A Survey on Hand-based Behavioral Activity Recognition

Conference Paper · December 2019

DOI: 10.1109/ICCWAMTIP47768.2019.9067524

Happy Nkanta Monday
Oxford Brookes College of Chengdu University of Technology

Grace U. Nneji
Oxford Brookes College of Chengdu University of Technology

Mordecai Raji
University of Electronic Science and Technology of China

Chiagoziem Chima Ukwuoma
University of Electronic Science and Technology of China

A SURVEY ON HAND-BASED BEHAVIORAL ACTIVITY RECOGNITION

HAPPY NKANTA MONDAY[1], *JIAN PING LI[1], GRACE UGOCHI NNEJI[2], MORDECAI FOLARIN RAJI[1], CHIMA C. UKWUOMA[2], JALALUDDIN KHAN[1], CHUKWUEBUKA J. EJIYI[2], AMIN ULHAQ[1], SAIFUN NAHAR[2], HASSAN S. ABUBAKAR[3], CHIKWENDU A. IJEOMA[3]

[1]School of Computer Science and engineering, University of Electronic Science and Technology of China, Chengdu 611731-China
[2]School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu
[3]School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu
E-MAIL: mh.nkanta@gmail.com, *jpli2222@uestc.edu.cn, ugochinneji@gmail.com, mraji@qq.com, jalal4amu@yahoo.com, cjejiyi@hotmail.com, khan.amin50@yahoo.com, sainahbd@163.com, hassansaniabubakar@gmail.com, amauchefelix@gmail.com

Abstract:

Wrist-worn devices such as smart-watches can be a perfect way to collect personalized user data, since these devices are embedded with various sensors such as the accelerometer and gyroscope which can sense user's hand movement, as well as help in identifying the activity performed by the user. Additionally, smart-watches are light and unobtrusive unlike the wristbands used in past studies which require the user to be connected by several wires and bulky sensors. In this study, we examined the different neural network adopted for the recognition of hand-based behavioral activity to recognize six complex hand motion activities such as smoking, drinking, eating, typing giving a talk and writing with a pen in a non-controlled environment, using accelerometer and gyroscope sensor signals acquired from smart-watch. Unlike other machine learning methods which use hand crafted features, CNN can learn features automatically from the data as well as capture local dependency.

Keywords:

Deep learning; Convolutional neural network; Hand motion; Behavioral activity

1. Introduction

The past decade has experienced great advances in sensor technology and wireless communication networks in terms of capacity increase, cost efficiency and power efficiency. For instance, computing devices such as mobile phones have had battery capacity increase from around 950mAh in around 2008 to over 1500 mAh now [1]. However, human activities have a complex hierarchical structure where an activity can be broken down into several simpler activities. Also a single activity can be performed in different styles. This complex and diverse nature of the activates makes it challenging to model as well as accurately recognition these activities [2-3]. fig.1 shows a general flow of activities involved in an activity recognition process. In the training phase, features are extracted from the raw time series data and then used to build and train a classification model. Features can be done using traditional statistical methods or using deep learning techniques. In the classification phase, key features are extracted from previously unseen raw data, and then the trained model is used to predict an activity class label.

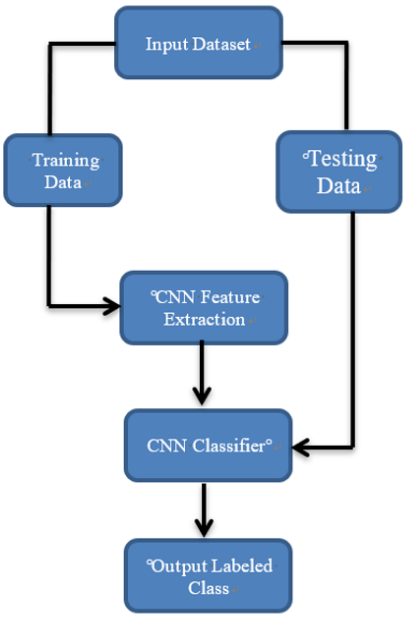

Fig.1 Activity recognition process

2. Background

Activity recognition from body worn sensors has been on the rise. Originally, activity recognition relied on the use of specially engineered devices which were distributed in the various parts of the subject's body[4-5]. However, in the recent years, commercial smartphones have been utilized to identify physical human activities[6-12]. Smartphone devices contain the sensors such as the accelerometer and the gyroscope which are mainly required for activity recognition. The use of smartphone brought about the increased success in the human activity recognition field and expanded its application. Nevertheless, it introduced limitations in terms of their placement as well as the orientation [13]. For example, [14] presents an activity recognition system to recognize nine different gestures usually performed in daily activities by means of ring and bracelets IMUs placed on the fingers and on the wrist, respectively. Two commonly used supervised machine learning techniques: Decision Tree and Support Vector Machine were used for classification. [15] Investigate the problem of activity recognition using accelerometer data generated from wrist-worn devices. Work focus on the recognition of ambulation activities like Walking, Standing, Sitting and Lying, which will enable various human performance evaluation applications. They used Naive Bayes, Support Vector Machines, Decision Tree, k-Nearest Neighbors, Multilayer Perceptron, and Random Forest classifiers for their study where sensor data from the accelerometer was fed to individual classifiers as well as their combinations for training and validation. Results show that data generated by wrist-worn accelerometer sensor are insufficient for ambulation activity recognition and can be used for human performance evaluation applications.

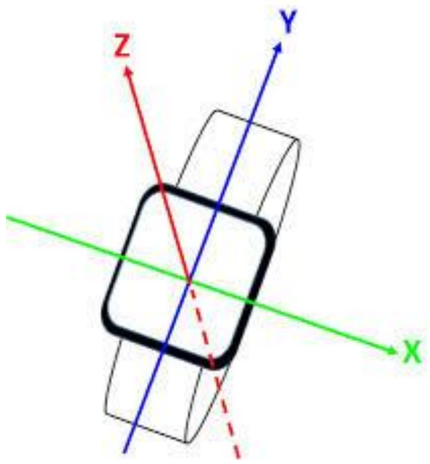

Fig.2 The coordinate system of a smart-watch

3. State of the art

Deep Learning is a new area of Machine Learning research, inspired by information processing and communication patterns in biological nervous systems. As it developed over the past several decades, deep learning has heavily depended on our understanding of human brain, statistics and applied mathematics [16-18]. Tremendous growth in terms of popularity and usefulness of deep learning approaches has been witnessed in recent years brought about by increasing computer power, ever increasing amounts of data as well as techniques for training deeper networks. Deep-learning methods are considered as some kind of representation-learning methods consisting several level of representation made up of simple but non-linear processing units and transformation [19-21]. Additionally, the predictive accuracy increases with more data availed [22]. Applying deep learning to sensor data has a number of rewards. To start with, deep learning enables us to automatically extract meaningful features from high-dimensional data thus eliminating the need for hand crafted or heuristic-based features [23]. Secondly, with the scarcity in labeled activity data, deep learning models can utilize unlabeled activity samples for model fitting in an unsupervised pre-training phase [24]. Finally, deep learning models are robust to over fitting problems [25] During classification tasks, features of the input essential for discrimination are amplified as you move to higher layers of representation while the irrelevant variations are suppressed.

4. Deep learning and hand-based behavioral activity

Recognition of human activities is still immature due to the poor recognition accuracy of existing recognition methods as well as scarcity of labeled training data [26]. In the past decade, deep learning has gained a lot of attention in various fields especially with its successful application in areas of image and speech recognition. Deep learning techniques such as CNNs, which is a powerful feature extractor and classifier, have set the latest state-of-the-art performance in the area of image and speech [28]. However, it is not only image and speech recognition that can benefit from such a powerful deep learning tool; Human Activity Recognition (HAR) is one field that is also a good match especially when it comes to handling hierarchical structure of activities, translation invariance and temporally correlated readings of time-series signals, as well as HAR feature extraction problems [27]. A deep convolutional neural network approach that uses data from the smartphone sensors to perform efficient and effective human activity recognition was proposed in [28]. The proposed approach exploits the inherent characteristics of activities and 1D time-series signals, at the same time providing a way to automatically and data-adaptively extract robust features from raw data. From their study, they concluded that CNN outperforms other state of the art techniques in HAR in terms of deriving relevant and more complex features as well as classification on moving activities, especially those that are very similar to one another, which were previously alleged to be very difficult to classify. Deep Learning (DL) techniques have been shown to offer excellent solutions since they allow an automatic extraction of features without any domain knowledge [29]. It is evident that Convolutional neural network (CNN) as a type of deep learning technique, through its alternating convolution and pooling operations has the ability to exploit these signal and activity characteristics as well as offer us hierarchical extraction of relevant features as presented in fig.3.

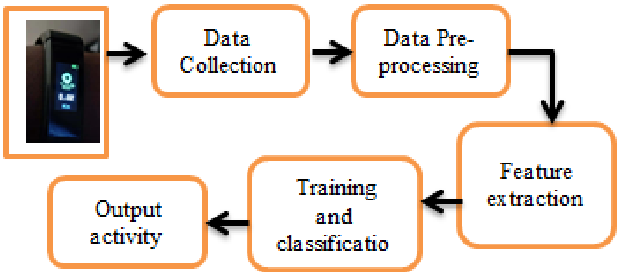

Fig.3 CNN architecture for hand motion classification

Table 2 Different methods of activity recognition report

Reference	Smartphone position	Activity numbers	Contributions	Algorithms and accuracy	Limitation
Anjum et al. [16] 2013	Pant pocket, hand, handbag, shirt pocket	7	Activity recognition with smartphone at multiple positions including pant pocket, hand, handbag and shirt pocket	Decision tree (98.5%)	Limited activity traced and thus would tradeoff the performance in external verification
Aril [17] 2014	Leg front pant pocket	6	Demonstrated of better activity classification accuracy	10-fold KNN (98.2%)	Position is fixed in front pants leg pockets
Roman et al. [18] 2013	Leg front pant pocket	9	Estimation of total energy expenditure with phone position independence by transform	Total energy expenditure (73.6%)	Low accuracy
Yongjin et al. [19] 2014	Pants pocket	5	Unsupervised learning without labels	Hierarchical clustering or DBSCAN (90. 3%)	Some important activities including going upstairs and downstairs were not studied.
Sourav et al. [20] 2014	Jacket pocket, pant pocket, backpack	8	Deal with unlabeled data	Sparse coding (80%)	Some important activities including going upstairs and downstairs were not studied.

Miao et al. [2] 2015	Any pocket	5	Automatically identifying the location of the smartphone and conveniently activity recognition with smartphone at any pocket	10-fold J48 (89.6%)	More situations as in the hand, should be further studied

5. Conclusion

In this paper, we discussed the basic ideas behind deep learning and specifically the Convolution Neural Networks. We also presented some of the past studies that used deep learning techniques to solve hand motion recognition problems. From our literature survey, we can conclude that CNNs have the potential of taking hand motion recognition research to a whole new level. Unlike other neural network techniques that require the user to do hand crafted feature extraction, CNNs can do both feature extraction as well as prediction/classification in one. Future work can consider deep transfer learning method for improved accuracy.

Acknowledgment

I want This paper was supported by the National Natural Science Foundation of China (Grant No. 61370073), the National High Technology Research and Development Program of China (Grant No. 2007AA01Z423), the project of Science and Technology Department of Sichuan Province, Chengdu Civil-military Integration Project Management Co., Ltd., and Sichuan Yin Ten Gu Technology Co., Ltd.

References

[1] Miao, Fen & He, Yi & Liu, Jinlei & Li, Ye & Ayoola, Idowu. (2015). Identifying typical physical activity on smartphone with varying positions and orientations. Biomedical engineering online. 14. 32. 10.1186/s12938-015-0026-4.

[2] M. Zeng *et al.*, "Convolutional Neural Networks for Human Activity Recognition using Mobile Sensors," *Proc. 6th Int. Conf. Mob. Comput. Appl. Serv.*, no. November, 2014.

[3] B. D. *et al.*, "Recognizing upper limb movements with wrist worn inertial sensors using k-means clustering classification," *Hum. Mov. Sci.*, vol. 40, pp. 59–76, 2015.

[4] J. F. Alsamora, "Hand gestures recognition using 3D-CNN A Degree Thesis Submitted to the Faculty of the Escola Tècnica d ' Enginyeria de Telecomunicació de Barcelona Universitat Politècnica de Catalunya by Josep Famadas Alsamora In partial fulfilment of the requirements f," no. June, 2017.

[5] C. A. Ronao and S. B. Cho, "Human activity recognition with smartphone sensors using deep learning neural networks," *Expert Syst. Appl.*, vol. 59, pp. 235–244, 2016.

[6] A. Bandar, A. Jalal, and A. Abdelmonim, "An Effective Deep Autoencoder Approach for Online Smartphone-Based Human Activity Recognition," *Int. J. Comput. Sci. Netw. Secur.*, vol. 17, no. 4, pp. 160–165, 2017.

[7] A. Moschetti, L. Fiorini, D. Esposito, P. Dario, and F. Cavallo, "Recognition of daily gestures with wearable inertial rings and bracelets," *Sensors (Switzerland)*, vol. 16, no. 8, 2016.

[8] M. Nguyen, L. Fan, and C. Shahabi, "Activity Recognition Using Wrist-Worn Sensors for Human Performance Evaluation," *Proc. - 15th IEEE Int. Conf. Data Min. Work. ICDMW 2015*, pp. 164–169, 2016.

[9] H. Wen, J. Ramos Rojas, and A. K. Dey, "Serendipity: Finger Gesture Recognition using an Off-the-Shelf Smartwatch," *Proc. 2016 CHI Conf. Hum. Factors Comput. Syst. - CHI '16*, pp. 3847–3851, 2016.

[10] F. Alharbi and K. Farrahi, "A Convolutional Neural Network for Smoking Activity Recognition," *2018 IEEE 20th Int. Conf. e-Health Networking, Appl. Serv.*, pp. 1–6, 2018.

[11] M. Shoaib, H. Scholten, P. J. M. Havinga, and O. D. Incel, "A hierarchical lazy smoking detection algorithm using smartwatch sensors," *2016 IEEE 18th Int. Conf. e-Health Networking, Appl. Serv. Heal. 2016*, 2016.

[12] G. M. Weiss, J. L. Timko, C. M. Gallagher, K. Yoneda, and A. J. Schreiber, "Smartwatch-based activity recognition: A machine learning approach," *3rd IEEE EMBS Int. Conf. Biomed. Heal. Informatics, BHI 2016*, pp. 426–429, 2016.

[13] M. Shoaib, S. Bosch, O. D. Incel, H. Scholten, and P. J. M. Havinga, "Complex human activity recognition using smartphone and wrist-worn motion sensors," *Sensors (Switzerland)*, vol. 16, no. 4, pp. 1–24, 2016.

[14] First paper (M. Shoaib, S. Bosch, O. D. Incel, H. Scholten, and P. J. M. Havinga, "Complex human

activity recognition using smartphone and wrist-worn motion sensors," Sensors (Switzerland), vol. 16, no. 4, pp. 1–24, 2016.) v.s. our proposed CNN

[15] K. Charalampous and A. Gasteratos, "On-line deep learning method for action recognition," *Pattern Anal. Appl.*, vol. 19, no. 2, pp. 337–354, 2016.
[16] Anjum A, Muhammad Usman I. Activity recognition using smartphone sensors. IEEE conference on ConsumerCommunications and Networking. 2013. p. 914–9.16.
[17] Arif M, Bilal M, Kattan A, Ahamed SI. Better physical activity classification using Smartphone acceleration sensor. JMed Syst. 2014;38:1–10.17.
[18] Guidoux R, Duclos M, Fleury G, Lacomme P, Lamaudièr N, Manenq P, et al. A smartphone-driven methodology forestimating physical activities and energy expenditure in free living conditions. J Biomed Inform. 2014;52:271–8.18
[19] Yongjin K, Kyuchang K, Changseok B. Unsupervised learning for human activity recognition using smartphonesensors. Expert Systems with Applications. 2014;41:6067–74.20.
[20] Bhattacharya S, Nurmi P, Hammerl N, Plötz T. Using unlabeled data in a sparse-coding framework for humanactivity recognition. Pervasive Mob Comput. 2014;15:242–62
[21] D. Wu *et al.*, "Deep Dynamic Neural Networks for Multimodal Gesture Segmentation and Recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 8, pp. 1583–1597, 2016.
[22] S. Shin and W. Sung, "Dynamic Hand Gesture Recognition for Wearable Devices with Low Complexity Recurrent Neural Networks," pp. 2274–2277, 2016.
[23] E. Nasr-Esfahani *et al.*, "Hand Gesture Recognition for Contactless Device Control in Operating Rooms," *Https://Arxiv.Org/Abs/1611.04138*, 2016.
[24] P. Molchanov, X. Yang, S. Gupta, K. Kim, S. Tyree, and J. Kautz, "Online Detection and Classification of Dynamic Hand Gestures with Recurrent 3D Convolutional Neural Networks," *2016 IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 4207–4215, 2016.
[25] D. Moazen, S. A. Sajjadi, and A. Nahapetian, "AirDraw: Leveraging smart watch motion sensors for mobile human computer interactions," *2016 13th IEEE Annu. Consum. Commun. Netw. Conf. CCNC 2016*, pp. 442–446, 2016.
[26] X. Ren, Y. G. Soo, M. Odagaki, and F. Duan, "sEMG-Based Hand Motion Recognition System Using RMSR and AR Model," pp. 5410–5415, 2017.
[27] S. Cho, M. Yoshikawa, M. Ding, J. Takamatsu, and T. Ogasawara, "Hand Motion Recognition Using a Distance Sensor Array," pp. 1459–1464, 2017.
[28] S. G. Cho, M. Yoshikawa, K. Baba, K. Ogawa, J. Takamatsu, and T. Ogasawara, "Hand motion recognition based on forearm deformation measured with a distance sensor array," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 2016-Octob, pp. 4955–4958, 2016.
[29] Y. Zhu, S. Jiang, and P. B. Shull, "Wrist-worn hand gesture recognition based on barometric pressure sensing," *2018 IEEE 15th Int. Conf. Wearable Implant. Body Sens. Networks, BSN 2018*, vol. 2018-Janua, no. March, pp. 11–184, 2018.