Exploring Human Activities Using eSense Earable Device

Md Shafiqul Islam, Tahera Hossain, Md Atiqur Rahman Ahad and Sozo Inoue

Abstract Detecting head- and mouth-related human activities of elderly people are very important for nurse care centers. They need to track different types of activities of elderly people like swallowing, eating, etc. to measure the health status of elderly people. In this regard, earable devices open up interesting possibilities for monitoring personal-scale behavioral activities. Here, we introduce activity recognition based on an earable device called 'eSense'. It has multiple sensors that can be used for human activity recognition. 'eSense' has a 6-axis inertial measurement unit with a microphone and Bluetooth. In this paper, we propose an activity recognition framework using eSense device. We collect accelerometer and gyroscope sensor data from eSense device to detect head- and mouth-related activities along with other normal human activities. We evaluated the classification performance of the classifier using both accelerometer and gyroscope data. For this work, we develop a smartphone application for data collection from the eSense. Several statistical features are exploited to recognize head- and mouth-related activities (e.g., head nodding, headshaking, eating and speaking), and regular activities (e.g., stay, walk and speaking while walking). We explored different types of machine learning approaches like Convolutional Neural Network (CNN), Random Forest (RnF), K-Nearest Neighbor (KNN), Linear

M.S. Islam
Dept. of Human Intelligence Systems, Kyushu Institute of Technology, Japan
e-mail: mdshafiqul.islam603@gmail.com

T. Hossain
Applied Science for Integrated System Engineering, Kyushu Institute of Technology, Japan
e-mail: tahera.moni@gmail.com

M.A.R Ahad
Dept. of Media Intelligent, Osaka University, Japan;
Dept. of EEE, University of Dhaka, Bangladesh
e-mail: atiqahad@du.ac.bd

S. Inoue
Dept. of Human Intelligence Systems, Kyushu Institute of Technology, Japan
e-mail: sozo@brain.kyutech.ac.jp

Discriminant Analysis (LDA), Support Vector Machine (SVM), etc. for classifying activities. We have achieved satisfactory results. Our results show that using both accelerometer and gyroscope sensors can improve performance. We achieve accuracy of 80.45% by LDA, 93.34% by SVM, 91.92% by RnF, 91.64% by KNN, and 93.76% by CNN while we exploit both accelerometer and gyroscope sensor data together. The results demonstrate the prospect of eSense device for detecting human activities in various healthcare monitoring system.

Key words: Healthcare, Earables, Wearable, Activity recognition, eSense

1 Introduction

The wearable device provides good scope to measure different types of human activities. To support elderly people, it is important to understand the regular and complex human activity. Researchers are widely working to improve sensor-based and vision-based activity recognition [1, 20]. It is important for the healthcare monitoring center [21, 14] assisted living facility center, surveillance, etc. Through advanced technology, it is possible to detect human body movement. In this regard, wearables open new possibilities for detecting various types of human activities. Through different kinds of wearable sensors accelerometer data, it is possible to detect human body movement. In most cases, low powered and low-cost wearable sensors are mounted on the human body for analyzing human behaviors [13]. Many researchers are actively working on human activity recognition over a decade [4]. In this system, first, the sensor data (e.g., accelerometer, gyroscope) are collected from different users. Then these data are sent to servers. Data labeling is a big challenge in this field. In this context, several mobile apps are developed for human activity recognition [15]. Researchers are also conscious of real-time applications for hospitals and nursing homes [16]. Collecting data in a controlled environment and wild is quite different [17]. In the healthcare monitoring center, it is important to detect the behavioral activities of elderly people related to head and mouth. It is quite difficult to detect some kind of mouth-related activities like speaking, eating, laughing, etc. through regular wearable devices. These activities have an impact to measure the well-being of a person and socialization ability. In assisted living and healthcare centers, caregivers also need to record and manage activities of elderly people like the amount of food intake, swallowing ability, etc. These are measurement criteria for good health for elderly people. Taking this into consideration, in this paper we mainly focus on detecting head- and mouth-related activities with some other regular physical activities using eSense device. It can be used for monitoring and analyzing personal scale activities [8, 22]. eSense is a new device from Nokia Bell Lab, UK [8]. We acknowledge them for providing the devices to explore in various applications and methods. In this paper, hence, we explored as in the paper. eSense is a very light-weight, wireless device. It performs all required functionalities of earbud like listening to music, receiving phone calls, etc. Besides it contains accelerometer and gyroscope sensors which enables it to track head and mouth-related activities drinking, eating, shaking, speaking, etc. These types of activities cannot be detected using smartphone sensors. eSense device can detect minute head and neck movements that have potential applications in clinical medicine related to head and neck injury. Besides, with the help of conversational activity monitoring capabilities of this device, social interactions can be detected. As a result, it will help to treat different types of mental health conditions. In a nutshell, eSense has potential usability in the areas of computational social science, healthcare, and well-being.

eSense is a multi-modal device. It has a 6-axis inertial measurement unit (IMU) with Bluetooth unit. Using eSense devices, it is possible to collect real-time data – audio, motion, and proximity. From eSense we can collect three-axis accelerometer and gyroscope data to detect various human activities. It is also possible to collect audio data from eSense which can be impactful to distinguish some similar pattern of mouth movement. There are various types of popular machine learning methods for sensor data exploration. Feature selection has a significant impact on several machine learning approaches for measuring accuracy [13]. For deep learning methods like Convolutional Neural Network (CNN), manual feature extraction is not required [6]. CNN can learn to identify complex patterns. In this paper, our goal is to detect head- and mouth-related human activities with other regular activity through eSense. We explored both traditional machine learning approaches and CNN to classify selected activities.

For detecting human activities different types of sensor data (accelerometer, gyroscope, etc.) are collected from the smartphone, smartwatch, Inertial Measurement Unit (IMU), etc. These devices are placed in different body parts i.e. wrist, leg, waist for detecting activities like walking, sitting, running, etc. From these devices head and mouth-related activities i.e. eating, nodding, drinking, shaking, etc. cannot be detected properly. eSense helps to detect such activities. It is an earbud that contains an accelerometer, gyroscope sensors. From these sensor data such head and mouth-related activities can be detected. So initially we have taken the challenge to detect some of the useful gesture types like speaking, head shaking, nodding, eating, walking, etc. Besides we have developed our own Android application "eSenseLog" to connect the smartphone to eSense device to fetch accelerometer, gyroscope, and audio data. We also presented a relative comparison of various models' performance for detecting head and mouth-related activities from the only accelerometer and both accelerometer, gyroscope data. There are other activities like measuring VO2 max, monitoring heart rate, chewing, swallowing, etc. These activities are important in nursing care. In the future, we will detect these activities. Our current work will play a vital role in detecting such activities.

The objectives of this paper are as follows: collecting accelerometer and gyroscope data from eSense device through Bluetooth and mobile application; detecting head- and mouth-related human activities alongside some other regular activities; using traditional machine learning and deep learning classifiers to detect activities and compare performances. We compare the result performance among machine learning technique and deep learning technique as well as we evaluate the result while having only accelerometer sensor data and having accelerometer and gyroscope sensor data together. Having both sensors' data demonstrate significantly higher recognition results (e.g., for LDA: 4%, for SVM: 6%, for RnF: 3.5%, and for CNN: 3%).

The organization of the paper is as follows: after providing a brief introduction in Section 1, we present related work in Section 2. In Section 3, we introduce the proposed framework and system architecture. In Section 4, we provide the details of the experimental setup and results. Finally, we conclude the paper with some analysis and future work guidelines.

2 Related Works

Sensor-based human activity recognition is a popular research area. The data is collected from mobile sensors; wearable sensors etc. and then machine learning algorithms are applied to detect human activities. Many deep learning algorithms like Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) [33], Transfer Learning [32], etc. have been used to recognize different types of activities from sensor data. There are different types of sensors in wearable devices. Especially, human movement data can be collected from an accelerometer and gyroscope sensor. Machine learning algorithms can be applied to these data to detect different types of human activities [10]. For recognizing simple and complex activities different types of approaches have been used by researchers. It is comparatively easier to detect human activity from controlled lab environment predefined activity class data. Inoue et al. [17] used real nurse care data and evaluated human activity recognition performance. They collected various activities' data from elderly people of a nursing care facility for 4 months. They used smartphones to collected data from the built-in sensors of smartphones. The authors introduce a problem of data annotation issues in a real-life scenario. On the other hand, choosing proper sensing modality and sensing location is difficult which is used to track head- and mouth-related activities. Earbuds are very popular for listening to music [34]. As earbud is very lightweight, many types of research have been done to integrate various sensors in it. From these sensors data, different types of activities can be detected. Measuring VO2max, monitoring heart rate, assessing energy expenditure, etc. were done by LeBoeuf et al. [35]. Taniguchi et al. [36] tried to develop an earphone type device to measure chewing count reliably. Monitoring cardiovascular and sweat parameters during physical exercise using the wireless ear-worn device was proposed by Bruno et al. [37]. Exploring audio and kinetic sensing using an earable device was done by Min et al. [38]. Louis et al. [39] used an ear-worn sensor for gait monitoring. Bojan et al. [42] analyzed jump performance using wearable inertial sensors. Bedri et al. [18] introduced 'EarBit' wearable sensor to detect eating episodes. The authors used a semi-controlled and outside lab environment for assessing the performance of EarBit device. The authors found a satisfactory result to recognize eating episodes. For detecting real-time swallowing a neck-worn system was used [19]. They achieved a precision of 67.7% and a recall performance of 79.9%. In Ref. [7], six wearable microphones were used to record chewing sounds, and then the performances were compared. It was found that the inner ear location provides the best acoustic signal intensity for detecting chewing sound through a wearable microphone.

Time and frequency domain features can be extracted from sensor data (accelerometer, gyroscope) for activity recognition [2]. For the classifying dynamic activities from accelerometer data, Preece et al. [11] made a comparison among some feature extraction methods. In [9], the authors used the Human Activity Sensing Consortium (HASC) tool. It is used for collecting data from different activities. They used eight features like mean, variance, fast Fourier transform (FFT) with zero-crossing rate, power of four bands, etc. to classify the activities. In [12], the authors proposed a method for HAR where the position of the wearable devices on the human body can be changed dynamically based on user preference. Gao et al. [3] used multiple sensors instead of a single sensor to recognize actions from accelerometer data. LoRaWAN technology was proposed by the authors in [5]. It is quite beneficial for healthcare monitoring service since LoRa can cover several kilometers through one gateway. Many types of research are ongoing in this area to explore this new stable platform for the healthcare domain.

Several research works have been done using eSense device. Tobias et al. [23] propose respiration rate monitoring using eSense device. They took several approaches to clean noisy data and calculate the respiratory rate on 20-second windows. They compared accelerometer and gyroscope data from twelve participants and employed pressure-based measurement with nasal cannulas as ground truth. Head motion tracking using accelerometer and gyroscope data was done by A. Ferlini et al. [24]. They increased accuracy up to a few degrees by leveraging accelerometer and gyroscope data for each earbud. M. Radhakrishnan et al. [25] proposed a low-cost system. It can monitor multiple users by fusing data from eSense device and other inertial sensors attached to exercise equipment. Robust step count using eSense device is proposed in [26]. While walking, the head movement can produce noise. They proposed a method to alleviate this noise and showed a 95% step count accuracy even in the most difficult scenarios. S. Rupavatharam et al. [27] proposed a jaw clenching detection technique using eSense device. It detects peaks/dips in the gyroscope vector magnitude. They showed an error rate of 1% when the person is stationary and 4% when moving. In paper [28], an experiment regarding the effect of acoustical manipulation with eSense was showed. They showed that the manipulation reduced deviation while walking straight in both subtle and overt conditions. Alexander proposed a method [29] to annotate activity data among multiple sensors using eSense device. They used cross-correlation to propagate the annotations across all sensor streams. They experimented with 7 persons, each performing 6 different physical activities. Their proposed algorithm synchronizes signals using cross-correlation within a second. J. E. Bardram et al. [30] integrated eSense computing platform with a programming framework to design custom mobile health (mHealth) applications. H. Odoemelem used eSense device [31] to control a lightweight robot arm.

There are many challenges in wearable sensors to explore it in some real fields [40, 41]. Sensor position, different methods, types of sensors are important factors to extract activities data properly. There are many challenges in introducing a new system for HAR. In this paper, we introduce a new platform for eSense devices to recognize head- and mouth-related activities alongside some regular activities.

3 Proposed Framework and System

The eSense device can be used to track head- and mouth-related activities like nodding, eating, swallowing, speaking, etc. It can be used for detecting other activities like walking, staying, etc. There are accelerometer and gyroscope sensors in eSense device. So these sensors data can be used to recognize human activities. Besides the audio data from eSense device can be used to track other activities like gossiping, laughing, etc. After connecting the eSense device to a mobile device, sensor data can be collected. The eSense device is regarded as a peripheral device and the mobile is regarded as a host device after establishing the connection between the mobile and the eSense device. Through an Android application, the sensor data and audio data are collected. Our developed Android application continuously maintains a connection with eSense device and helps to collect data from the device. In this section, we present our developed Android application, system architecture, and framework.

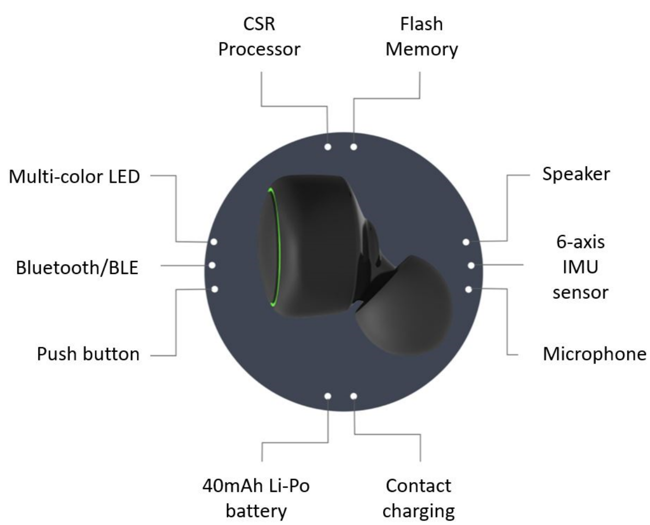

Fig. 1: Overview of eSense device

In Fig. 1, an overview of eSense device is presented. Advanced Audio Distribution Profile (A2DP) is used in the firmware of eSense device. A2DP is used for mono channel recording and high-definition audio streaming [8]. A thin middleware is used in eSense to support application development in Android and iOS platforms.

The Node.js middleware allows developers to connect eSense devices with desktop applications, configure and ingest sensory data in real-time.

The size of both eSense device and standard wireless earbud is equal and it contains a battery, electronics, etc. The eSense device contains a custom-designed PCB (Printed Circuit Board) whose size is 15 x 15 x 3 mm [8]. The eSense device has System-on-Chip (SoC), an Inertial Measurement Unit (IMU). The IMU contains a three-axis gyroscope and accelerometer. It also contains a two-state button; a circular LED, a digital motion processor, and battery-charging circuitry. The eSense device is powered by a 40-mAh LiPo battery. The weight is each earbud is 20g and the size is 18 x 20 x 20 mm. There is no flash memory in eSense device.

The right earbud does not have a Bluetooth Low Energy (BLE) interface. On the other hand, the left earbud has a BLE interface that can be used to configure various aspects of the IMU sensor and collect accelerometer and gyroscope data. By default, It transmits periodic BLE advertisement packets. The interval is between 625ms and 750ms. The advertisement packets contain the Complete List of 16-bit Service Class UUIDs (Universally Unique Identifier). All the standard UUIDs are used in eSense [8] device.

eSense is a Bluetooth Low Energy (BLE) device. It is based on a specification called “General ATTribute profile” (GATT). The specifications to send and receive attributes between a server and a client are defined by GATT. Attributes are short pieces of data. In this case, eSense is the client and the connected host device (mobile) is the server. The Attribute Protocol (ATT) is used as the base to build GATT on top of ATT. GATT consists of “characteristics”, “services”, “profiles”, “descriptors”. A profile consists of one or more services.

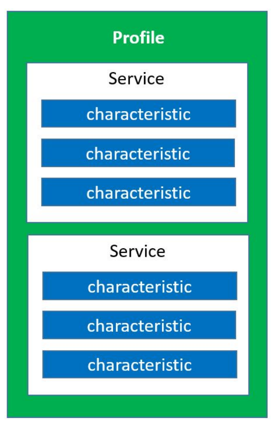

Fig. 2: GATT diagram

In Fig. 2, the GATT diagram is explained. Each service consists of one or more characteristics that encapsulate data. In eSense, there are such services that provide accelerometer, gyroscope data. A characteristic has a single value and 0-n descriptors.

The descriptors describe the characteristic's value. For example, a descriptor might specify a unit of measure which is specific to the characteristic's value, a human-readable description, etc. It is similar to a class.

When the connection gets established between a Bluetooth Low Energy (BLE) device to a host device like a smartphone, then the Generic Attribute Profile (GATT) is used. This general specification is used for sending and receiving 'attributes' over a BLE link. Attributes are short pieces of information.

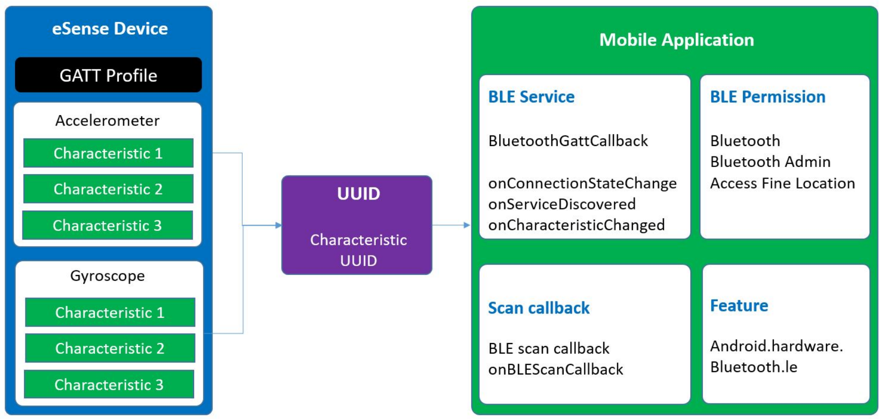

Fig. 3: A general system architecture

In Fig. 3, an overall system architecture is explained. It is for the BLE device and a host device. The eSense has accelerometer and gyroscope sensors. Each sensor (accelerometer or gyroscope) sends some characteristic data. This data contains a single value and 0-n descriptors. The descriptor describes the characteristic's value. Each GATT profile contains one or more services. The service has some characteristics. A Universally Unique Identifier (UUID) is used to identify each attribute uniquely. The UUID has a string ID which size is 128-bit.

The host device (for example, a mobile phone) scans for nearby devices to connect. There are some callback methods, which are used in the application. These callback methods are used in different scenarios. After establishing the connection between mobile and eSense device, onConnectionStateChanged() method is called. The onCharacteristicChanged() callback method is responsible when data is sent from the eSense device to a mobile device. The whole process runs in service to make it runnable in the background. Inside the application, different types of callback methods and normal methods are used for establishing a connection, sending and receiving data.

We have developed an Android application named 'eSenseLog' for collecting audio and sensor data from the eSense device (Fig. 4 shows a screenshot of it). There are several features of this application. A user can connect eSense device with mobile phone by clicking the 'Connect' button as shown at the top-left corner of Fig. 4. If the device gets connected, its name and connected status will be shown.

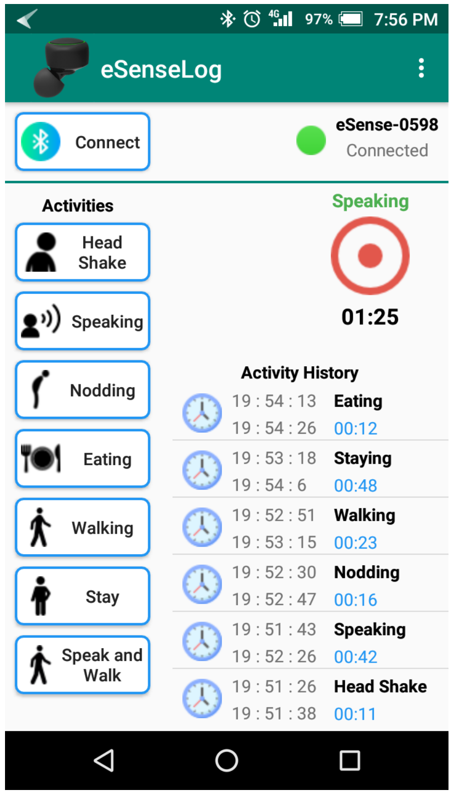

Fig. 4: Overview of eSenseLog Android application

In Fig. 4, the overview of our developed eSenseLog application is presented. After connecting the device, if someone wants to collect any activity data, he needs to click an activity button from the list of activity buttons. There is a toggle 'start' and 'stop' button icon. After selecting the activity, if a user clicks 'start' toggle button, the data collection will be started. The elapsed time will be shown under the toggle button. If the user clicks the toggle button again, data collection will be stopped and saved in the mobile storage. The sensor data will be saved as '*.csv' format and the audio data will be saved as '*.3gpp' format. In the app, there is an activity history list from where the user can easily track the collected activities and their durations.

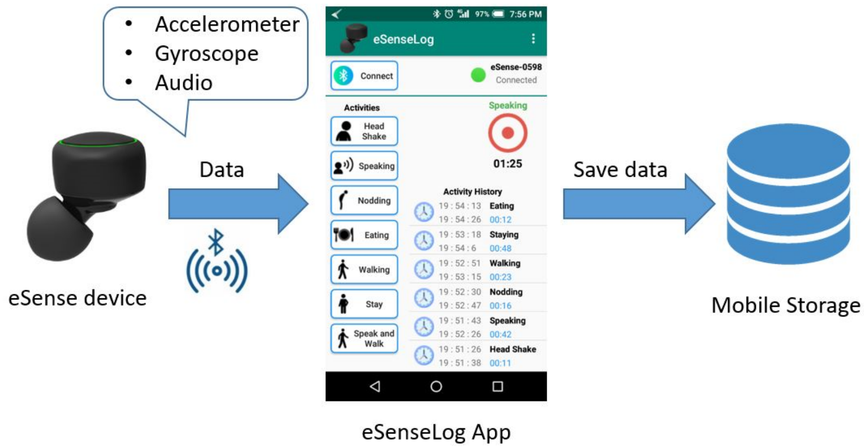

Fig. 5: The proposed framework for this work

In Fig. 5 our proposed framework is shown. The audio and sensor data from the accelerometer and gyroscope will be collected simultaneously. These data are saved in the mobile phone's internal storage. However, if there is not enough space and if any external storage (e.g., a microSD card) is available then the data will be saved in external storage. When the data collection is finished, a user can fetch the data from the mobile device and then apply any machine learning algorithm on the collected data for recognizing activities.

4 Experimental Result and Analysis

We performed experiments for seven activities. Four activities are related to head and mouth (namely, eating, speaking, headshaking, and head nodding). Three regular activities (i.e. walk, stay, and speaking while walking). These activities are detected from eSense device's accelerometer and gyroscope data. Each activity was performed for 3 minutes. One-person data collection time was 21 minutes for all activities in total. We have collected these activities data from 6 persons (having age range from 25 to 35 years, 4 males and 2 females). There are around 105881 records for these seven activities. The sampling rate was 50Hz while the data was collected from 3-axial accelerometer and gyroscope sensors of the eSense device. Figure 4 shows each activity class count in the dataset.

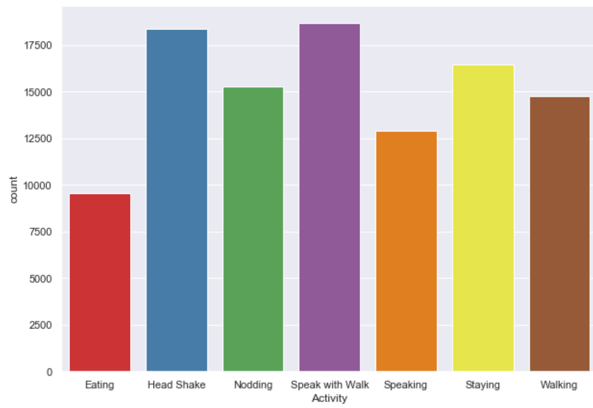

Fig. 6: Bar plot for Activities in the dataset

In Fig. 6, the activity count from the collected data is represented. We collected data from multiple persons to split train and test data. Based on different users, we separated training and test data with different trials. These are not continuous classes. K-fold cross-validation technique was used to separate train and test data. We split the training dataset into k-folds based on the completely separate person data for training and a different person for testing. During the training procedure, we use first kth folds for model training and holdout another fold to use in testing. We use this k-fold splitting technique because in our dataset there are 6 persons' data and we want to isolate test and train data in different persons with different trials. We have used all person's data for training and testing. Normally training a model needs lots of data. We have used 1 person out cross-validation. We did not use only one person's data for training and then testing on the rest of the person's data. Rather, we used 5 persons' data for training and 1 person's data for testing.

We did not use any kind of feature selection techniques. We have calculated the magnitude value of raw accelerometer and gyroscope data for extracting features. Some of the extracted time-domain statistical features are maximum value, minimum value, mean, standard deviation, etc. We used a sliding window approach with 50% overlap. We used several traditional machine-learning algorithms, i.e., Random Forest (RnF), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM) with radial basis function kernel, K-Nearest Neighbor (KNN). For the deep learning model, we used 1D CNN. For 1D CNN model we did not extract any features as CNN automatically extract features. Among the four traditional machine-learning methods, KNN achieved the highest accuracy of 92.43% when we used only accelerometer sensor data. On the other hand, SVM performs better while we used both accelerometer and gyroscope sensor data together for evaluation.

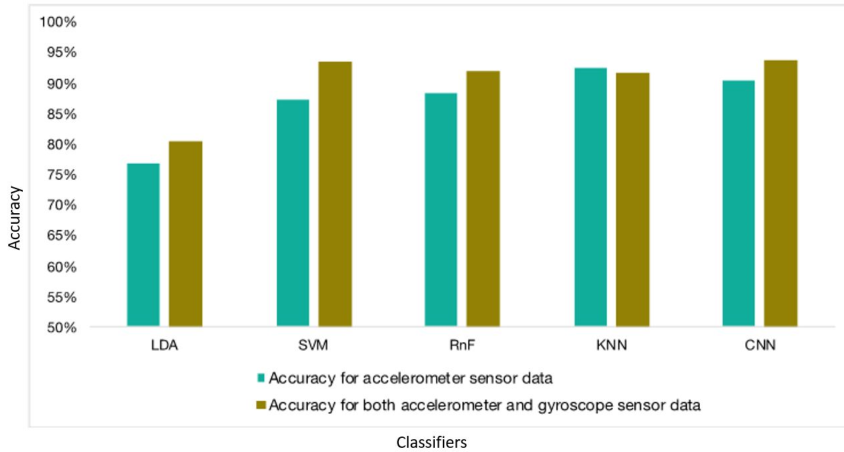

Fig. 7: Accuracies comparison obtained from accelerometer sensor data and both accelerometer and gyroscope sensor data

In Fig. 7, the performance comparison among different classifiers is explained. Afterward, we used 1D CNN model. CNN model achieved the highest accuracy (93.76%) among all other classifiers, with the combined accelerometer and gyroscope data. In the CNN model, 2 convolutional layers were used. Each convolutional layer has 64 feature maps and the kernel size is 1x3. A max-pooling layer is used after the second convolutional layer. The pool size of the max-pooling layer is 1x2. Then two fully connected (dense) layers are used after the max-pooling layer. Finally, a softmax layer is employed. In this model, we used ReLu activation function and Adam optimizer. These are very well-established methods for optimization and activation. In this model, there is no need to extract hand-crafted features from the sensors' data, because the CNN model can do it as per the model.

The comparison of accuracy from different classifiers by using accelerometer sensor data, as well as, by using both accelerometer and gyroscope sensor data are demonstrated in Fig. 7. On the other hand, with accelerometer-only data, the achieved recognition results are 76.75% by LDA, 87.39% by SVM, 88.23% by RnF, 92.43% by KNN, and 90.55% by CNN. We can notice from these achieved results in five classifiers that when we exploit both sensors' data, the results are superior to using accelerometer-only sensor data. Only one exception is the case with KNN. Using KNN classifier, the results are slightly better with accelerometer-only data. However, the other four cases, having both sensors' data demonstrate significantly higher recognition results (e.g., for LDA: 4%, for SVM: 6%, for RnF: 3.5%, and for CNN: 3%). Hence, we can conclude that both sensors are more suitable for better recognition results in this case.

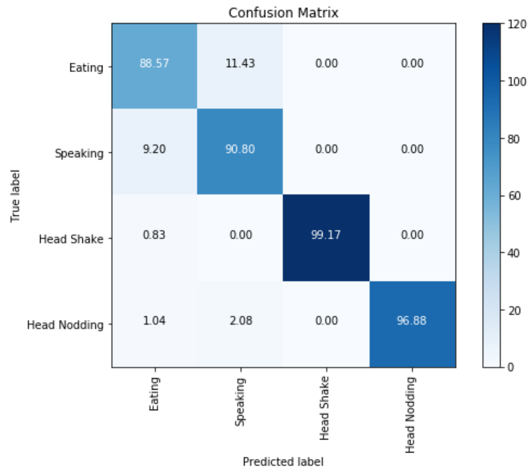

Fig. 8: Confusion matrix only head-mouth activities by RnF

In Fig. 8, we observe that only head- and mouth-related activity perform better results with the eSense. In this case, major miss-classification occurred among 'Eating' and 'Speaking' classes as these classes have closely similar patterns of muscle movement of the mouth.

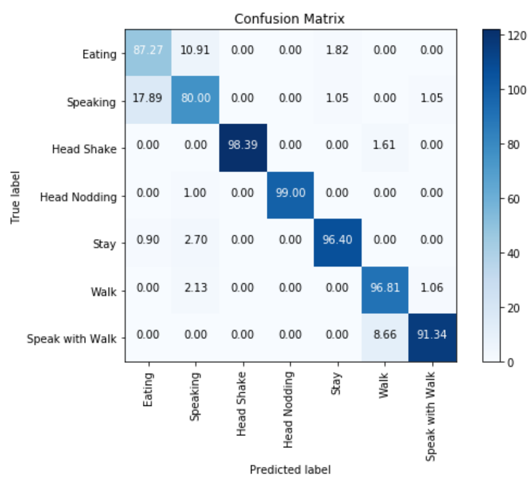

Fig. 9: Confusion matrix for SVM with regular activities

From Fig. 9, it is found that if regular activity classes (i.e. 'Walk', 'Stay' and 'Speak with Walk') are added with the previous four head- and mouth-related activity classes, then the performance of the classifiers drops. From the confusion matrix, we noticed that the ’Speaking’ class is confused with ‘Eating’ class ‘Eating’ class is mostly confused with ‘Stay’ class (Fig. 9). Both ‘Eating’ and ‘Stay’ actions need less movement and these activities can be performed simultaneously. So there is a good chance of getting confused with each other for these activities. Besides, ’Eating’ and ’Speaking’ activities require the mouth’s movement. Therefore, these 2 activities have similar patterns. Moreover, ‘Speaking’ activity data was recorded in the sitting state (no movement) for this study. This is the main reason for the confusion between activities. Sometimes, ‘Head shake’, ‘Nodding’ activities also get confused with ‘Stay’ class. The reason is, these activities were also performed while sitting on a chair (Fig. 10).

Fig. 10: Confusion matrix for CNN classifier with other regular activities

From Fig. 10 it is seen that the recognition results for the head- and mouth-related activities are better than the regular activities (i.e. ‘Walk’, ‘Stay’ and ‘Speak with Walk’).

5 Conclusions and Future Work

Wearables sensors are used to recognize activities of daily life. Accelerometer and gyroscope sensors are mostly used for this purpose. There are several types of behavioral activities (i.e. nodding, drinking, eating, shaking, etc.) which are related to head and mouth. Detecting these activities accurately using only accelerometer and gyroscope sensors data quite challenging. Also, the sensor position plays a vital role to collect proper data for the head- and mouth-related behavioral activities. In this regard, an earable device like 'eSense' can play a vital role to detect head- and mouth-related activities. In this paper, we used several machine-learning methods to evaluate the performance of detecting head- and mouth-related activities as well as other regular activities from sensor data of eSense device. We compared the recognition performance of several classifiers using only accelerometer data and using both accelerometer and gyroscope sensor data. We used both traditional and deep learning methods for classification. We used Random Forest, Support Vector Machine, Linear Discriminant Analysis, K-Nearest Neighbor, and 1D Convolutional Neural Network. Having both accelerometer and gyroscope data, Convolutional Neural Network achieves the best accuracy (93.76%) among all classifiers, which is slightly better than Support Vector Machine accuracy (93.34%) among many machine learning classifiers. Our results demonstrate that using both accelerometer and gyroscope significantly improves recognition performance.

However, there exist some misclassifications among different activities that require body movement and which don't require any kind of body movement. Also, head- and mouth-related activities demonstrate slightly better results than normal activities. Therefore, using our proposed framework, it will be easier for collecting head- and mouth-related activity data. Any machine-learning model can be applied to this data to detect head- and mouth-related activities. It will create a new scope in research and applications using earable device. For this experiment, we used a lab environment. In the future, we will deploy our system to a nursing care center to collect real-life data. Using our proposed system, we also collected audio data. In the future, we will use audio data to reduce misclassification related to 'Speaking' and 'Eating' activities, which have some similar patterns due to the muscle's movement in the face. We will also increase the number of activities including complicated activities, which will be impactful for real nursing care/healthcare centers.

Acknowledgements We acknowledge Pervasive Research Unit Nokia Bell Labs, Cambridge and Prof. Fahim Kawsar 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, and 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, M.A.R Ahad, and 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. O. Amft, M. Stager, P. Lukowicz, and G. Troster. “Analysis of Chewing Sounds for Dietary Monitoring”. In Proceedings of the 7th International Conference on Ubiquitous Computing (UbiComp’05), 2005.
8. F. Kawsar, C. Min, A. Mathur, and A. Montanari, “Earables for Personal-Scale Behavior Analytics”, IEEE Pervasive Computing, 17, 2018.
9. N. Kawaguchi, N. Ogawa, Y. Iwasaki, K. Kaji, T. Terada, K. Murao, S. Inoue, Y. Kawahara, Y. Sumi, and N. Nishio. “HASC Challenge: gathering large scale human activity corpus for the real-world activity understandings”. In Proceedings of the 2nd augmented human international conference. ACM, 2011.
10. G. Ogbuabor and R. La, “Human Activity Recognition for Healthcare 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.
14. S. Abbate, M. Avvenuti, P. Corsini, J. Light, and A. Vecchio, “Monitoring of Human Movements for Fall Detection and Activities Recognition in Elderly Care Using Wireless Sensor Network: a Survey”, INTECH, 2015.
15. Lara, D. Oscar, and A. Miguel, “A mobile platform for real-time human activity recognition,” Consumer Communications and Networking Conference (CCNC), 2012, IEEE, 2012.
16. N. Mairittha, T. Mairittha, S. Inoue, “FonLog - A Mobile App for Nursing Activity Recognition,” ACM Int’l Conf. Pervasive and Ubiquitous Computing (Ubicomp) Demo, 2018.
17. 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.
18. A. Bedri, R. Li, M. Haynes, R. P. Kosaraju, I. Grover, T. Prioleau, M. Y. Beh, M. Goel, T. Starner, and G. Abowd. “EarBit: Using Wearable Sensors to Detect Eating Episodes in Unconstrained Environments,” Proc. ACM Interact. Mob. Wearable Ubiquitous Technologies, 2017.
19. T. Olubanjo and M. Ghovanloo. “Real-time swallowing detection based on tracheal acoustics,” IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP), 2014.
20. M.A.R Ahad, “Motion History Images for Action Recognition and Understanding,” ISBN: 978-1-4471-4730-5, Springer, 2013.
21. M.A.R Ahad, “Computer Vision and Action Recognition: A Guide for Image Processing and Computer Vision Community for Action Understanding,” ISBN: 978-94-91216-20-6, Springer, 2011.
22. T. Hossain, M. S. Islam, M. A. R. Ahad and S. Inoue, “Human Activity Recognition Using Earable Device,” ACM International Joint Conference on Pervasive and Ubiquitous Computing and the International Symposium on Wearable Computers (UbiComp/ISWC ’19 Adjunct), 2019.
23. T. Röddiger, D. Wolffram, D. Laubenstein, M. Budde and M. Beigl, “Towards Respiration Rate Monitoring Using an In-Ear Headphone Inertial Measurement Unit,” 1st Int. Workshop on Earable Computing, London, UK, September 10, 2019.
24. A. Ferlini, A. Montanari, C. Mascolo and R. Harle, “Head Motion Tracking Through in-Ear Wearables,” 1st Int. Workshop on Earable Computing, London, UK. September 10, 2019.
25. M. Radhakrishnan and A. Misra, “Can Earables Support Effective User Engagement during Weight-Based Gym Exercises?” 1st Int. Workshop on Earable Computing, 2019, London, UK. September 10.

26. J. Prakash, Z. Yang, Y. Wei and R. R. Choudhury, “EStep: Earabales as opportunity for physio-analytics,” 1st Int. Workshop on Earable Computing, London, UK. September 10, 2019.
27. S. Rupavatharam and M. Gruteser, “Towards In-Ear Inertial Jaw Clenching Detection,” 1st Int. Workshop on Earable Computing, London, UK. September 10, 2019
28. K. Matsumura and K. Okada, “eSense Veers: a Case Study of Acoustical Manipulation in Walking without Sight both on Subtle and Overt Conditions,” 1st Int. Workshop on Earable Computing, London, UK. September 10, 2019
29. A. Hölzemann, H. Odoemelemand and K. V. Laerhoven, “Using an in-ear device to annotate activity data across multiple wearable sensors,” 1st Int. Workshop on Earable Computing, London, UK. September 10, 2019.
30. J. E. Bardram, “The CAMS eSense Framework – Enabling Earable Computing for mHealth Apps and Digital Phenotyping,” 1st Int. Workshop on Earable Computing, London, UK. September 10, 2019.
31. H. Odoemelem, A. Hölzemann and K. V. Laerhoven, “Using the eSense Wearable Earbud as a Light-Weight Robot Arm Controller,” 1st Int. Workshop on Earable Computing, London, UK. September 10, 2019.
32. M.S. Islam, T. Okita, and S. Inoue, “Evaluation of Transfer Learning for Human Activity Recognition Among Different Datasets”, 2019 IEEE International Conference on Pervasive Intelligence and Computing (PiCom), pp. 6 pages, Fukuoka, Japan, August 5, 2019.
33. M. Inoue, S. Inoue, T. Nishida, “Deep recurrent neural network for mobile human activity recognition with high throughput,” arXiv:1611.03607, 2016
34. S. Nirjon et al., “MusicalHeart: A Hearty Way of Listening to Music,” Proc. 10th ACM Conf. Embedded Network Sensor Systems (SenSys 12), pp. 43–56, 2012.
35. S.F. LeBoeuf et al., “Earbud-Based Sensor for the Assessment of Energy Expenditure, Heart Rate, and VO2max,” Medicine and Science in Sports and Exercise, vol. 46, no. 5, pp. 1046–1052, 2014.
36. K. Taniguchi et al., “Earable RCC: Development of an Earphone-Type Reliable Chewing-Count Measurement Device,” J. Healthcare Eng; doi.org/10.1155/2018/6161525; 2018.
37. B. Gil, S. Anastasova, and G. Z Yang. “A Smart Wireless Ear-Worn Device for Cardiovascular and Sweat Parameter Monitoring During Physical Exercise: Design and Performance Results,” Sensors 19, 7, 1616, 2019.
38. C. Min, A. Mathur, and F. Kawsar, “Exploring audio and kinetic sensing on earable devices,” Proceedings of the 4th ACM Workshop on Wearable Systems and Applications, WearSys ’18, pages 5–10, New York, NY, USA, 2018.
39. L. Atallah, AnatoleWiik, G. G Jones, B. Lo, J. P Cobb, A. Amis, and G. Yang. “Validation of an ear-worn sensor for gait monitoring using a force-plate instrumented treadmill,” Gait and posture 35, 4; 674–676, (2012).
40. M.A.R. Ahad, “Vision and Sensor based Human Activity Recognition: Challenges Ahead,” Chapter 2, in "Advancements in Instrumentation and Control in Applied System Applications", IGI Global, 2020.
41. M.A.R. Ahad, Anindya Das Antar, and Omar Shahid, “Vision-based Action Understanding for Assistive Healthcare: A Short Review”, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, USA, pp. 1-11, 2019.
42. B. Milosevic and E. Farella, “Wearable Inertial Sensor for Jump Performance Analysis,” Proceedings of the 2015 Workshop on Wearable Systems and Applications (WearSys ’15). ACM, New York, NY, USA, 15–20. DOI:https://doi.org/10.1145/2753509.2753512