Poster: Human Activity Recognition Using Earable Device

Tahera Hossain[†]
Applied Science for Integrated System Eng.
Kyushu Institute of Technology, Japan
taheramoni@gmail.com

Md Shafiqul Islam
Dept. of Human Intelligence Systems
Kyushu Institute of Technology, Japan
shafiqul@sozolab.jp

Md Atiqur Rahman Ahad
University of Dhaka, Bangladesh
Osaka University, Japan
atiqahad@du.ac.bd

Sozo Inoue
Dept. of Human Intelligence Systems
Kyushu Institute of Technology, Japan
sozo@sozolab.jp

ABSTRACT

Wearable sensors are monumental for human activity recognition. Researchers are continuously inventing new technology to detect human activity properly. Earable opens up interesting possibilities of monitoring personal scale behavioral activities. In this paper, we explore earables device 'eSense' multisensory stereo device for personal scale behavior analysis. We propose an activity recognition framework by exploiting eSense based multi-sensory device. It has a microphone, 6-axis inertial measurement unit, and a dual-mode Bluetooth. We use eSense accelerometer sensor data for detecting head and mouth related behavioral activities. We develop a data collection framework from the eSense through our smartphone application via Bluetooth. Then from the collected data, a few statistical features are computed to classify six personal scale activities related to head and neck movement such as speaking, eating, headshaking and head nodding, as well as, stay and walk. We aggregate the time series data into different action labels that summarize the user activity over a time interval. After, we train the data to induce a predictive model for activity recognition. We explore both machine learning and deep learning approach for data classification. For classification, we use the Support Vector Machine, Random Forest, and K-Nearest Neighbor and Convolutional Neural Network and achieve satisfactory recognition accuracy. The findings provide promising prospect for eSense for personal scale activity recognition in healthcare monitoring service. Based on our study, this kid of work is done for the first time with satisfactory findings.

CCS CONCEPTS

• **Human-centered computing → Ubiquitous and mobile computing.**

KEYWORDS

Wearable, Activity recognition, Earables, eSense, Healthcare

UbiComp/ISWC '19 Adjunct, September 9–13, 2019, London, United Kingdom

ACM ISBN 978-1-4503-6869-8/19/09.
https://doi.org/10.1145/3341162.3343822

ACM Reference Format:
Tahera Hossain, Md Shafiqul Islam, Md Atiqur Rahman Ahad, and Sozo Inoue. 2019. Poster: Human Activity Recognition using Earables Device. *In Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2019 International Symposium on Wearable Computers (UbiComp/ISWC '19 Adjunct), September 9--13, 2019, London, United Kingdom.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3341162.3343822

1 INTRODUCTION

Wearables opens interesting possibilities of monitoring different type of human actions [1]. To support elderly population in the world, it is important to know elderly people's regular activity as well as personal behavioral activity to monitor their life quality. Through the efficient technology, it is possible to support elderly people's life in nursing care center, assisted living center or healthcare monitoring center. For behavior analysis, low-power and low-cost wearable sensors are used in most of the cases, which are mounted on the human body [13]. Sensor-based activity recognition systems are explored for over a decade by many researchers [4]. In this system, data are collected from various wearable sensors (e.g., accelerometer, gyroscope) from each individual, then these are transferred to near-by access points, then to servers. But Data annotation is a big challenge in this field. Along with this side, a controlled environment data collection and wild is different [7]. Also, through wearables sensors it is bit difficult to detect personal activities like speaking, eating and other head and mouth related activities. Earable opens the possibility to detect this kind of critical activity which are important to detect for nurse care and elderly people. Because they need to record and manage the amount of intake, and swallowing ability is also one of the measures of health for elderly people. By taking this concern as an important issue, in this paper, we focus on the analysis of the signals obtained via ear-mounted wearable sensor called eSense earables, which can be explored in personal scale behavior analysis [8]. These devices are provided by Nokia-Bell Lab, UK. eSense is a high-definition stereo earable equipped with a microphone, a 6-axis inertial measurement unit, and a dual-mode Bluetooth. Through eSense earables, it is possible to record real-time data streams of three modalities – audio, motion, and proximity (BLE RSSI). From eSense 6-axis inertial measurement unit (IMU), we can get three-axis accelerometer and three-axis gyroscope data. eSense can be effectively used to monitor head- and mouth-related behavioral activities, which we focus in this paper.

On the other hand, researchers have exploited different machine learning methods including Support Vector Machines (SVM), Random Forest (RnF), and K-Nearest Neighbor, etc. for sensor data exploration. In machine learning approach, it is essential to extract significant features properly. This feature selection has great impact on accuracy measurement [13]. On the other hand, Convolutional Neural Network (CNN) has the ability to learn complex patterns without the need for prior feature extraction by an expert [6]. Motivated by the achievement of CNN, we explored both CNN and other classical machine learning approaches for classify head- and mouth-related human activities, e.g., Speaking, Eating, Head Shaking and Head Nodding along with stay and walk class. We propose a framework for some activity recognition after collecting data from eSense.

The objectives of this paper are as follows: To explore eSense multi-sensory earables for personal behavioral data collection through Bluetooth and smartphone application; to use eSense sensors for human activity recognition; and to identify supervised machine learning and deep neural network classifier to achieve good recognition performance.

The paper is organized as follows: after introducing the works in Section 1, we present a related work in short in Section 2. Then we introduce the proposed framework with a system architecture. We explain the experimental dataset and results after that. Finally, we conclude the paper with some future work guidelines.

2 RELATED WORK

In the wearable device, the presence of sensors specially accelerometer and gyroscope sensor make it possible to collect human movement data. The signals from both sensors are commonly used for Human Activity Recognition (HAR) [10]. Some of the important time and frequency domain features can be extracted from accelerometer measurements for the purpose of action recognition [2]. Preece et al. [11] made a comparison among some feature extraction methods for the classification of dynamic activities from accelerometer data. Three types of features, time domain, frequency domain, and time-frequency domain, are extracted from the data.

Kwapisz et al. [9] studied accelerometer data from smartphone. They collected data from 29 subjects. Though they achieved a good accuracy in recognizing jogging action, they hardly succeeded to recognize upstairs and downstairs actions. Sztyler et al. [12] proposed a method for HAR whereby the position of the wearable devices on the human body can change depending on user preference. Multiple-sensors were used instead of single-sensor to recognize actions from accelerometer data correctly by Gao et al. [3]. Authors in [5] introduced LoRaWAN technology for human activity recognition. It is promising for healthcare monitoring service as LoRa has capability to cover multiple kilometer through one gateway. Researcher are introducing new technology for better human activity recognition through any suitable platform. There are many scope in this area to explore a stable new platform for healthcare domain.

Activity recognition from wearable sensors have many realistic challenges to explore. Therefore, it is required to explore different kinds of sensors, locations and methods so that we can extract activities in a better manner and through different directions. So, it is required to introduce new system and framework for better human activity recognition. It is challenging to monitor personal scale human activity through wearable devices. On this aspect, we try to introduce a new system eSense multi-sensory stereo device for recognize head and mouth related behavioral activities which can be impactful for healthcare monitoring center.

3 PROPOSED FRAMEWORK AND SYSTEM

The eSense earable device contains accelerometer and gyroscope sensors. These sensors data can be used to track a set of head- and mouth-related behavioral activities including speaking, eating, drinking, shaking, and nodding. To get this data, the eSense device needs to be connected to a smart phone via Bluetooth. When the device is connected, it acts like a peripheral device and the mobile acts as a central or host device. The sensor data of eSense device cannot be directly received. So, an Android application is needed, which will continuously maintain connection with the eSense device. We have developed an Android application to access the sensor data from the device. In this section, we introduce the system architecture, Android application and our proposed framework.

3.1 System Architecture

When a Bluetooth Low Energy (BLE) device gets connected with a central device (i.e., mobile, tablet), it uses Generic Attribute Profile (GATT). This is a general specification for sending and receiving short pieces of data known as ‘attributes’ over a BLE link.

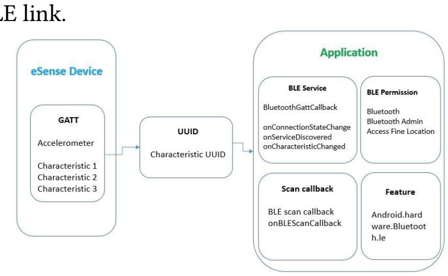

Figure 1: System Architecture

In Fig. 1, a general system architecture of BLE device and central device is stated. Each sensor inside the eSense device sends some characteristic data, which contains a single value and 0-n descriptors that describe the characteristics’ value. Each GATT profile contains a service and each service has some characteristics. Each of these attributes is uniquely identified by a Universally Unique Identifier (UUID), which is a standardized 128-bit format for a string ID.

When the mobile device tries to connect eSense device, first it scans the device. There are some callback methods which are called based on different states. When the connection between mobile and eSense is established, the onConnectionStateChanged method is called. While sending data from BLE device to central device, the onCharacteristicChanged method gets called. The whole process runs in a service. Using the callback methods, all types of communication and data sending/receiving are done.

3.2 Description of the Android Application

We have developed an Android application that maintains connection and receives sensor data from eSense device. It has several features and list of activities (Fig. 2). When Connect button is clicked, it scans for eSense device and connects it.

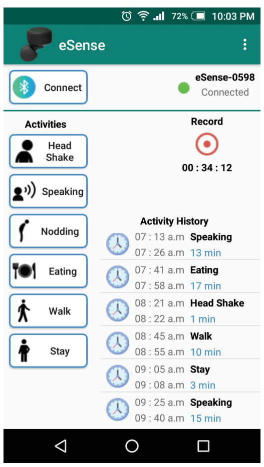

Figure 2: Android application for eSense device

In Fig. 2, a snap of the Android application screen is shown. It contains several features to connect and receive sensor data from eSense device. If a user wants to receive the sensor data from eSense device, the user will have to click the particular activity and record button. This record button works for start and stop recording of a particular activity. Then after performing the activity while wearing the device in an ear, data is saved when user clicks record button second time (which means stop). It depicts a list of performed activities.

3.3 Framework

As shown in Fig. 3, the application saves data in mobile storage as '.csv' format. To make the application more robust and easier to use, we propose a new framework. In our framework, after establishing connection, the sensor data from eSense device are saved in application database temporarily. If the Wi-Fi network becomes available, then the data is uploaded to a cloud server. After collecting big amount of sensor data, several statistical features and machine learning algorithm are applied to train and test the data.

Figure 3: Proposed framework

4 EXPERIMENTAL RESULT AND ANALYSIS

In this section, we accomplish experiments for six activities: target activities (namely, speaking, eating, headshaking and head nodding); and two regular activities (i.e., stay and walk) from eSense earables accelerometer data. We analyze this data by using several machine learning algorithms and Convolutional Neural Network (CNN) for activity prediction. The duration for performing each activity was around 3 minutes. There are around 34708 records for these 4 targeted activities, and 50467 records for all 6 activities. Data was collected at a constant rate of 50Hz from 3 axial accelerometer. For separating training and test data, we collected different person data for training and testing. We separated training and test data by different user with different trials. In the data collection apps, there are option to once connect the earables with eSense apps for data collection and stop the connection while we finished data collection. While we stopped and start again the data collection, then it starts with another new session for data collection for particular activity class. This way first we collected training data for specific user and separated those data for training purpose. Afterwards, we collected data again for different activity for testing purpose. This way, we used different trials for different classes for different person.

For feature extraction, we have calculated the magnitude value of raw acceleration data. Time domain statistical features such as mean and standard deviation for 3-axes of the accelerometer have been extracted. By using a windowing approach, we calculate the features for 80% overlapping data. We used several traditional machine learning algorithms, i.e., Random Forest (RnF), Support Vector Machine (SVM) with radial basis function kernel, K-Nearest Neighbor (KNN), as well as deep learning model 1D CNN. Among three machine learning classifiers, KNN achieved best accuracy which is 90.95%.

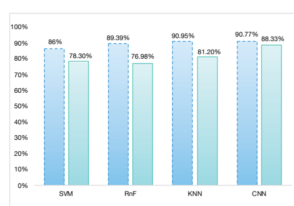

Figure 4: Accuracies comparison obtained from the classifiers for 4 activities and 6 activities respectively.

Afterwards, we computed 1D CNN and among the all classifiers, it provides best accuracy (90.77%). We used 2 convolutional layers, each of which contains 64 filter maps and kernel size is 1x3. We have used 1 max-pooling layer after convolutional layer and its pool size is 1x2. After pooling layer, we have used 2 fully-connected layers. The last layer is softmax layer. We have used Adam optimizer and ReLu activation function. We did not manually extract features from accelerometer sensor data since CNN automatically extract features. When 4 activities are considered, maximum misclassification is found between 'Head nodding' and 'Speaking' activities.

In Fig. 4, we demonstrated the comparison of accuracy for 4 and 6 activity classes by exploring several classifiers and CNN. We have found the accuracy of 78.3% by SVM, 76.98% by RnF, 81.20% by KNN, ad 88.33% by CNN for 6 activity classes. While CNN performs slightly better than KNN for 4 activities, the later has significantly higher accuracy when only head and mouth related four classes are considered for classification.

As observed in Fig. 4, when 'Stay' and 'Walk' classes are added with the previous four head- and mouth-related classes, the recognition accuracy falls. We notice that 'Eating' class is mostly confused with 'Stay' class and 'Speaking' class confused with 'Eating' (Fig. 5). As 'Eating' action is done while in 'Stay' or less movement situation, so they are confused. Also 'Eating' and 'Speaking' has similar pattern for mouth movement. Moreover, in this study, 'Speaking' activity is done while in a sitting (no movement status). This is the reason for main confusion. Sometimes, 'Nodding' has more similarity with 'Speaking' status, as these are taken while sitting on a chair (Fig. 6).

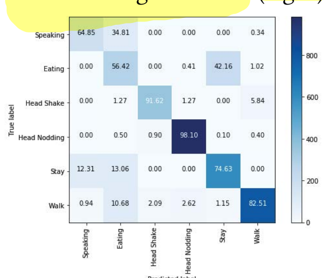

Figure 5: Confusion matrix with other regular activities by KNN

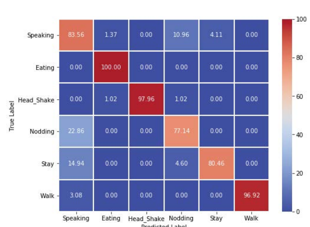

Figure 6: Confusion matrix with other regular activities by CNN

However, recognition results only for head- and mouth-related activity classes are better without 'Stay' and 'Walk' classes.

5 CONCLUSION AND FUTURE WORK

For recognizing activities of daily life, wearable devices or mobile phones are widely used. There are certain types of behavioral activities related to head and mouth, i.e., drinking, eating, nodding, shaking, etc., which are hard to recognize using sensors in mobile devices. Use of earable device like eSense is a promising new technology to detect these head- and mouth-related activities. It has huge potential to create immense impact on healthcare specially for elderly people. In this paper, we have exploited the accelerometer data for detecting different types of activity classes. By using some statistical features and machine learning approaches, we recognized them and KNN performed better. We also implemented CNN model and achieved better results. However, there are misclassifications related to moving and less-moving related activities. Therefore, by applying our proposed framework, it will be easier and efficient to detect head and mouth related activities with good accuracies. Based on our knowledge, there is no work on earables and activity recognition. We hope that it will enhance a new and important dimensions of research and applications.

This experiment was done in lab environment. In future, we will explore this experiment in nursing care center to collect real field data. With eSense earables, if we only use accelerometer sensor data then it is hard to detect normal activity class along with head and mouth movements. In future, we will also explore the potentials of gyroscope and audio data along with accelerometer data. We will also enhance activity classes with more complicated activities and similar types of activities which have subtle differences.

ACKNOWLEDGMENTS

We acknowledge Prof. Fahim Kawsar and Pervasive Research Unit Nokia Bell Labs, Cambridge for providing the devices.

REFERENCES

[1] A. D. Antar, M. Ahmed, and M. A. R. Ahad, "Challenges in Sensor-based Human Activity Recognition and a Comparative Analysis of Benchmark Datasets: A Review," Intl. Conf. on Activity and Behavior Computing (ABC), 2019.

[2] W. Dargie, "Analysis of Time and Frequency Domain Features of Accelerometer Measurements," 18th Intl. Conf. on Computer Communications and Networks, 2009.

[3] L. Gao, A. Bourke, J. Nelson, "Evaluation of Accelerometer Based Multi-sensor versus Single-sensor Activity Recognition Systems," Medical Engineering and Physics, 2014.

[4] Z. Hajihashemi and M. Popescu, "Detection of Abnormal Sensor Patterns in Eldercare", 4th IEEE Int. Conf. on E-Health and Bioengineering, 2013.

[5] T. Hossain, T. Tazin, A. Ahad, S. Inoue, "Activity Recognition by Using LoRaWAN Sensor", ACM Int'l Conf. Pervasive and Ubiquitous Computing (Ubicomp), 2018.

[6] A. Ignatov, "Real-time human activity recognition from accelerometer data using Convolutional Neural Networks", Applied Soft Computing, 2018.

[7] S. Inoue, T. Mairittha, N. Fah, T. Hossain, "Integrating Activity Recognition and Nursing Care Records: the System, Experiment, and the Dataset", International Conference on Activity and Behavior Computing, 2019.

[8] F. Kawsar, C. Min, A. Mathur, and A. Montanari, "Earables for Personal-Scale Behavior Analytics", IEEE Pervasive Computing, 17, 2018.

[9] J. R. Kwapisz, G. M. Weiss, S. A. Moore, "Activity Recognition using Cell Phone Accelerometers," SIGKDD Explorations, 2010.

[10] G. Ogbuabor and R. La, "Human Activity Recognition for Health- care Using Smartphones". Intl. Conf. on Machine Learning and Computing, 2018.

[11] S. J. Preece, J. Y. Goulermas, P. J. Kenney, D. Howard, "A Comparison of Feature Extraction Methods for the Classification of Dynamic Activities from Accelerometer Data," IEEE Trans. on Biomedical Engineering, 2009.

[12] T. Sztyler, H. Stuckenschmidt, and W. Petrich, "Position-aware activity recognition with wearable devices," Pervasive and Mobile Computing, 38(2), 2017.

[13] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep learning for sensor-based activity recognition: A Survey", Pattern Recognition Letters, 2018.