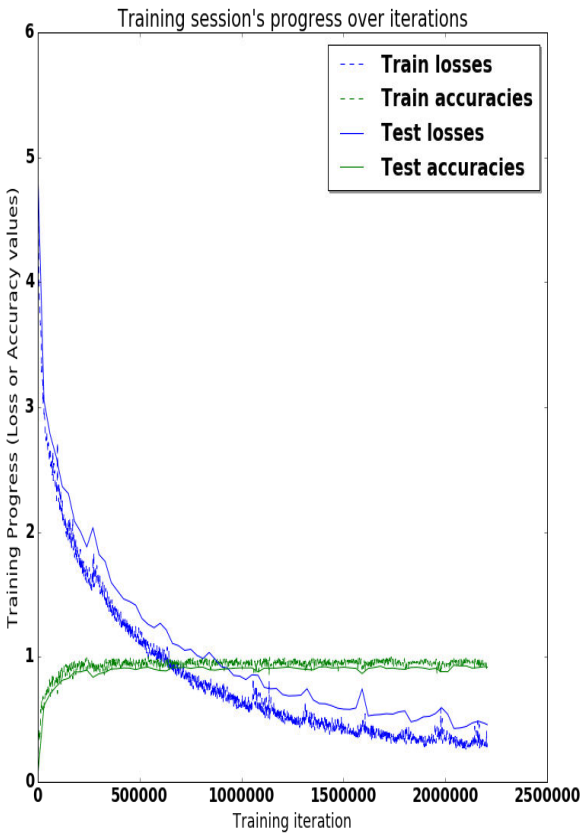

Fig 4: Training Iteration

5 **Conclusion**

In this paper, a platform to combine sensors of smartphones and smartwatches to classify various human activities was proposed. It recognizes activities in real-time Moreover, this approach is light-weight, computationally inexpensive, and able to run on handheld devices. The results showed that there is no clear winner, but naive Bayes performs best in our experiment in both the classification accuracy and efficiency. The overall accuracy lies between 84.6% and 89.4%, at which the differences are negligible. Thus, this platform is able to recognize various human activities. However, all of the tested classifiers confused walking and using the stairs activities. The second conclusion is that adding the smartwatch's sensor data to the recognition system improves it's accuracy with at least six percentage point. Finally, it is computations that the best sampling frequency is in the field of 10 Hz. Some questions still require to be answered. Most important is the conducting of larger experiments with more people in order to perform more robust evaluation to clearify if indeed one method is better than the other, or whether, any off-the-shelf method can do well in this classification task. This work could be furhter extended by incorporating more sensors (e.g. heart rate sensor), recognizing high-level activities (e.g. shopping or eating dinner) or extrapolating these trained classifiers to other people.

6 References

[1]H.Wang,A .Klaser,C.Schmid and C-L.Liu, "Action recognition by dense trajectories ," in Proc. IEEE Conf. Comput. Vis.Pattern Recog.,Jun. 2011,pp 3169-3176.

[2] H.Wang and C Schmiid , "Action recognition with improved trajectories," in Proc.IEEE Int. Conf.Comput. Vis., Dec 2013,,pp 3551- 3558.

[3] Y-G Jiang,Q.Dai,X.Xue,W.Liu and C-W Ngo. "Trajectory-based modeling of human actions with motion reference points," inProc. Eur.Conf .Comput.Vis.,Oct 2012,Vol 7576,pp.425-438.

[4] M. D. Rodriguez, J. Ahmed, and M. Shah, "Action MACH: A spatiotemporal maximum average correlation height filter for action recognition,"in Proc. IEEE Conf. Comput. Vis. Pattern Recog., Jun. 2008,pp. 1–8.

[5] Y. Hu, L. Cao, F. Lv, S. Yan, Y. Gong, and T. S. Huang, "Action detectionin complex scenes with spatial and temporal ambiguities," in Proc. IEEE Conf. Comput. Vis. Pattern Recog., Sep.–Oct. 2009, pp.128–135.

[6] T. Lan, Y. Wang, and G. Mori, "Discriminative figure-centric models for joint action localization and recognition," in Proc. IEEE Int. Conf. Comput. Vis., Nov. 2011, pp. 2003–210.

[7] M.Raptis, I.Kokkinos and S.Soatto," Discovering discriminative action parts from mid-level video representations" ,in Proc ,IEEE Conf.Comput.Vis,.Pattern Recog.,Jun 2012,pp.1242-1249

[8] W. Brendel and S. Todorovic, "Learning spatiotemporal graphs of human activities," in Proc. IEEE Int. Conf. Comput. Vis., Nov. 2011,

[9] M.Jain,J.van Genmert,H.Jegou ,P.Bouthemy and C.Snoek ,"Action localization with tubelets from motion" ,in Proc IEEE Conf, Comput.Vis.Pattern Recog. Jun 2014 pp 740-747