Journal of Discrete Mathematical Sciences and Cryptography

ISSN: 0972-0529 (Print) 2169-0065 (Online) Journal homepage: https://www.tandfonline.com/loi/tdmc20

HARGRURNN: Human activity recognition using inertial body sensor gated recurrent units recurrent neural network

Mahendra Kumar Jangir & Karan Singh

To cite this article: Mahendra Kumar Jangir & Karan Singh (2019) HARGRURNN: Human activity recognition using inertial body sensor gated recurrent units recurrent neural network, Journal of Discrete Mathematical Sciences and Cryptography, 22:8, 1577-1587, DOI: 10.1080/09720529.2019.1696552

To link to this article: https://doi.org/10.1080/09720529.2019.1696552

Journal of Discrete Mathematical Sciences & Cryptography
ISSN 0972-0529 (Print), ISSN 2169-0065 (Online)
Vol. 22 (2019), No. 8, pp. 1577–1587
DOI : 10.1080/09720529.2019.1697055

HARGRURNN : Human activity recognition using inertial body sensor gated recurrent units recurrent neural network

Mahendra Kumar Jangir *
Karan Singh †
School of Computer and Systems Sciences
Jawaharlal Nehru University
New Delhi 110067
India

Abstract

Internet of Things (IoT), wearable devices, and smartphones are more ubiquitous and available at a reasonable cost. Fitbit, Microsoft Band, Apple Watch, MI band, and Smartphone apps namely Strava and Runkeeper are some of the commercial products which are already available in the market. These products are embedded with the sensors that facilitate them to regularly sense and capture the environmental, physiological, functional data for applications in healthcare, wellbeing, and sports. These sensing devices have mediocre computing capabilities for data processing and transfer. The size of devices is so compact that can be worn on the body put in the pocket or locate in the house. This large scale recorded a wealth of physiological data require an efficient and meaningful interpretation as well as the proper and proficient method of analysis and the classification of data.

Human Activity Recognition is classified into two features say Shallow and Deep features. Shallow features are extracted conventionally with the help of a simple machine learning approach. The large-scale time series situ data which require high computation power, processing capabilities, activities, and real-time classification, Deep Learning (DL) is the promising technique to deal with these activities.

Weexercise the triaxial Accelerometers and triaxial Gyroscope to capture the data for HAR. The data set measured by the inertial sensor is then divided into the segments of 4 to 10s. The performance will be shown on the parameters of precision (%), Recall (%) and

"This work is carried out in Security and Computing Laboratory, SC&SS, JNU, New Delhi, India and sponsored by the UPE-II grant". We want to state that there is no conflict of interest and during this research, there is no harm of nature.

**E-mail:* `Jangirnawalgarh@gmail.com (Corresponding Author)`
†*E-mail:* `karancs12@gmail.com`

Accuracy (%).Parameter optimization and extraction of the segments are the two areas where we can carry forward our research.

Subject Classification: *Primary 93A30, Secondary 49K15*

Keywords: *HAR, Shallow and deep features, Smartphone, Feature extraction, Neural Network*

1. Introduction

The Technology, as well as the Population of the globe, is increasing day by day at very high speed. This is the era of information, which makes one's life very fast and informative. The forecast by Cisco says that by 2020 the 5.4 billion people will be on a mobile phone. Internet connectivity is also increasing in a parallel manner so it will obviously create a big amount of data. In this very competitive and quick life, to a regular monitor and diagnose human behavior, one should have to recognize human activities as a time series analysis.

Currently, the Internet of Things (IoT), wearable devices, and smartphones are more ubiquitous and available at a reasonable cost. Fitbit, Microsoft Band, Apple Watch, MI band, and smartphone apps, namely Strava and Runkeeper are some of the commercial products which are already available in the market. These products are embedded with the sensors that facilitate them to regularly sense and capture the environmental, physiological, functional data for application in healthcare, wellbeing, and sports. These sensing devices have mediocre computing capabilities for data processing and transfer. The size of devices is so compact that can be worn on the body put in the pocket or locate in the house. These large scales recorded a wealth of physiological data require an efficient and meaningful interpretation as well as the proper and proficient method of analysis and classification of data.

Human Activity Recognition is classified into two features say Shallow and Deep features. Some of the shallow features like Interquartile range, Variance, Standard deviation, Skewness, Root mean square, Median, Amplitude, Kurtosis, Mean, Min, Max and so on are extracted conventionally with the help of a simple machine learning approach. As we know the inertial sensor devices provide the large scale time series situ data which require high computation power, processing capabilities, activities, and real-time classification. To implement these successfully, DL is a promising technique. "Deep learning is a particular kind of machine learning that achieves great power and flexibility by learning to represent the world as a nested hierarchy of concepts, with each concept defined in relation to simpler concepts, and more abstract representations computed in terms of less abstract ones".

The motivation behind the DL method for HAR is that the conventional method provides poor HAR accuracy, whereas the DL method promises high accuracy, flexibility, and better system performance. It is difficult to find the entire hierarchy of features using DL methods with less computational layers, In such circumstances, shallow features, deliver better performance than the DL approach.

The DL methodology that combines shallow features and the features learned from inertial sensor which are complementary to each other achieves accurate and real-time activity classification.Before passing data onto the Deep learning module the spectral domain preprocessing is used to optimize the activity for the real-time scenario.

We will use the triaxial Accelerometers and triaxial Gyroscope to capture the data for HAR. The data set measured by the inertial sensor is then divided into the segments of 4 to 10s. Then DL and Shallow features are extracted separately. For automatically extracted deep features we use spectrogram which will help to ease the data of Accelerometer as we know these data are of Multifrequency, Fluctuating signals, Aperiodic. Finally, a soft-max layer andfully connected layer of the DL model is used to combine Deep and shallow features together and classification. The experiment of the framework will be performed on the five well-defined data sets viz. Activemiles, WISDM v1.1, WISDM v2.0, Daphnet FoG and Skoda. The performance will be shown on the parameters of precision (%), Recall (%) and Accuracy (%).

2. Literature Survey

Data processing techniques, the wireless and wired communication network [1-2], and sensor technology have made considerable growth and evolution during the last decade. For precise and realistic Human Activity Recognition (HAR), the high capacity, low cost, low power, and miniaturized sensor are used. For activity recognition and monitoring the idea of the sensor was very first used in the late 1990s. This idea was implemented for location-based applications and home automation targeting to adapt systems to users [15-16]. For time series analysis, while designing the classification method, the prime challenge is the selection of an appropriate set of features for subsequent classification. There are numbers of surveys related to the activity recognition using a varied range of classification methods and features are available in research discourse [3], [4]. The freezing of gait in Parkinson patients is detected by the frequency analysis of input data using the energy threshold method [5].

In some applications, to describe the time series data, we have a set of shallow features viz. statistical parameters [6], basis transform coding [7], and symbolic representation [8]. For classification of data, the above mentioned features are then used to train methods such as Support Vector Machine (SVM) and Decision Trees [9–10].

To attained maximum accuracy Catal et al. [11] proposed a HAR model called the ensemble of classifiers, which combines several classification methods. Each classification method contributes to achieve accuracy. There are several DL methods *viz.* Convolution Neural Network (CNN), Restricted Boltzmann Machine (RBM), and Deep Belief Network (DBN), which can be used to learn the discriminative set of features directly from the data [12-13]. Based on DL methods such as RBNs and DBNs which constitute multiple hidden layer Alsheikh et al. [14] demonstrate the activity recognition of data.

3. Method

Shallow and deep features are extracted from the data sets to address the computational problems. For extracting deep features, we have

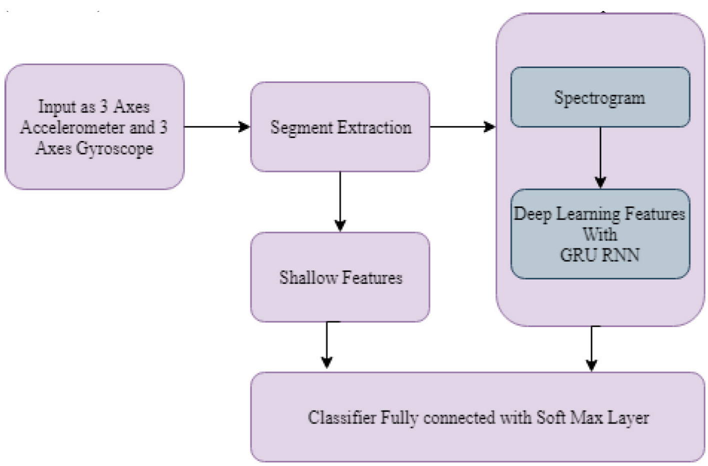

Figure 1

Schematic Diagram of Deep and Shallow features extraction using GRU RNN

to use deep learning approach. Although the deep features are more useful to data extraction but Ravi et.al. [17] “shown that if a device has limited resources for deep learning then the features derived from DL are sometimes less discriminative than the complete set of predefined shallow features”. Shallow and deep features are complementary to each other that are jointly used for data classification in [18]. The joint venture of deep and shallow features gives better accuracy to time series data classification as compare to individual shallow or deep features methodology, but this joint approach using CNN encounters more complexity and it was comparatively difficult to modify. In the proposed approach we use Gated Recurrent Units Recurrent Neural Network (GRU RNN). The workflow of the proposed model for data extraction is depicted in the figure 1.

The first block is showing the Input as 3Axes Accelerometer and 3 Axes Gyroscope data, collected from inertial sensors which were placed on the body of the object. The collected big data from the body sensors require high computational processing as well as a large amount of time to be processed so to make it easy typically, the segments of 4 to 10 seconds are used for human activity recognition (HAR) in the second block as a segment extraction (SE) block. The output of the SE block is classified into shallow and deep features.

Shallow features that we considered are predefined that were extracted separately from each segment of 3 axes accelerometer and 3 axes gyroscope data. A number of Shallow features can be directly extracted from the raw data whereas other remaining can be collected by first derivatives. Kurtosis, Mean, Min, Max, Median, Amplitude, Variance, Skewness, Interquartile Range, Standard Deviation, Root Mean Square, Zero-Cross and Mean –Cross are some of the examples of shallow features.

The Spectrogram is the representation of the inertial signal as a function of frequency and time and it is the magnitude square of the short-time Fourier transformation (STFT).The spectrogram is used because Accelerometer generates Multi frequency, Aperiodic and Fluctuating signals which are complicated to Activity Recognition using time series data so Spectrogram is used which helps Deep activity recognition model to capture variation in data.

We have embedded the Gated Recurrent Units Recurrent Neural Network for computation. In this methodology, the deep and shallow features are computed in a parallel manner, finally, a fully connected soft-max layer is used to merge and classify these two sets of features.

4. Experimental Setup

To carry out the experimental setup, the data set of Human Activities is taken as input, so to train and evaluate the proposed model we consider the well-known public benchmark data set of human activity recognition (HAR). Inertial sensors are attached to the human body that continuously captures the diverse movement of human activities in continuous time

Table 1

Human Activity Database summary.

Data set	Description	#of Classes	Subjects	Samples	Sampling Rate	Reference
Daphnet FoG	Parkinson's patients freezing of gait	2	10	1,917,887	64HZ	[19]
Skoda	Manipulative gestures performed in a car maintenance scenario	10	1	~701, 440	98HZ	[20]
ActiveMiles	Humana activity captured by Smartphone in uncontrolled environment	7	10	4,390,726	50-200 HZ	[17]
WISDM v1.1	Human activity captured by Smartphone in lab.	6	29	1,098,207	20 HZ	[21]
WISDM v2.0	Humana activity captured by Smartphone in uncontrolled environment	6	563	2,980,765	20 HZ	[22]

series manner. Table 1 shows the various well-known data sets that are used as input for our proposed model.

Daphnet FoG [19] is the data set of movement of patient of a Parkinson's disease; these patients are suffering from the gait or FoG symptoms. 3 Axes Accelerometer is attached to the body of patients to collect the data, this data is then divided into the two classes, viz. normal and freeze. If the gait of the patient was frozen it is considered as the freeze otherwise normal.Skoda [20] is the dataset of an employee while maintaining car, the 3 Axes Accelerometer is attached to the right hand of the employee which arrests 10 different activities.

ActiveMiles [17] is a completely free phone app in a smartphone that allows tracking daily human activity in an uncontrolled environment. The data set is divided into 7 different numbers of classes. WISDM v1.1 is the captured data of human activity under controlled laboratory condition. This dataset has 1,098,207 samples of 3 Axes accelerometer at 20HZ sample rate.

5. Training and Performance Metrics

We have five different types of data sets by that we have to extract mainly two types of features say deep and shallow features. To train our model we use 80% of each data set as training and the remaining 80% for testing purpose. The model uses a different hidden layer for training the data set and the weight is initialized randomly. During each phase of training, we find the difference between the obtained and target value that is labeled as the cost function. To minimize this cost function, we use the backward propagation routine; it will update the weight of different layers.

We have many evaluation metrics such as Precision, Recall, Accuracy and F1-score for the verification of the performance of our model [23].

Precision: Precision is the closeness of two of more different measurements. Mathematically Precision is defined as the ratio of true samples out of the classified positive.

$$\text{Per class precision}_c = \frac{tp_c}{tp_c + fp_c}$$

$$\text{Over all Precision} = \frac{1}{\varphi}\left(\sum_{c=1}^{\varphi} \frac{tp_c}{tp_c + fp_c}\right)$$

Where "φ is the number of classes in our data set, fp_c is the false positive rate of class c and fp_c is a true positive rate of c class".

Recall: "Recall is the ratio of correctly classified samples to a total number of samples of class whereas overall recall is the average recall of every class". The mathematical equation of Recall is as follow.

$$\text{Per class Recall}_c = \frac{tp_c}{tp_c + fn_c}$$

$$\text{Over all Recall} = \frac{1}{\varphi}\left(\sum\nolimits_{c=1}^{\varphi} \frac{tp_c}{tp_c + fn_c}\right)$$

fp_c denotes the false negative rate of class c

Accuracy: Accuracy is the ratio of correctly predicted label to all predicted labels of class c. The mathematically Accuracy is defined as follow.

$$\text{Overall accuracy} = \frac{TP + TN}{TP + TN + FP + FN'}$$

Where TP is the overall true positive rate, TN is overall true negative, FP is an overall false positive rate, and FN is the overall false negative rate.

$$F1 - \text{Score} = \sum\nolimits_{c=1}^{\varphi} 2\left(\frac{n_c}{N}\right) * \frac{\text{precision}_c * \text{recall}_c}{\text{precision}_c + \text{recall}_c}$$

F1-score: F1-score is the weighted harmonic mean of recall and precision. The mathematical representation of F1 score is represented as above.

6. Conclusion and Future Scope

In recent years human being opted the trend of wearable devices that capture continuous functional and physiological data. For classification and analyzing these huge data, we need an efficient technology. DL plays a vital role to do so.The spectral domain preprocessing is used to optimize the activity for real-time scenario. The state of art method using real world data set and laboratory experiments are used to find the classification accuracy of DL.

We used the triaxial Accelerometers and triaxial Gyroscope to capture the data for HAR. The data set measured by the inertial sensor is then divided into the segments of 4 to 10s. The performance will be shown on the parameters of precision (%), Recall (%) and Accuracy (%). For future scope one can recognize some additional activities and to improve the real time results the sample size of the people may increase. The performance [24-26] can be evaluated on different position of the user by keeping his smart devices on different orientation.

References

[1] A. Pantelopoulos and N. G. Bourbakis, "A survey on wearable sensorbased systems for health monitoring and prognosis," IEEE Trans. Syst., Man, Cybern. C, Appl. Rev., vol. 40, no. 1, pp. 1–12, Jan. 2010.

[2] D. Ding, R. A. Cooper, P. F. Pasquina, and L. Fici-Pasquina, "Sensor technology for smart homes," Maturitas, vol. 69, no. 2, pp. 131–136, 2011.

[3] M. Shoaib, S. Bosch,O.D. Incel, H. Scholten, and P. J. Havinga, "A survey of online activity recognition using mobile phones," Sensors, vol. 15, no. 1, pp. 2059–2085, 2015.

[4] C. C. T. Clark, C. M. Barnes, G. Stratton, M. A. McNarry, K. A. Mackintosh, and H. D. Summers, "A review of emerging analytical techniques for objective physical activity measurement in humans," Sports Med., pp. 1–9, 2016.

[5] M. B¨achlin et al., "Wearable assistant for Parkinson's disease patients with the freezing of gait symptom," *IEEE Trans. Inf. Technol. Biomed*., vol. 14, no. 2, pp. 436–446, Mar. 2010.

[6] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," *ACM Comput. Survey*, vol. 46, no. 3, Jan. 2014, Art. no. 33.

[7] T. Huynh and B. Schiele, "Analyzing features for activity recognition," in Proc. 2005 Joint Conf. Smart Objects Ambient Intell.: Innovative Context- Aware Services: Usages Technol., 2005, pp. 159–163.

[8] J. Lin, E. Keogh, S. Lonardi, and B. Chiu, "A symbolic representation of time series, with implications for streaming algorithms," in Proc. 8th ACM SIGMOD Workshop Res. Issues Data Mining Knowl. Discovery, 2003, pp. 2–11.

[9] Y. Lu, Y. Wei, L. Liu, J. Zhong, L. Sun, and Y. Liu, "Towards unsupervised physical activity recognition using smartphone accelerometers," Multimedia Tools and Appl., pp. 1–19, 2016.

[10] K. H. Walse, R. V. Dharaskar, and V. M. Thakare, "A study of human activity recognition using AdaBoost classifiers on WISDM dataset," *Inst. Integrative Omics Appl. Biotechnol. J.*, vol. 7, no. 2, pp. 68–76, Jan. 2016.

[11] C. Catal, S. Tufekci, E. Pirmit, and G. Kocabag, "On the use of ensemble of classifiers for accelerometer-based activity recognition," *Appl. Soft Comput.*, vol. 37, pp. 1018–1022, Dec. 2015.

[12] M. Zeng et al., "Convolutional neural networks for human activity recognition using mobile sensors," in Proc. 2014 6th Int. Conf. Mobile Comput., Appl. Services, Nov. 2014, pp. 197–205.

[13] Y. Chen and Y. Xue, "A deep learning approach to human activity recognition based on single accelerometer," in Proc. 2015 *IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 2015, pp. 1488–1492.

[14] M. A. Alsheikh, A. Selim, D. Niyato, L. Doyle, S. Lin, and H.-P. Tan, "Deep activity recognition models with triaxial accelerometers," Nov. 2015.

[15] U. Leonhardt and J. Magee, "Multi-sensor location tracking," in Proc. 4th ACM/*IEEE Int. Conf. Mobile Comput. Netw.*, 1998, pp. 203–214.

[16] Khan, Tayyab, Karan Singh, Mohamed Abdel-Basset, Hoang Viet Long, Satya P. Singh, and Manisha Manjul. "A Novel and Comprehensive Trust Estimation Clustering-Based Approach for Large Scale Wireless Sensor Networks." *IEEE Access*, vol. 7, pp. 58221-58240, (2019).

[17] D. Ravi, C. Wong, B. Lo, and G.-Z. Yang, "Deep learning for human activity recognition: A resource efficient implementation on low-power devices," in Proc. 2016 IEEE 13th Int. Conf. Wearable Implantable Body Sensor Netw., Jun. 2016, pp. 71–76.

[18] D. Ravi, C. Wong, B. Lo, and G.-Z. Yang, "A deep learning approach to on-node sensor data analytics for mobile or wearable devices. *IEEE journal of biomedical and health informatics,* 2016, 21(1), 56-64.

[19] M. B¨achlin et al., "Wearable assistant for Parkinson's disease patients with the freezing of gait symptom," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 436–446, Mar. 2010.

[20] P. Zappi et al., "Activity recognition from on-body sensors: Accuracypower trade-off by dynamic sensor selection," in Wireless Sensor Networks. Berlin, Germany: Springer, 2008, vol. 4913, ch. 2, pp. 17–33.

[21] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," SIGKDD Explor. Newslett., vol. 12, no. 2, pp. 74–82, Mar. 2011.

[22] G. M. Weiss and J. W. Lockhart, "The impact of personalization on smartphone-based activity recognition," in Proc. AAAI Workshop Activity Context Representation: Techn. Lang., 2012, pp. 98–104.

[23] Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. Inf. Process. Manag. 2009, 45, 427–437.

[24] Zhang, Qianqian, and Shifeng Liu. "Urban traffic flow prediction model based on BP artificial neural network in Beijing area." *Journal of Discrete Mathematical Sciences and Cryptography* 21, no. 4 (2018): 849-858.

[25] Applying the artificial neural networks to the renewal function estimation procedure for equipment replacements in preventive maintenance planning and management

[26] Lv, Liping. "An improved neural network algorithm for acceleration sensor to recognize human posture." *Journal of Discrete Mathematical Sciences and Cryptography* 21, no. 3 (2018): 789-795.