



Particle Swarm Optimization of Convolutional Neural Networks for Human Activity Prediction

Preethi Gunishetty Devarakonda and Bojan Božić

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

November 26, 2020

Particle Swarm Optimization of Convolutional Neural Networks for Human Activity Prediction

Preethi Gunishetty Devarakonda and Bojan Bozic

School of Computer Science, Technological University Dublin
D18128978@mytudublin.ie, bojan.bozic@tudublin.ie

Abstract. The increased usage of smartphones for daily activities has created a huge demand and opportunities in the field of ubiquitous computing to provide personalized services and support to the user. In this aspect, Sensor-Based Human Activity Recognition (HAR) has seen an immense growth in the last decade playing a major role in the field of pervasive computing by detecting the activity performed by the user. Thus, accurate prediction of user activity can be valuable input to several applications like health monitoring systems, wellness and fit tracking, emergency communication systems etc.,

Thus, the current research performs Human Activity Recognition using a Particle Swarm Optimization (PSO) based Convolutional Neural Network which converges faster and searches the best CNN architecture. Using PSO for the training process intends to optimize the results of the solution vectors on CNN which in turn improve the classification accuracy to reach the quality performance compared to the state-of-the-art designs. The study investigates the performances PSO-CNN algorithm and compared with that of classical machine learning algorithms and deep learning algorithms. The experiment results showed that the PSO-CNN algorithm was able to achieve the performance almost equal to the state-of-the-art designs with a accuracy of 93.64%. Among machine learning algorithms, Support Vector machine found to be best classifier with accuracy of 95.05% and a Deep CNN model achieved 92.64% accuracy score.

Keywords: Human Activity Recognition, Particle Swarm Optimisation, Convolutional Neural Network, Time Series Classification, Deep Learning, Sensors

1 Introduction

Activity Recognition aims at identifying the activity of users based on series of observations collected during the activity in a definite context environment. Applications that are enabled with activity recognition are gaining huge attention, as users get personalized services and support based on their contextual behaviour. The proliferation of wearable devices and smartphones has provided real-time monitoring of human activities through sensors that are embedded in smart devices such as proximity sensors, cameras, microphone, magnetometers,

accelerometers, gyroscopes, GPS etc., Thus, understanding human activities in inferring the gesture or position has created a competitive challenge in building personal health care systems, examining wellness and fit characteristics, and most pre-dominantly in elderly care, abnormal activity detection, diabetes or epilepsy disorders etc.,

Initially, Human Activity Recognition (HAR) experiment was carried out by attaching one or more dedicated on-body sensors to specific parts of human body to collect time series data [8] As, the usage of smart phones for daily activities has increased extensively, HAR research has employed to collect data from built-in sensors embedded in smart phones [17]. The raw data from the sensors are analysed using several machine learning and deep learning algorithms to classify the activity with appropriate evaluation metric. The activity recognition performance has significantly made strides since the of research, but the experiment set up can be varied, for example, the types of exercises performed by human subjects, the sorts of sensors utilized, the rate at which signal is sampled , the segment length of time series data. Apart from choosing classifier learning algorithms, the approaches are varied in terms of applying various feature processing techniques namely feature selection, extraction and transformation. These choices made comparative evaluation of different Human Activity Recognition (HAR) approaches complex. Thus, Human Activity Recognition (HAR) plays a significant part in enhancing people's lifestyle, as it should be competent enough in learning high level quality information from raw sensor data. Effective HAR applications are incorporated for contextual behaviour analysis [1], video surveillance analysis [9], gait investigation (to determine any abnormalities in walking or running), gesture and position recognition [10].

2 Related Work

This section provides an overview on Human Activity Recognition and its applications. Various approaches for HAR task are discussed. Particularly, Sensor based HAR is detailed with different sensor modalities. This chapter also gives an overview of the modelling approaches for HAR. Each modelling approach is discussed with its theory and its applicability in HAR.

2.1 Sensor Based Human Activity Recognition

Due to the advancement in ubiquitous computing, Activity Recognition has been one of the major research areas in mobile technology that has seen the rapid demand over the past few years. This covers major areas like smart homes [11], wellness and fitness monitoring [14], video surveillance analysis for security purposes [18], behavioural analysis, emergency services [20] etc., Due to the immense growth of sensor technology and ubiquitous computing, sensor-based Human Activity Recognition is gaining attention which is widely used with enhanced protection and privacy. According to [3], the HAR task can be achieved by placing the sensors at different locations to recognize human activity for specific

context. Wearable sensors are one of the widely used sensor modalities in HAR. These sensors are often worn or attached to the users, namely an accelerometer, gyroscope, and magnetometer. As the human body moves, the acceleration and angular velocity are varied, this data is further analysed to predict the activity. Thus, wearable sensors were widely used for HAR [19] in various health monitoring systems. In recent days, inertial sensing, that uses movement-based Sensors which can be attached on user's body has been studied widely [21].

2.2 Modelling Approaches for HAR

Due to the natural ordering of the temporal feature data, the Human Activity Recognition is considered as a typical pattern-recognition system where it involves classifying the human activity based on the series of data. The main difference between Machine Learning Algorithms and Deep Learning Algorithms in recognising human activity is the way the input features are extracted.

Machine Learning: In case of Machine Learning Approach, the raw inertial activity signals received from the sensors are subjected to feature -extraction process by domain knowledge experts [2]. The features that are usually extracted are based on two main domain features namely; time domain and frequency domain. Some research works employed Machine Learning Approach to perform HAR with hand-crafted features faced low performance as only shallow features are explored and learned by the classifiers [19]. Before deep learning was used extensively, shallow neural network classifiers, that is Multi-Layer Perceptron (MLP), was considered to be a promising algorithm for HAR. In this aspect, [6] performed HAR with algorithms like logistic regression, decision tree and MLP and MLP outperformed the other two models. This process requires domain expertise . Sometimes this process may lead to loss of significant data points

Deep Learning: Where as in Deep Learning Algorithms, the raw sensor signals collected from inertial sensors (accelerometer, gyroscope etc,.) are directly subjected to modelling, where no feature extraction step is performed [5]. Several Deep Learning Algorithms like LSTM [7], RNN [4], Restricted Boltzmann Machines [13], CNN [16, 12, 15] was utilized to perform HAR tasks. However, the trial and error method for selecting parameters to select the best model does not guarantee an optimal performance, which requires user to continuously observe the performance trend. Thus, in order to get outstanding results, one has to be expert in the model architecture and also domain.

2.3 Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a nature inspired, meta-heuristic algorithm often used for discrete, continuous and sometimes for combination optimization problems. The main ideology behind PSO is that each particle is well known of its velocity and the best configuration achieved in the past (pBest),

and the particle which is the current global best configuration in the swarm of particles(gBest). Hence, at every current iteration, each particle updates its velocity in such a way that its new position will be close enough to global gBest and its own pBest at the same time. The velocity and particle vector are adjusted based to the following equations 1 and 2 respectively:

$$v_{id}(t + 1) = w * v_{id}(t) + c_1 * r_1 * (P_{id} - x_{id}(t)) + c_2 * r_2 * (P_{gd} - x_{id}(t)) \quad (1)$$

$$x_{id}(t + 1) = x_{id}(t) + v_{id}(t + 1) \quad (2)$$

where v_{id} indicates the velocity of i_{th} particle in the d_{th} dimension, z_{id} indicates the position of i_{th} particle in the d_{th} dimension, P_{id} and P_{gd} represents the local best and the global best in the d_{th} dimension, r_1 and r_2 are the random numbers between the range 0 and 1, c_1 , c_2 and w , are acceleration coefficient for exploitation, acceleration coefficient for exploration and inertia weight respectively.

Deep learning networks have gained better results with less efforts in parameter settings. In particular, Deep Convolutional Neural Networks are used extensively due to its flexibility in both data driven approach (Using 1D Convolution for signal data) and model driven approach (data transformation of signal data to a 2D image). In order to gain higher performance of the model, several layers has to be used and parameter initialization has to be done carefully. This needs a detailed knowledge on CNN architecture and also on the dataset.

Thus, to find the optimal CNN architecture automatically without human intervention, a meta heuristic algorithm Particle Swarm Optimisation is utilized which is easy to implement with lower computational cost.

3 Dataset Used

The dataset used in this study is downloaded from UCI Machine Learning Repository created at SmartLab, one of the Research Laboratories at DIBRIS at University of Genova. They experimented on a group of 30 volunteers within a range of age be-

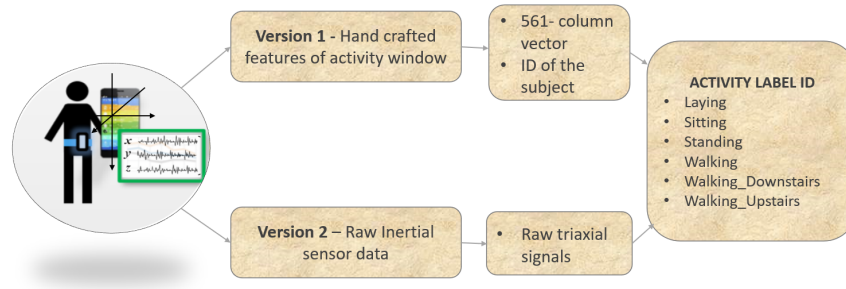


Fig. 1. Dataset

tween 19-48 years who were performing daily activities like Sitting, Standing, Laying, Walking, Walking Upstairs and Walking Downstairs. Each subject are volunteer performed daily activities like while carrying a smartphone that is waist-mounted. The smartphone was embedded with inertial sensors. Thus, With the help of this embedded accelerometer and gyroscope, 3-axial linear acceleration and 3-axial angular velocity were captured. For experimental purposes, two versions of the dataset is provided which is showed in figure 1

For version - 1 With the help of domain experts, time and frequency domain features are extracted to get 561 columns data. Along with that, it consists of ID of six activities and also the subject ID who is performing the experiment.

For version - 2 The signals collected from raw inertial data is taken without any feature extraction and respective activity ID is labelled. Thus, as the target ID six activities are identified

4 Methodology

The Cross Industry Standard Process for Data Mining (CRISP-DM), is well proven methodology with a structured approach. This is employed to conduct the current study and methodology possess flexibility, practicality and is idealised with a sequence of events. The figure 2 shows the design flow to be followed for the current research.

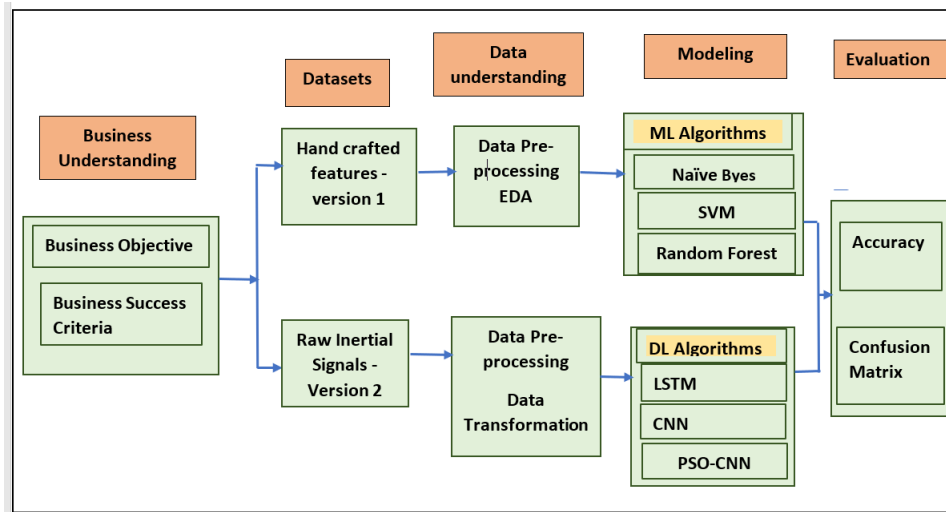


Fig. 2. Design Flow

The experiment begins with the Business Understanding phase, which indicates what is to be accomplished from a business perspective. The expected outputs of this phase

form the main objectives of the project. Here the insights and goals of the project are defined. In order to answer the research question, the experiment is conducted with two versions of the datasets which is explained in Data Understanding Phase. Additionally, data description report is prepared to understand each filed description. This is done separately for both the datasets.

The third phase is the Data preparation stage. Here the data is checked for duplicates, null records and appropriate action is taken to address them. Further, new derived fields can be formed based on the domain knowledge. Data from multiple databases are integrated to form the final dataset for modelling. The fourth step is the modeling stage.

Based on initial analysis done from the literature review, suitable modelling technique is chosen and applied on the two versions of the dataset. Next phase is Evaluation phase. The performance of the modelling algorithms is done using various measures. HAR is a Multilabel classification problem. The main challenge in classification task is to correctly classify the target variables. Only accuracy score cannot give us the overall performance of the model. Hence, confusion matrix which gives the actual number of correct and incorrect predictions made for each target class is considered. Additionally, precision, recall and f1 score is computed. But for comparisons accuracy and f1 score are considered. Based on the evaluation criteria, models are evaluated to see if it meets the business objective.

4.1 Algorithm Of Particle Swarm Optimization Based CNN

The Figure 3 shows the working of the model. Though CNN's have showed good results in HAR, there are multiple parameters to take care to find the optimal CNN architecture. The main focus of any neural network is to minimize the error between training targets and predicted outputs. It is cross-entropy in case of CNN's, which is carried out by backpropagation and gradient descent. Even a simple CNN's have many parameters to tune them. Thus, it is significant to find algorithms which finds and evaluates CNN architecture with less time. Thus, motivated from this, a new PSO-CNN is utilized for Human Activity Recognition. The working of PSO-CNN can be divided into five stages as below -

- CNN Training - The CNN is trained with some pre-defined weights initialized. . It uses a CNN with 1D convolutional layer, since the HAR dataset consist of signals in shape [samples, time.steps, no of features]. The output is one hot vector encoded which is 6 (target activity to be predicted).
- Pre-PSO Training - Here weights are captured from CNN training and it is converted to particle.
- Particle Swarm Optimization Training - After initializing the values of convergence, cognitive value, social value, number of particles, stopping condition and number of epochs, PSO algorithms searches the hyperplane for optimized vector using the CNN loss function .
- Update CNN Architecture - Using the values of weight in previous phase, the final results are computed. A new CNN architecture is created based on these weights rather than basis of the output.
- Computation of Prediction Accuracy and Results - The evaluation itself is done by comparing the loss function of each particle Thus, the objective of the algorithm is to find a particle architecture with the smallest loss, regardless of the number of parameters or other criteria.

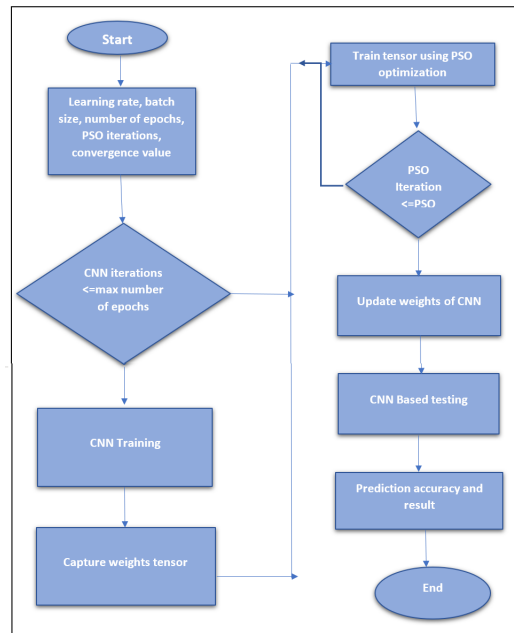


Fig. 3. Particle Swarm Optimization Training for Convolutional Neural Network

5 Implementation

The implementation can be divided into parts that is with ML using Hand crafted features and other experiment using DL with raw inertial sensor data.

5.1 With Version 1 - Hand crafted features dataset

The steps involved performing this is illustrated in figure 3 .The data is merged from various files present in the repository. As a part of data understanding, using matplotlib in python, the distribution of data is analyzed. where it is observed that target level distribution is uniform. The range of time and frequency domain variables were seen between +1 and -1, indicating normalization were performed while creating the dataset. The data is checked for missing values and duplicate rows. No such scenarios were found over the entire data. The data is already split to train and test and size of the data is (7352, 564) samples and 564 is the number of features. Further the data is subjected to modeling algorithms like Random Forest The three algorithms were executed with default parameter setting.

5.2 With Version 2 - Raw Inertial sensor data

The steps involved performing with Version -2 Data is illustrated in figure 4. Raw inertial signal data is gathered and merged from different files. It is checked for null values and duplicates. The data shape indicates that 7352 is the number of samples

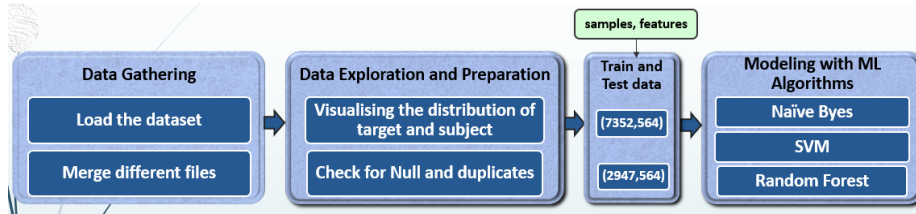


Fig. 4. Modelling With Version 1 - Hand crafted feature dataset

from a raw signal data file, 128 indicates the number of of timesteps that series of data is partitioned. 9 indicates total variables for each time steps. Finally it is subjected to modeling with LSTM, CNN, PSO-CNN - it is ensured that the model is not overfitted.

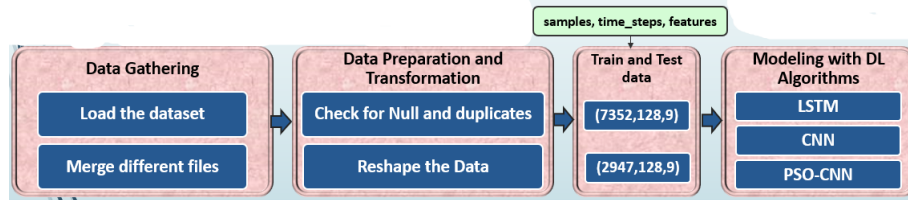


Fig. 5. Modeling With Version 2 - Raw Inertial sensor data

5.3 Modeling With PSO-CNN

The modelling for PSO-CNN can be categorized into three parts namely Particle Setting for Particle Swarm Optimization (Number of iterations-10, Swarm Size - 20, Cg - 0.5) , Parameter Settings for initializing CNN architecture (Displayed in figure 6),

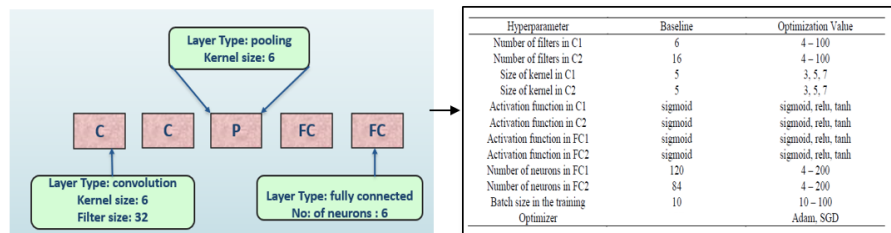


Fig. 6. Each Particle as CNN architecture (left) and Parameters Settings used for PSO-CNN (right)

Parameter Settings for training Convolutional Neural Network(epochs for particle evaluation - 20, epochs for global best - 256,Dropout rate - 0.5, Batch normalizer layer outputs - yes). The parameters used is described in figure below 6 (right). Thus each particle is nothing but a CNN architecture as displayed in figure 6 (left).

6 Results and Discussion

The performance of PSO-CNN is evaluated against Machine Learning and Deep Learning Algorithms

6.1 Comparing PSO-CNN with machine learning algorithms

The analysis of results is performed by comparing PSO-CNN with Machine Learning Algorithms. The below Figure 5.13 shows the results. From the table, it is evident that

Dataset	Algorithms	Accuracy	F1 score
Hand-crafted features -Version 1	Naïve Byes	77.02%	71.59%
	Support Vector Machine	95.04%	95.1%
	Random Forest	92.36%	91.78%
Raw Inertial Signal – Version 2	PSO-CNN	93.64%	93.62%

Fig. 7. Comparison of PSO-CNN with Machine Learning Algorithms

Support Vector Machine achieved accuracy of 95.04% and F1 score 95.1%. The machine learning models were built using Hand-crafted features -Version 1 Dataset. The model achieved satisfactory results without performing any Data Dimensionality reduction techniques. On the other hand, PSO-CNN also achieved considerable results with raw sensor data with accuracy of 93.64%. However, the hand-crafted feature extraction process requires human effort to manually design the features.

6.2 Comparing PSO-CNN with Deep learning algorithms

The analysis of results is performed by comparing PSO-CNN with Deep Learning Algorithms. The below Figure 5.14 shows the results. From the table, it is clear that PSO-CNN was able to achieve high performance of accuracy when compared with LSTM and CNN models. LSTM performance was low with accuracy 84.71% and F1 score with 84.42%. This, PSO-CNN gained better results than the state-of-the art CNN model. For a classification problem, the capability of the modeling algorithm to classify each target class correctly also plays a major role. Each The algorithm’s ability to classify each activity like walking, sitting, laying are discussed in section 5.1 . From the classification report of PSO-CNN Figure 5.11,it is evident that PSO-CNN was able to classify most number of activities correctly.

Dataset	Algorithms	Accuracy	F1 score
Raw Inertial Signal – Version 2	LSTM	84.71%	84.42%
	CNN	92.64%	92.71%
	PSO-CNN	93.64%	93.62%

Fig. 8. Comparison of PSO-CNN with Deep Learning Algorithms

6.3 Evaluation of PSO-CNN

As a further discussion, the classification accuracy for each activity is compared between PSO-CNN and CNN. We can observe that PSO-CNN was able to classify all the



Fig. 9. Confusion Matrix of CNN and PSO-CNN

activity with least errors except for activities STANDING and SITTING. These are less compared to actual CNN model. This can be illustrated by Confusion Matrix of PSO-CNN and CNN model in figure 5. PSO-CNN model was built to find the best CNN architecture with minimum effort. This also overcomes the local minima problem of the backpropagation training algorithms. The experiment was conducted with 20 epochs, which is less than base CNN model. On an overall note, we can say that PSO-CNN achieved higher accuracy with that of the Deep Learning algorithms but failed to reach the accuracy of Machine learning models.

7 Future Work and Recommendations

Detailed literature review was performed emphasizing on the applications of Human Activity Recognition in various fields. In particular, Sensor Based HAR is highlighted for the readers. This also detailed about the current state of the art techniques in HAR.

A systematic investigation is done for importing two versions of the sensor datasets. This can be used as reference for future works. Illustrated that PSO based CNN proved to be the best classifier for data where human-engineered feature knowledge is not needed. Additionally, the work tries to enhance the performance of state-of-the-art design of the CNN model by using Optimisation. This adds up to the generalization of using PSO-CNN model for other Activity Recognition tasks.

In the current research PSO algorithm is used to find the optimal architectures in deep convolutional neural network. Furthermore it make use of the benefits of global and local exploration capabilities of the particle swarm optimization technique PSO and the gradient descent back-propagation thereby to form a efficient searching algorithm this is because the performance of of deep convolution network extremely depends on their network structure used and hyper-parameter selections.

In order to find the best hyper parameters lot of training time is employed which requires the deep understanding of CNN architecture and also the domain knowledge. Hence PSO-CNN is employed to optimize these parameter configurations and through which efficient parameters are evolved that would increase the performance with less training time.

The current research can be explored and improved in many ways so as to improve the human activity recognition tasks. The proposed approach also provides a flexible methodology where one can change the initial parameter settings of both PSO and CNN. In this way a trade-off between the model generalization capabilities and complexity of the model can be justified. From the experimental results it is illustrated that PSO has been shown to converge faster and find the best configuration with less training time. This exceed he performance of state-of-the-art results obtained in the domain of HAR. To some extent, the algorithm failed to recognize the similar he activities like WALKING_UPSTAIRS and WALKING, LAYING and SITTING . This may be due to the insufficient data. The solution can be further explored with large time series data.

The experiment is conducted to explore the capability of deep learning algorithms in HAR tasks. In order to generalize the model capability, this can be applied to other Activity Recognition tasks which includes Time Series data.

References

1. Aurangzeb, K., Haider, I., Khan, M.A., Saba, T., Javed, K., Iqbal, T., Rehman, A., Ali, H., Sarfraz, M.S.: Human behavior analysis based on multi-types features fusion and von nauman entropy based features reduction. *Journal of Medical Imaging and Health Informatics* **9**(4), 662–669 (2019)
2. Bengio, Y.: *International conference on statistical language and speech processing* (2013)
3. Chavarriaga, R., Sagha, H., Calatroni, A., Digumarti, S.T., Tröster, G., Millán, J.d.R., Roggen, D.: The opportunity challenge: A benchmark database for on-body sensor-based activity recognition. *Pattern Recognition Letters* **34**(15), 2033–2042 (2013)
4. Edel, M., Köppe, E.: Binarized-blstm-rnn based human activity recognition. In: *2016 International conference on indoor positioning and indoor navigation (IPIN)*. pp. 1–7. IEEE (2016)
5. Figo, D., Diniz, P.C., Ferreira, D.R., Cardoso, J.M.: Preprocessing techniques for context recognition from accelerometer data. *Personal and Ubiquitous Computing* **14**(7), 645–662 (2010)

6. Godino-Llorente, J.I., Gomez-Vilda, P.: Automatic detection of voice impairments by means of short-term cepstral parameters and neural network based detectors. *IEEE Transactions on Biomedical Engineering* **51**(2), 380–384 (2004)
7. Guan, Y., Plötz, T.: Ensembles of deep lstm learners for activity recognition using wearables. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* **1**(2), 1–28 (2017)
8. He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: *Proceedings of the IEEE international conference on computer vision*. pp. 1026–1034 (2015)
9. Hwang, M.G., Park, H.J., Har, D.H.: A method of smart phone original video identification by using unique compression ratio pattern. *Forensic science international* **304**, 109889 (2019)
10. Kang, W., Zhang, Y., Dong, X.: Body gesture recognition based on polarimetric micro-doppler signature and using deep convolutional neural network. *Progress In Electromagnetics Research* **79**, 71–80 (2019)
11. Li, Q., Ning, H., Zhu, T., Cui, S., Chen, L.: A hybrid approach to inferring the internet of things for complex activity recognition. *EURASIP Journal on Wireless Communications and Networking* **2019**(1), 1–11 (2019)
12. Li, X., Zhang, Y., Marsic, I., Sarcevic, A., Burd, R.S.: Deep learning for rfid-based activity recognition. In: *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM*. pp. 164–175 (2016)
13. Liu, Z., Wu, M., Zhu, K., Zhang, L.: Sensafe: A smartphone-based traffic safety framework by sensing vehicle and pedestrian behaviors. *Mobile Information Systems* **2016** (2016)
14. Nawaratne, R., Alahakoon, D., De Silva, D., Kumara, H., Yu, X.: Hierarchical two-stream growing self-organizing maps with transience for human activity recognition. *IEEE Transactions on Industrial Informatics* (2019)
15. Ravi, D., Wong, C., Lo, B., Yang, G.Z.: Deep learning for human activity recognition: A resource efficient implementation on low-power devices. In: *2016 IEEE 13th international conference on wearable and implantable body sensor networks (BSN)*. pp. 71–76. IEEE (2016)
16. Singh, M.S., Pondenkandath, V., Zhou, B., Lukowicz, P., Liwicki, M.: Transforming sensor data to the image domain for deep learning—an application to footstep detection. In: *2017 International Joint Conference on Neural Networks (IJCNN)*. pp. 2665–2672. IEEE (2017)
17. Subasi, A., Fllatah, A., Alzobidi, K., Brahim, T., Sarirete, A.: Smartphone-based human activity recognition using bagging and boosting. *Procedia Computer Science* **163**, 54–61 (2019)
18. Voicu, R.A., Dobre, C., Bajenaru, L., Ciobanu, R.I.: Human physical activity recognition using smartphone sensors. *Sensors* **19**(3), 458 (2019)
19. Yang, J., Nguyen, M.N., San, P.P., Li, X., Krishnaswamy, S.: Deep convolutional neural networks on multichannel time series for human activity recognition. In: *Ijcai*. vol. 15, pp. 3995–4001. Citeseer (2015)
20. Yin, J., Yang, Q., Pan, J.J.: Sensor-based abnormal human-activity detection. *IEEE Transactions on Knowledge and Data Engineering* **20**(8), 1082–1090 (2008)
21. Yu, H., Cang, S., Wang, Y.: A review of sensor selection, sensor devices and sensor deployment for wearable sensor-based human activity recognition systems. In: *2016 10th International Conference on Software, Knowledge, Information Management & Applications (SKIMA)*. pp. 250–257. IEEE (2016)