HAL
open science

Wearable Sensor Data Based Human Activity Recognition using Deep Learning: A new approach

Phuong Hanh Tran, Thong Nguyen, Kim Phuc Tran, Cédric Heuchenne

To cite this version:

Phuong Hanh Tran, Thong Nguyen, Kim Phuc Tran, Cédric Heuchenne. Wearable Sensor Data Based Human Activity Recognition using Deep Learning: A new approach. FLINS/ISKE 2020, Sep 2020, Cologne (on line), Germany. pp.581-588, 10.1142/9789811223334_0070 . hal-02547033v2

HAL Id: hal-02547033

https://hal.science/hal-02547033v2

Submitted on 3 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Wearable Sensor Data Based Human Activity Recognition using Deep Learning: A new approach

Phuong Hanh Tran

HEC Liège - Management School of the University of Liège,
Liège, 4000, Belgium
** Corresponding author; E-mail: phtran@uliege.be*

Quoc Thong Nguyen

HEC Liège - Management School of the University of Liège,
Liège, 4000, Belgium

GEMTEX Laboratory, Ecole Nationale Supérieure des Arts et Industries Textiles,
Roubaix, 59560, France
E-mail: qtnguyen@uliege.be

Kim Phuc Tran

GEMTEX Laboratory, Ecole Nationale Supérieure des Arts et Industries Textiles,
Roubaix, 59560, France
E-mail: kim-phuc.tran@ensait.fr

Cédric Heuchenne

HEC Liège - Management School of the University of Liège,
Liège, 4000, Belgium
E-mail: C.Heuchenne@ulg.ac.be

With a tremendous increase in mobile and wearable devices, the study of sensor-based activity recognition has drawn a lot of attention in the past years. In recent years, the applications of Human Activity Recognition are getting more and more attention, especially in eldercare and healthcare as an assistive technology when combined with the Internet of Things. In this paper, we propose three deep learning approaches to improve the accuracy of activity detection on the WISDM dataset. Particularly, we apply a convolutional neural network to extract the interesting features, then we use softmax function, support vector machine, and random forest for classification tasks. The results show that the hybrid algorithm, convolutional neural network combined with the support vector machine, outperforms all the previous methods in classifying every activity. In addition, not only the support vector machine but also the random forest shows better accuracy in classification task than the neural

network classification and the former approaches do.

Keywords: Human activities recognition, Smart healthcare, wearable technologies, Time series data, sensor.

1. Introduction

Nowadays, mobile devices and especially smart wearable devices, such as cellphones, smartwatch; these devices are equipped with many sensors (e.g. GPS, accelerometers, gyroscope). Human activity data become more popular and easier to collect. Innumerable applications from the recognition of human activities can be visualized in medical, security, entertainment, and tactical scenarios. Especially in the healthcare system, one can apply Human Activity Recognition (HAR) to the detection and prevention of many chronic diseases[1]. Moreover, with the emergence of the Internet of Things (IoT) tracking and reacting to a special patient using wearable devices is more practical and more reliable[2].

The recognition of human activities has been approached in two main ways based on the procedure to gather the data, namely using *external* and *wearable* sensors. In the former, the devices are fixed in points of interest, so the activities information entirely depends on the voluntary of the users with the sensors. On the other hand, in the latter, the sensors are attached to the human body, the smartwatch is a typical example. There are numerous studies of HAR conducted in the literature. With the extensive use of cellphones, limitations and suggestions for online activity recognition using mobile phones are described in[3]. A survey on HAR using wearable sensors describes the problem clearly by providing a generic data acquisition architecture for HAR systems[2]. A tutorial was given on HAR using body-worn inertial sensors in[4]. The more recent reviews on human activity monitoring using wearable sensors can be referred to[5].

In the garment industry, there are some studies on using wireless sensor technologies to develop a smart clothes service platform for physiological monitoring[6,7]. With the development of thin film-based sensors, the implantable wireless sensors on the garment have an innovative application in healthcare monitoring[8].

From the raw data, the features are generated based on the n readings contained within each t-second segment. After having the set of selected features, there are a large number of machine learning algorithms suggested such as support vector machine[9], and K-Nearest Neighbors[10]. The other models like decision trees, logistic regression, and multilayer neural net-

works are applied for HAR based on the accelerometer of a smartphone[11]. In this paper, we use a convolutional neural network to extract interesting features. In the classification step, we compare the performance of softmax, random forest, and multi-class support vector machine. The paper is organized as follows. Section 2 describes the dataset, recall the technique of machine learning algorithms, present the proposed methods, and analyze their performance. Conclusion is given in Section 3.

2. Approaches for Activity Recognition with deep learning

In this section, we propose approaches combining deep learning and machine learning algorithms, such as support vector machine and random forest, to improve the accuracy of the HAR algorithm.

2.1. *Dataset*

The dataset we used in this paper is WISDM_raw[11], which covers 1.098.207 records from a tri-axial accelerometer. The dataset contains the measures corresponding to user activity, such as jogging, walking, upstairs, downstairs, sitting, and standing. The accelerometer data is collected every 50ms, which means 20 samples per second. Each reading contains an x, y, and z values corresponding to the three axes.

As mentioned, the raw time series data is needed to be transformed into examples. To accomplish this, we divided the data into 90-readings segments. Each segment is overlapped on the previous one, particularly, the next segment contains 45 readings of the previous segment. The features are then generated from the 90 readings contained within each 4.5-second segment. The number of transformed examples is 24403. The distribution of the activities in the raw dataset is presented in Figure 1.

2.2. *The machine learning algorithms*

As famously applied in image recognition[12], a convolutional neural network (CNN) is a feed-forward artificial neural network. A convolutional neural network consists of several feature extraction stages and a classifier. The classifier of CNN generally uses a back-propagation network[13] and a radial basis function network. In each feature extraction stage, a higher level of features is obtained through convolution operations. Each feature extraction stage includes a convolution layer and a sub-sampling layer. The sub-sampling layer uses the region filter to reduce the resolution

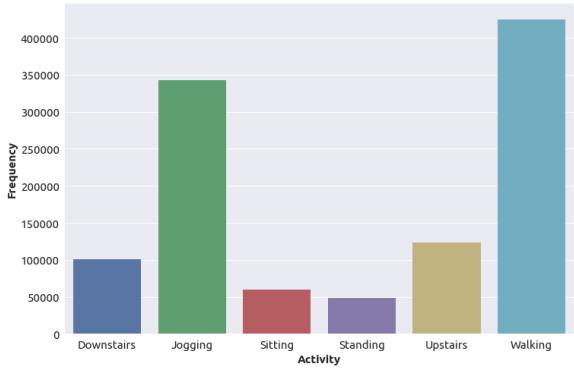

Fig. 1. The distribution of the activities.

of the feature map in the convolution layer. The sub-sampling operation can extract the important features for classification and ignore the useless details and noise. The mapping relationship between input and output is obtained through learning. This mapping relationship is reflected in the network weight, which helps CNN have the ability to extract features of input data, layer by layer.

The Support Vector Machine (SVM) is a supervised learning algorithm applied frequently in classification and regression analysis. In the context of binary classification, the outputs belong to $\{-1, 1\}$. The idea of the SVM is to find a hyperplane with the maximum margin which separates two classes. In the case of a non-linear SVM, the input data are mapped to higher dimensional space where the discriminating hyperplane can be linearly constructed. The SVM can also be extended to multi-class classification [14]. Particularly, the multi-class classifier is constructed by combining several binary classifiers. In practice, some algorithms were derived including the one-against-rest method, one-against-one method, DAG-SVM, ECOC-SVM, and SVM-BTA. We use the one-against-one method for classification procedure since it is practically suitable with similar performance.

Random Forest is an algorithm that learns from multiple decision trees driven on slightly different subsets of data. The procedure of the algorithm consists of three stages [15]. The first stage is to create n_{tree} bootstrap samples from the data. In particular, each sample (bag) contains N observations which are uniformly selected (with replacement) out of N original observations using bootstrap. Then for each sample, we grow a decision CART (Classification and Regression Tree). Instead of using all predictors,

at each node of each tree, m_{try} of the predictors are randomly selected, and the best split is chosen from those variables.

2.3. *Activity recognition task*

The features are extracted from each segment by CNN. With the way we divide the data, each input of the network has shape 3×90. After the input layer, one convolutional layer with 32-5×5 filters interleaves with a max-pooling layer with a kernel of size 2×2. A fully connected network follows, composed of two dense layers of 128 units each. To avoid overfitting, the dense layers are regularized with dropout with a 0.8 probability of keeping each neuron. The activation function used within the whole network is the ReLU function, and the loss is calculated with the cross-entropy function. The stochastic optimization method used in our paper is Adam optimizer. In our learning process, the batch size of 30 is used since this value speeds up the training time. The number of training epoch is 15. The output layer is composed of 6 units corresponding to 6 activities. The softmax function will classify the input to the most likely class.

The SVM and random forest are two other classifiers used for the classification tasks. In particular, the weights of the second fully connected layer are the features for the classification task of SVM and random forest algorithms. The hybrid classifiers have been used in recognizing handwritten to improve the performance[16,17]. The CNN approach is well known in extracting feature, especially in the domain of image processing. On the other hand, the SVM and random forest are two most common used in classification tasks, then the combination of CNN and these algorithms would well improve the performance. The Gaussian or radical basis function (RBF) kernel is used for SVM. The parameters of both SVM and random forest algorithms are optimized by cross-validation.

2.4. *Experiment and results*

The summary results for our activity experiments are presented in Table 1 and Figure 2. As presented in Table 1, the CNN-SVM approach outperforms the other methods in most cases; the CNN-softmax gives the best accuracy in classifying 'downstairs'. While in other cases, both CNN-SVM and CNN-RF give more accurate results than the neural network classification does. When we compare with the best result in[11], our method combining CNN and SVM outperforms the multilayer perceptron in all activities recognition. Therefore, the overall accuracy of CNN-SVM (with

92.5%) is higher than the overall accuracy of the multilayer perceptron (with 91.7%). Furthermore, the CNN-RF also shows the better results in most of the activities recognition, besides the slightly less accurate in classifying 'jogging', 97.59% of CNN-RF compared to 98.3% of multilayer perceptron. The overall accuracy of CNN-RF is 91.77%, which is higher than the multilayer perceptron with 91.7%.

Figure 2 demonstrates the confusion matrices (in percentage) of our three algorithms. In both Figure 2 and Table 1, it is observed that most of the misclassified records are 'upstairs' and 'downstairs' activities. The reason is that those two activities are similar and often confused with one another. In addition, from the confusion matrices, we observed that the 'upstairs' and 'downstairs' activities are also misclassified a lot with 'walking' activity, especially the CNN-RF.

Table 1. The accuracies of the activity recognition algorithms, the best results are bolded.

	% of records correctly predicted			
	CNN-softmax	CNN-SVM	CNN-RF	Multilayer Perceptron [11]
Downstairs	**78.01**	70.12	67.63	44.3
Jogging	92.50	**98.35**	97.59	98.3
Sitting	96.30	**99.07**	**99.07**	95.0
Standing	94.29	**98.10**	**98.10**	91.9
Upstairs	66.78	**74.48**	73.08	61.5
Walking	94.06	**97.45**	97.24	91.7
Overall	88.94	**92.50**	91.77	91.7

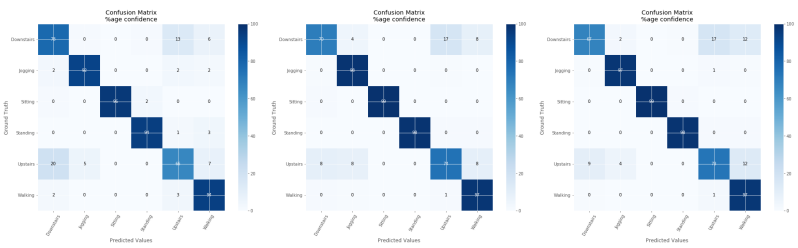

Fig. 2. The confusion matrix for the activity classification using CNN-softmax (left), CNN-SVM (center), and CNN-Random Forest (right).

3. Conclusions

In this paper, we presented three new approaches using deep learning for human activity recognition. The advanced deep learning convolutional neural network combined with the strong machine learning algorithms, support vector machine, and random forest, has improved the accuracy of activity detection. Numerical results have shown that the proposed approach, CNN-SVM, outperforms the previous methods in detecting all the activities as well as the overall accuracy. The CNN-RF also performs very well compare to the previous study and the neural network classification.

In the future study, we would like to apply these algorithms on the data with more than three attributes, and healthcare data. Moreover, these proposed methods can also be used for numerous classification problems. We are also interested in applying human activities recognition in health monitoring on patients and elders. On the other hand, we can develop other hybrid algorithms to increase the accuracy and explore a way to automatically determine the sizes and numbers of the convolution kernels in the network during the training process.

References

1. E. Cero Dinarević, J. Baraković Husić and S. Baraković, Step by step towards effective human activity recognition: A balance between energy consumption and latency in health and wellbeing applications, *Sensors* **19**, p. 5206 (2019).
2. O. D. Lara and M. A. Labrador, A survey on human activity recognition using wearable sensors, *IEEE communications surveys & tutorials* **15**, 1192 (2012).
3. M. Shoaib, S. Bosch, O. D. Incel, H. Scholten and P. J. Havinga, A survey of online activity recognition using mobile phones, *Sensors* **15**, 2059 (2015).
4. A. Bulling, U. Blanke and B. Schiele, A tutorial on human activity recognition using body-worn inertial sensors, *ACM Computing Surveys (CSUR)* **46**, 1 (2014).
5. A. Stisen, H. Blunck, S. Bhattacharya, T. S. Prentow, M. B. Kjærgaard, A. Dey, T. Sonne and M. M. Jensen, Smart devices are different: Assessing and mitigatingmobile sensing heterogeneities for activity recognition, in *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems*, 2015.
6. J. Wang, C.-C. Lin, Y.-S. Yu and T.-C. Yu, Wireless sensor-based

smart-clothing platform for ecg monitoring, *Computational and mathematical methods in medicine* **2015** (2015).

7. Y. Charlon, N. Fourty and E. Campo, A telemetry system embedded in clothes for indoor localization and elderly health monitoring, *Sensors* **13**, 11728 (2013).
8. A. Darwish and A. E. Hassanien, Wearable and implantable wireless sensor network solutions for healthcare monitoring, *Sensors* **11**, 5561 (2011).
9. D. Anguita, A. Ghio, L. Oneto, X. Parra and J. L. Reyes-Ortiz, Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine, in *International workshop on ambient assisted living*, 2012.
10. M. A. Ayu, S. A. Ismail, A. F. A. Matin and T. Mantoro, A comparison study of classifier algorithms for mobile-phone's accelerometer based activity recognition, *Procedia Engineering* **41**, 224 (2012).
11. J. R. Kwapisz, G. M. Weiss and S. A. Moore, Activity recognition using cell phone accelerometers, *ACM SigKDD Explorations Newsletter* **12**, 74 (2011).
12. A. Krizhevsky, I. Sutskever and G. E. Hinton, Imagenet classification with deep convolutional neural networks, in *Advances in neural information processing systems*, 2012.
13. Y. LeCun, B. E. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. E. Hubbard and L. D. Jackel, Handwritten digit recognition with a back-propagation network, in *Advances in neural information processing systems*, 1990.
14. H. Qian, Y. Mao, W. Xiang and Z. Wang, Recognition of human activities using svm multi-class classifier, *Pattern Recognition Letters* **31**, 100 (2010).
15. A. Liaw, M. Wiener *et al.*, Classification and regression by randomforest.
16. M. Elleuch, R. Maalej and M. Kherallah, A new design based-svm of the cnn classifier architecture with dropout for offline arabic handwritten recognition, *Procedia Computer Science* **80**, 1712 (2016).
17. X.-X. Niu and C. Y. Suen, A novel hybrid cnn–svm classifier for recognizing handwritten digits, *Pattern Recognition* **45**, 1318 (2012).