Comparative Performance Study on Human Activity Recognition with Deep Neural Networks

Gajjala Kavya sree[1], Kothamachu Ramesh Ashika[1], Kotaro Nakano[2], and Basabi Chakraborty[3]

[1] Graduate School of Software and Information Science, Iwate Prefectural University, Takizawa, Iwate, Japan.
[2] Research and Regional Co-operative Division, Iwate Prefectural University, Takizawa, Iwate, Japan.
[3] Faculty of Software and information science, Iwate Prefectural University, Takizawa, Iwate, Japan.

Abstract. Human activity recognition (HAR) plays an important role in every spheres of life as it assists in fitness tracking, health monitoring, elderly care, user authentication and management of smart homes. The assistive applications can be implemented on smartphones and wearable watches which are easily accessible and affordable. Now-a-days use of smart phones is ubiquitous, sensor data of diverse physical activities can be easily collected by in-built motion sensors. Many research works are proposed in this area using machine learning techniques including deep neural networks to develop smartphone based applications for human activity recognition. Our objective is to find an effective method from a variety of machine learning including deep learning models for low cost as well as high accuracy activity recognition. To fulfill the objective, a comparative performance study has been done in this work by simulation experiments on five publicly available bench mark data sets. The simulation results show that deep learning models, especially, 2D CNN and BI GRU can be promising candidates for developing smartphone based applications using motion sensor data.

Keywords: Human activity recognition · smart phone sensors · machine learning · deep neural network

1 Introduction

Human actions such as walking, running, cooking, eating, lying down, sitting and so on are referred as human activities and monitoring them with emerging technologies for some useful purpose is the main objective of the research area of Human activity recognition(HAR). It is supportive in health monitoring[1], medical care, authentication, advance computing, sports and smart homes[2]. HAR is one of the supporting technology for daily life care of aged people and the development of smart patient monitoring systems[3].

Activity recognition can be done through video based systems as well as other sensor based systems.Video based system[4] uses surveillance cameras to capture image or video for recognizing movements of people to identify their activities. But there are few shortcomings of video based systems such as intrusion to privacy and high cost. Sensor based systems use body sensors[5] or ambient sensors for the identification of person's movements. Now a days many sensors are embedded in smart phones[6], smart watches, spectacles, shoes and some non-movable objects like wall, furniture etc. Smartphones and watches have in-built accelerometer, gyroscope and magnetometer. Accelerometer helps to capture acceleration and velocity of the movement[7][8]. Gyroscope helps to capture orientation and angular velocity. Magnetometer is embedded with accelerometer and gyroscope. It identifies change in the magnetic field at certain position. From these sensors we can acquire tri variate time series data which can be used to recognize human activity with reasonable accuracy. Various research works have been done already in this field using several machine learning techniques including deep neural networks.

In this work, activity recognition from smartphone sensor data by machine learning including several deep neural networks models has been studied in order to find out an effective method suitable for developing a cost effective smart phone based health monitoring application. A comparative performance study of different techniques has been done by simulation experiments with multiple benchmark data sets. The following section describes some of the related works. Section 3 describes performance study of different models followed by the results and analysis in the section 4. Last section comprises of summary and conclusion.

2 Related Work

Due to low cost, high degree of portability, and wide range of real world applications, wearable sensors based activity recognition with machine learning techniques became a popular research area. With the increasing popularity of deep neural networks, several researchers have proposed different deep neural network based techniques for human activity recognition. Here a brief summary of research works on traditional machine learning techniques and deep neural network-based techniques for human activity recognition problem has been presented.

2.1 Machine Learning methods

Among traditional machine learning algorithms, Support Vector Machine (SVM)[9], K-Nearest Neighbor (KNN)[10], Decision tree(DT) [11], Naive Bayes(NB), Hidden Markov Model (HMM)[12], Random forest[13] classifiers have been used to recognize human activity in several works. Fan [11] used decision tree for classification of daily activities collected from wearable accelerometers. In [9], the authors presented a system architecture based on support vector machine (SVM) for HAR. HMM is used in[12] to classify the physical activities. In this approach the authors combine shape and optical flow features extracted from videos. Random forest is used in [13] for classifying daily activities and achieved accuracy more than 90%. Comparative study of HAR on machine learning algorithms can be found in [14].

2.2 Deep Neural Network Methods

A comprehensive study on HAR using deep neural networks can be found in [15]. Works on deep network based HAR can be categorized into three types. The first category , the most popular, uses Convolutional Neural Network (CNN). Ronao [16] used 1 dimensional CNN to classify the activity data recorded by smartphone sensors and compared their proposed model using SVM and DT. Authors in [17] used 2 dimensional CNN to classify six daily activities recorded from 12 volunteers. In [18], authors presented new architecture for CNN and handcrafted feature based methods to reduce the computation cost.

The second category uses Recurrent Neural Network (RNN) to capture time dependency of sensor data [19]. Among the RNN models, Long Short Term Memory (LSTM) network is the most popular one. In [20], authors applied LSTM to recognize daily activities and found LSTM is more accurate than machine learning methods. In[21], authors proposed BI LSTM for detecting activities.This method has two LSTM layers for extracting temporal features from both forward and backward direction. In [22], the authors used BI LSTM to classify 12 different activities recorded for 10 subjects. There is one more RNN model namely Gated Recurrent Unit (GRU). In [23], authors used GRU instead of LSTM and applied it to activity recognition. From the experiment they found GRU is efficient than CNN. In [24], authors presented new approach for using bidirection GRU model in human activity recognition and found BI GRU model gives promising and high quality recognition results.

The third category uses mixed model to identify human activity. This category represents the combination of CNN and RNN networks[25]. They observed that the combination attains higher accuracy because this model can utilize power of CNN in feature extraction and RNN in temporal dependencies among activities. Comparative study of HAR using hybrid models can be found in [26].

3 Comparative Performance Study

The performance study has been done by simulation experiments with five different bench mark data sets.

3.1 Datasets

The data sets used in this study are presented briefly in Table 1, the details can be found in the references noted in the Table.

3.2 Methodology

Here a very brief introduction of the methods used in our study has been presented. K-Nearest Neighbour (KNN), Naive Bayes (NB), Decision tree (DT), eXtreme Gradient Boosting (XG Boost), Random Forest (RF) and Support Vector Machine (SVM) are used for our study. KNN is easily implementable,

Table 1. Details of Datasets

Data Sets	Sensors	Sampling Frequency(HZ)	Activities	Subjects
WISDM[27]	Accelerometer	20	6	36
UCI HAR[28]	Accelerometer, Gyroscope	50	6	30
UCI HHAR[29]	Accelerometer, Gyroscope	50	6	9
Motion Sensor[30]	Accelerometer, Gyroscope	48	6	24
PAMAP2 [31]	Accelerometer, Gyroscope, Magnetometer	100	18	9

popular and computationally cheap instance based classifier. Naive Bayes is a probabilistic classification technique based on Bayes theorem. DT is another popular nonparametric supervised learning method used for classification while RF represents ensemble of DT. XG Boost is a newly developed, highly efficient and portable implementation of gradient boosted decision trees which provides high accuracy in pattern classification and regression problems. SVM is a discriminative classifier which aims to find a hyperplane with maximum margin so that error rate for classification is the least.

Among deep networks, we have used Convolutional Neural Networks (CNN) and two variants of Recurrent Neural Networks (RNN) : Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU). CNN belongs to the class of multilayer feed forward network composed of convolution layer, max pooling layer, flatten layer and dense layer. These layers are stacked to form deep architecture for feature extraction from raw sensor data before classification. We have used one dimensional (1D CNN) and two dimensional (2D CNN) for our study. LSTM belongs to the class of recurrent neural network (RNN). This network captures temporal dependencies which can be used for prediction problems. Bi directional LSTM (BI LSTM) is an extension of LSTM which is comprised of 2 LSTM cells, and information flows both forward and backward direction. GRU is another variant of RNN similar to LSTM. But there are a few differences, such as LSTM has 4 gates and GRU has 2 gates. It is observed in certain tasks that GRU exhibits better performance than LSTM. Bi directional GRU (BI GRU) is a combination of two GRU's.

4 Simulation Experiments and Results

Simulation experiments with bench mark data sets have been done for selected traditional machine learning techniques and deep network models 1D CNN, 2D CNN, LSTM , BI LSTM, GRU and BI GRU mentioned in the previous section using Python 3.7 in Anaconda3. Performance of machine learning models in terms of classification accuracy and computational time is represented in Fig 1 & Fig 2 respectively. From Fig 1 we can infer that **SVM** achieved the best performance with an average of 82% over all data set while Naive Bayes achieved accuracy of 71% which is the worst compared to the other models. We also observed that UCI HAR dataset achieved the highest accuracy for all the models. From Fig 2, we can interpret that SVM takes longer time compared to the other models. Decision tree and XG Boost take lesser time to predict activity. By considering accuracy and time complexity we can conclude that **XG Boost** yields the best performance.

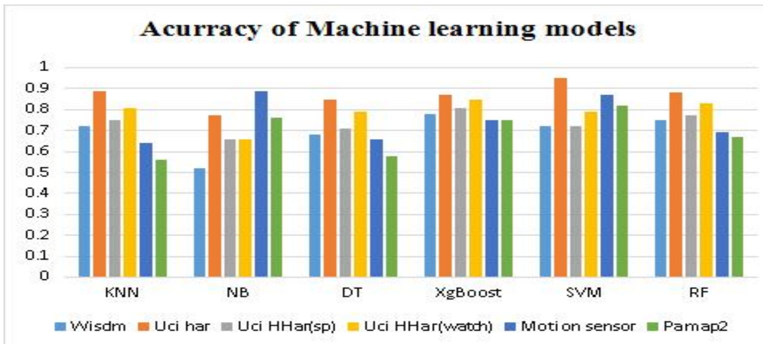

Fig. 1. Classification accuracy of traditional machine learning models

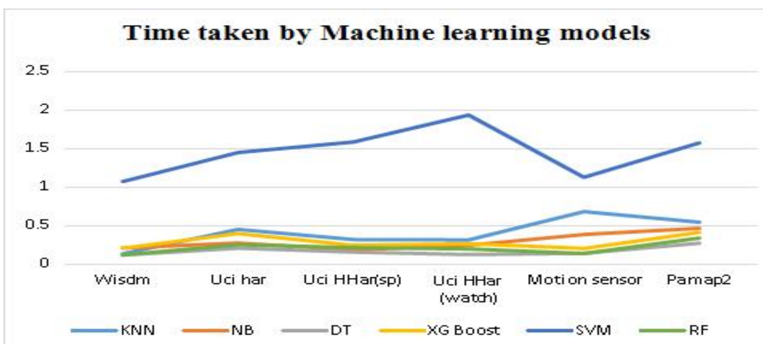

Fig. 2. Computation time for machine learning models

Deep learning models are implemented using tensor flow and Keras library for efficient performance. To stabilize and speed up training phase, batch normalization is used with a batch size of 80 data segments. The learning rate of the training is 10^{-3}. Sliding windows are used in this experiment to generate epochs with duration of 4s with 50% overlap. Each model is configured to run over 20-100 epochs using "sparse categorical loss entropy" as the loss function.

After experimenting multiple times it is observed that every model have least loss value at epoch 50. Results of deep learning models are presented in Fig 3 & Fig. 4. Fig 3. represents classification accuracy of deep learning models for multiple data sets. It is observed that the performance of all the models are better compared to traditional machine learning models. Among deep learning models **BI GRU, BI LSTM, 2D-CNN** achieved better results for all the data sets. 1D-CNN achieved the worst accuracy among DL models. Fig 4 represents computation time of different models. It is noticeable that BI LSTM consumes longer time compared to the other models. LSTM and BI GRU take almost the same time. GRU takes shorter time among RNN models whereas CNN models take less time compared to RNN models. Among all the models **2D CNN** takes the shortest time. UCI Har data set takes relatively higher time compared to other data sets.

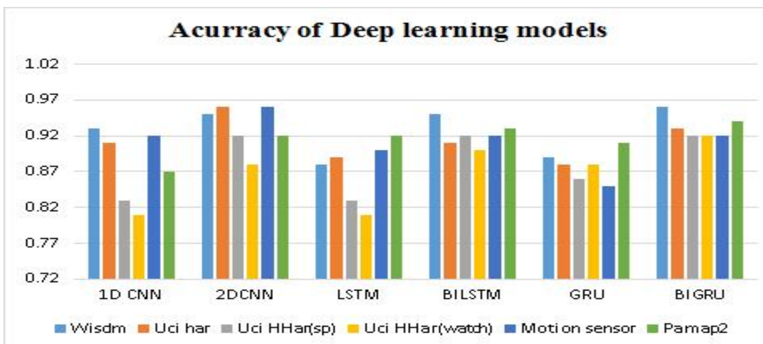

Fig. 3. Classification accuracy of deep learning models

Table 2. summarizes recognition accuracy of all the methods studied for all the data sets. For all the data sets, deep learning models produced better classification accuracy than traditional machine learning models. Tables 3-5. represent activity wise accuracies of 2D CNN, BI LSTM, and BI GRU respectively. It seems that **BI-GRU** is the best model for discriminating different activities. All three models achieved high accuracy for sitting, standing and lying ,as signal is constant throughout the activity. Upstairs and downstairs activities attained less accuracy compared to other activities, due to more similarity in them but BI GRU can distinguish these two activities better than other models. Walking attained satisfactory results in all the models. Riding bike achieved good results with all the models but there is slight inaccuracy as the signal values are similar to downstairs.

Fig 5. represents average accuracy of each activity for 2D CNN, BI LSTM and BI GRU. Three models achieved same accuracy for sitting, laying, jogging and bike. For walking, BI LSTM and BI GRU has slight difference. For upstairs and

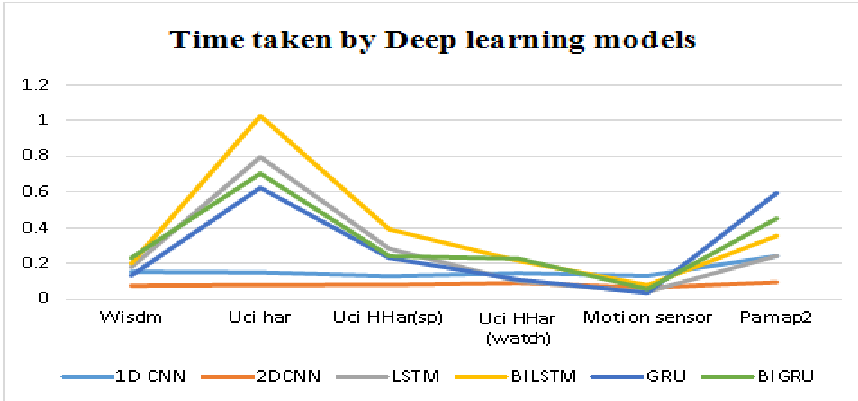

Fig. 4. Computation time of deep learning models

Table 2. Classification accuracy of ML and DL models with multiple datasets

Classifiers	Classification Accuracy(%) for below Data sets					
	Wisdm	**Uci har**	**Uci HHar(sp)**	**Uci HHar(watch)**	**Motion sensor**	**Pamap2**
KNN	0.72	0.89	0.75	0.81	0.64	0.56
NB	0.52	0.77	0.66	0.66	0.89	0.76
DT	0.68	0.85	0.71	0.79	0.66	0.58
XG Boost	0.78	0.87	0.71	0.85	0.75	0.72
SVM	0.72	***0.95***	0.72	0.79	0.87	0.82
RF	0.75	0.88	0.77	0.83	0.69	0.67
1D CNN	0.93	0.91	0.83	0.81	0.85	0.87
2D CNN	***0.95***	**0.96**	**0.92**	0.88	0.89	0.92
LSTM	0.88	0.89	0.83	0.81	0.90	0.92
BILSTM	***0.95***	0.91	**0.92**	***0.90***	0.92	0.94
GRU	0.92	0.88	***0.86***	0.88	***0.93***	***0.95***
BIGRU	**0.96**	0.93	**0.92**	**0.92**	**0.94**	**0.96**

*sp-smartphone

Table 3. Activity wise accuracy of 2D CNN

Activities	Activity wise accuracy(%)for below Data sets					
	Wisdm	**Uci har**	**Uci HHar(sp)**	**Uci HHar(watch)**	**Motion sensor**	**Pamap2**
Walking	0.94	0.92	0.80	0.73	0.99	0.82
Sitting	1.00	0.96	1.00	0.93	1.00	1.00
Standing	1.00	0.98	1.00	0.93	1.00	1.00
Downstairs	1.00	0.88	0.66	0.59	0.90	0.90
Upstairs	0.83	0.96	0.80	0.80	0.94	0.84
Laying	NA	1.00	NA	NA	NA	1.00
Jogging	0.94	NA	NA	NA	0.97	0.92
Bike	NA	NA	0.97	0.90	NA	NA

*NA-Not Available

downstairs, BI GRU gained superiority compared to the other models. From the

Table 4. Activity wise accuracy of BI LSTM

Activities	Activity wise accuracy(%)for below Data sets					
	Wisdm	**Uci har**	**Uci HHar(sp)**	**Uci HHar(watch)**	**Motion sensor**	**Pamap2**
Walking	0.95	0.90	0.80	0.93	0.94	0.95
Sitting	1.00	0.96	1.00	1.00	1.00	1.00
Standing	1.00	0.98	1.00	0.93	1.00	1.00
Downstairs	0.80	0.76	0.66	0.86	0.89	0.90
Upstairs	1.00	0.91	0.80	0.65	0.94	0.84
Laying	NA	1.00	NA	NA	NA	1.00
Jogging	0.95	NA	NA	NA	0.97	0.94
Bike	NA	NA	0.97	1.00	NA	NA

*NA-Not Available

Table 5. Activity wise accuracy of BI GRU

Activities	Activity wise accuracy(%)for below Data sets					
	Wisdm	**Uci har**	**Uci HHar(sp)**	**Uci HHar(watch)**	**Motion sensor**	**Pamap2**
Walking	0.94	0.90	0.83	0.80	0.89	0.94
Sitting	1.00	0.96	0.97	1.00	1.00	1.00
Standing	1.00	0.98	1.00	1.00	1.00	1.00
Downstairs	0.94	0.86	0.90	0.90	0.89	0.96
Upstairs	0.94	0.92	0.86	0.83	0.94	0.87
Laying	NA	1.00	NA	NA	NA	1.00
Jogging	0.94	NA	NA	NA	1.00	0.94
Bike	NA	NA	0.97	0.97	NA	NA

*NA-Not Available

results, it can be concluded that **BI GRU** model can discriminate well compared to the other models and this model can predict quickly and accurately.

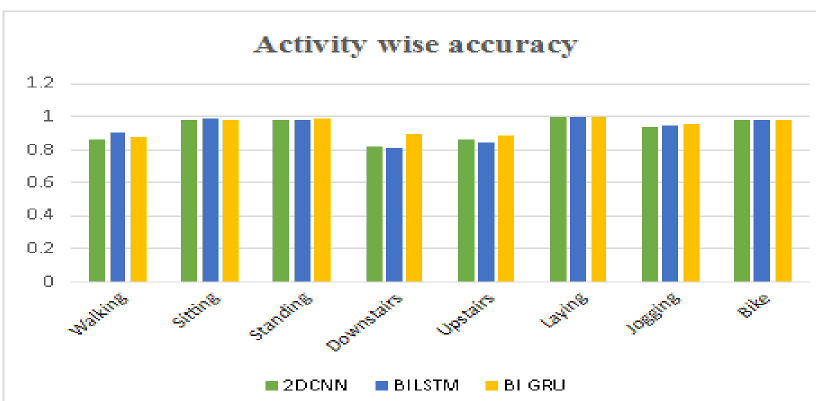

Fig. 5. Activity Wise accuracy for 2D CNN ,BI LSTM,BI GRU

5 Conclusion

In this work we have studied performance of machine learning including deep learning tools for human activity recognition. This study has been done on multiple benchmark data sets and analysis of those results are presented. We observed that deep learning tools produce better classification accuracy for all the data sets. It is also observed that bidirectional LSTM and GRU perform better than uni-directional LSTM and GRU though the computational cost is slightly high. Activity wise accuracy for the models which achieved higher accuracy are represented. From that analysis, it is observed that BI GRU model provides the best performance and can be utilized for the implementation of HAR applications.

References

1. S. Majumder, T. Mondal and M. Jamal Deen, Wearable sensors for remote health monitoring, Sensors, Vol 17(1):Pp.130–175, 2017.
2. Jehad Sarkar, La The Vinh, Young-Koo Lee, et al. GPARS: A general purpose activity recognition system ,Applied Intelligence, Vol 35(2):Pp.242- 259, 2011.
3. A.Lentzas, D.Vrakas,Non-intrusive human activity recognition and abnormal behavior detection on elderly people: A review. Artif. Intell. Rev. 53,Pp.1975–2021, 2020.
4. T.Subetha and S.Chitrakala, A survey on human activity recognition from videos, International Conference on Information Communication and Embedded Systems(ICICES), Chennai,Pp.1-7, 2016.
5. A.Bulling,U.Blanke, B.Schiele, A tutorial on human activity recognition using body worn inertial sensors. ACM Comput. Surveys, Article No. 33, 2014.
6. W.Sousa Lima , E.Souto, K.El Khatib,R.Jalali,J.Gama. Human Activity Recognition Using Inertial Sensors in a Smartphone: An Overview. Sensors, 19(14):3213,2019.
7. J.Wang,R.Chen,X.Sun,M. F.She,Y. Wu, Recognizing human daily activities from accelerometer signal. Procedia Engineering, 15,Pp.1780-1786, 2011.
8. A.Bayat,M.Pomplun, & D. A.Tran, A study on human activity recognition using accelerometer data from smartphones. Procedia Computer Science, Vol.34, Pp.450-457, 2014.
9. D.Anguita, A.Ghio, L.Oneto, X.Parra, and J.L.Reyes Ortiz. Human activity recognition on smartphones using a multi class hardware-friendly support vector machine.International workshop on ambient assisted living, Pp. 216–223, 2012.
10. P.Paul and T. George, An effective approach for human activity recognition on smartphone, IEEE International Conference on Engineering and Technology (ICETECH), Coimbatore, Pp. 1-3, 2015.
11. L. Fan, Z. Wang and H. Wang, Human Activity Recognition Model Based on Decision Tree, International Conference on Advanced Cloud and Big Data, Nanjing, Pp.64-68, 2013.
12. M.H.Kolekar,D.P. Dash, Hidden markov model based human activity recognition using shape and optical flow based features.Proc. of the IEEE Region 10 Conference (TENCON), Pp.393–397, 2016.
13. P.Casale, O.Pujol,P.Radeva, Human activity recognition from accelerometer data using a wearable device,Proc. of the Iberian Conference on Pattern Recognition and Image Analysis, Pp. 289–296, 2011.

14. W.Wu,S.Dasgupta,E.E.Ramirez,C.Peterson,G.J.Norman,Classification accuracy's of physical activities using smartphone motion sensors, J. Med Internet Res, 14(5), e130, 2012
15. K.Chen,D. Zhang,L.Yao, B.Guo, Z.Yu,Y.Liu, Deep learning for sensor based human activity recognition: overview, challenges and opportunities. https://arxiv.org/abs/2001.07416 ,2020.
16. C.A.Ronao, , & S. B. Cho,, Human activity recognition with smartphone sensors using deep learning neural networks. Expert systems with applications, 59,Pp.235-244,2016.
17. T.Zebin, P.J.Scully,K.B.Ozanyan, Human activity recognition with inertial sensors using a deep learning approach. In Proc. of the IEEE SENSORS, Pp. 1–3, 2016.
18. M. Gholamrezaii and S. M. Taghi Almodarresi, Human Activity Recognition Using 2D Convolutional Neural Networks, 27th Iranian Conference on Electrical Engineering (ICEE), Yazd, Iran,Pp.1682-1686, 2019.
19. N.Y.Hammerla , S.Halloran, T.Ploetz, Deep Convolutional and Recurrent Models for Human Activity Recognition using Wearable's,Journal of Scientific Computing, 61(2):Pp.454-476,2016.
20. D.Arifoglu,A. Bouchachia,Activity recognition and abnormal behaviour detection with recurrent neural networks. Procedia Comput.Sci.,110, Pp.86–93,2017.
21. S. Ishimaru, K. Hoshika, K. Kunze, et. al.,Detecting reading activities by EOG glasses and deep neural networks, In Proc. of the ACM International Joint Conference on Pervasive and Ubiquitous Computing, Pp.704–711, 2017.
22. A. A. Aljarrah and A. H. Ali, Human Activity Recognition using PCA and BiLSTM Recurrent Neural Networks,2nd International Conference on Engineering Technology and its Applications (IICETA), Al-Najef, Iraq,Pp.156-160, 2019.
23. S. Yao, S. Hu, Y. Zhao, A. Zhang, and T. Abdelzaher. Deepsense: A unified deep learning framework for time-series mobile sensing data processing, In Proc. of the 26th International Conference on World Wide Web, Pp.351–360, 2017.
24. T. Alsarhan, L. Alawneh, M. Al-Zinati and M. Al-Ayyoub, Bidirectional Gated Recurrent Units For Human Activity Recognition Using Accelerometer Data, In Proc. of IEEE SENSORS, Pp. 1-4, 2019,
25. K. Xia, J. Huang and H. Wang, LSTM-CNN Architecture for Human Activity Recognition, in IEEE Access, vol.8, Pp. 56855-56866, 2020.
26. S. Abbaspour ,F. Fotouhi, A. Sedaghatbaf ,H. Fotouhi et.al., A Comparative Analysis of Hybrid Deep Learning Models for Human Activity Recognition, Sensors (Basel), 20(19), 5707, 2020.
27. J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," ACM SigKDD Explorations Newsletter, Vol. 12(2), Pp. 74–82, 2011.
28. D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes Ortiz. A public domain dataset for human activity recognition using smartphones.,In Proc. ESANN, 2013.
29. A. Stisen, et.al., Smart devices are different: Assessing and mitigatingmobile sensing heterogeneities for activity recognition. In Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems. ACM,Pp.127–140. 2015.
30. M. Malekzadeh, R. G. Clegg, A. Cavallaro, and H. Haddadi, "Protecting sensory data against sensitive inferences," in Proceedings of the Workshop on Privacy by Design in Distributed Systems (W-P2DS18), 2018.
31. Reiss, A.; Stricker, D. Introducing a new benchmarked dataset for activity monitoring. In Proceedings of the 16th International Symposium on Wearable Computers, Pp.108–109, 2012.