Detection of smoking events from confounding activities of daily living

Jianchao Lu
Department of Computing
Macquarie University
Sydney, Australia
jianchao.lu@hdr.mq.edu.au

Jiaxing Wang
Department of Computing
Macquarie University
Sydney, Australia
jiaxing.wang@ mq.edu.au

Xi Zheng
Department of Computing
Macquarie University
Sydney, Australia
james.zheng@mq.edu.au

Chandan Karmakar
School of IT, Deakin University
Geelong, Victoria, Australia
Karmakar @deakin.edu.au

Sutharshan Rajasegarar
School of IT, Deakin University
Geelong, Victoria, Australia
sutharshan.rajasegarar@deakin.edu.au

ABSTRACT

Although smoking prevalence is declining in many countries, smoking related health problems still leads the preventable causes of death in the world. Several smoking intervention mechanisms have been introduced to help smoking cessation such as counselling program, motivational interview and pharmacotherapy. However, these methods lack providing real time personalized intervention messages to the smoking addicted users. The challenge is to develop an automated smoking behavior detection. We address this challenge by proposing a non-invasive sensor based automated framework for smoking behavior detection. We used a wristband based accelerometer and gyroscope sensors to detect smoking activities, differentiating with the closely confounding activities. We extract several features using learning algorithms and the empirical results with our participants show good accuracy in detecting the smoking activity in terms of precision, recall, and F1-score.

1 Introduction

Tobacco consumption, primarily cigarette smoking plays an adverse effect on human health and its increasing consumption during last couple of decades has made it well known to be harmful for human beings health. One of the study claims that one third of the population who try smoking become addicted to nicotine, the leading cause of mortality and preventable morbidity [1]. A government study in Australia [2] reveals the detrimental effects of tobacco smoking in the human health. According to the report, along with nicotine, smokers inhale about 7000 other chemicals, and among which over 60 are known cancer-causing chemicals. A recent report by World Health Organization [3] points out that smoking is considered to be one of the major public health threats, which results around 6 million deaths annually. It is estimated [4] that in Australia alone, smoking is responsible for 15; 000 deaths and costs 31.5 billion AUD each year. Although the negative effects of smoking are known to majority of smokers, smokers' addiction to it has forced them to ignore the resulting effects. The increasing numbers of smokers and the resulting mortality rates have indicated the importance of understanding smoking behaviors such as frequency of smoking for evaluating and improving the effectiveness of behavioral smoking interventions [5].

Smoking monitoring and intervention systems have been well studied since last couple of decades with the use of different available technologies ranging from computer vision, self-report to wearable sensing and mobile technology [6] [7]. However, computer vision systems were considered to be an inaccurate and impractical way of monitoring (especially when privacy is considered). With the development of the miscreation of sensors technology, there has been a shift towards using inertial sensors which are integrated into a wearable device, such as smart phone and wristband, for detecting smoking behavior.

In this paper, we aim to analyze the performance of smoking gesture recognition from other confounding gestures using different sensors. The main idea is to compare the recognition performance of an accelerator-only approach with a gyroscope-only approach, and a combination of accelerator- gyroscope approach. For this purpose, we categorize the features into two classes: orientation features and rotation features. Features from the accelerometer and gyroscope are used as the orientation features, whereas features from the Euler angles pitch, yaw, and roll computed by fusing the accelerometer and gyroscope measurements are used as the rotation features.

In order to investigate the performance of smoking gesture recognition from other confounding gestures, we collected a dataset from five participants. An Arduino 101 development board was used for the data collection under a sampling rate of 20Hz. Five confounding gestures were performed: smoking, drinking, eating, scratching head, and biting nails. Both accelerometer and gyroscope readings were used for extracting features. An ensemble based feature selection was employed for constructing relevant feature subset. In the experiment, we used classifiers such as SVM and random forest, and their classification performance of the confounding activities are analyzed.

Our evaluation reveal that the use of both accelerometer and gyroscope achieve the highest accuracy, while of the accuracy using only one accelerometer sensor shows comparable performance, However, using only a gyroscope did not perform well. Moreover, random forest classifier provided the best performance in our experiment. The following lists the main contributions of this paper:

1. We considered confounding activities for analysis such as smoking, drinking, eating, scratching head, and biting nails rather than concurrent activity such as smoking while standing, smoking while sitting, smoking while in a group conversation, and smoking while walking.
2. We categorized the set of features into two categories: orientation features and rotation features. We analyzed the performance of confounding activity recognition using orientation and rotation features, both alone and in combination.
3. We analyzed the performance of accelerometer only solution, gyroscope only solution and accelerometer and gyroscope combined solution (includes rotation features only solution and both orientation and rotation features solution), and compared their performance with different classifiers (SVM and random forest).

The paper is organized as follows. Section 2 talks about the state of art in smoking behavior detection using wearable device. Section 3 talks about the details of our solution, which involves data collection, feature extraction and selection, and data Interpretation. Section 4 discusses the experiment setup and our evaluation results and in Section V we discuss our overall contribution and some future work.

2 Related Work

In this section, we walk through a few research works on smoking activity recognition.

In recent years, several works have been done in recognizing smoking activity and puffing (hand-to-mouth gestures). In [8], a smoking intervention system based on video camera is used to capture the movement and behavior of smokers. Captured image frames were subsequently processed and analyzed to produce a time-series, which is termed as smoking topography. Another work in [5] proposes a smoking intervention system using hand to mouth proximity sensor to detect the hand gesture, and using airflow sensor, thoracic and abdominal respiratory bands for detecting breathing pattern. They evaluated the uniqueness of the breathing patterns and the hand gesture movement through signal processing of these multiple sensors. They also evaluated their approach in a laboratory setting where 20 participants preformed 12 different activities including smoking cigarette. The results show a prediction rate of 90% and show a clear correlation between hand-to-mouth gesture and the respiratory signals. Similarly in [9], a smoking intervention system is created that utilizes respiratory signal to automatically detect the smoking puffs and the entire smoking episodes. The system has been tested with the 91% accuracy rate where 10 daily smokers were used in the evaluation. Although these works achieved satisfactory accuracy rates in their experiments, it is obviously that the presented detection systems with the required vest, various sensors, and electronics are the intrusive smoking detection, which is not convenient for the practical use in daily life.

With the development of the miscreation of sensors technology, there has been a shift towards using inertial sensors which are integrated into a wearable device, such as smart phone and wristband, for detecting smoking and other similar activities. For example, in [10], the authors used a wrist-worn sensor (an accelerometer) for smoking detection. They selected mean and variance as their features and Gaussian model as their classifier. This study highlighted the feasibility into activity recognizing by using wrist-worn inertial sensors. However they only considered smoking while standing in their evaluations, and failed to construct good feature engineering in their experiment which caused a low-accurate performance for smoking recognition (a user-specific recall of 70% and precision of 51%). Similarly, the authors in [18] also used the accelerometer sensor (but two at both wrists) for the smoking detection. Various forms of smoking activities, such as smoking while sitting, smoking while standing, smoking while eating, smoking while walking, smoking while using a phone, and smoking while talking in a group, were considered in their evaluation. 51 candidate features for each wrist were computed which included mean, standard deviation, maximum, minimum, median, kurtosis, skew, percentile, SNR and RMS of each window, peak-peak amplitude, peak rate, local peak point, correlation coefficients, crossing rate between axes, slope, MSE and R-squared. They designed a two-layer model as their activity classifier in which random forest classifier as the first layer is used to calculate the frequency of puff for a recent window and then a threshold-based rule as the second layer is applied to identify a smoking session. Compare with [10], the authors improved a lot in their work. However, they still did not achieve a high-accurate performance for smoking recognition (79% on F-measure). As the result, we are working on the assumption that only using the accelerometer sensor (s) is not enough for detecting concurrent activity while smoking and the confounding gestures such as smoking, eating and drinking.

In [11], the authors present a smoking gesture detection system using a smart watch with an accelerometer and a gyroscope. Data collection was performed with 30 participants of distinct age group and genders performing 4 different gestures such as smoking, drinking, phone calls, and crossing arms. The raw sensor data were recorded and from which a number of features were extracted including max speed, median speed, and mean speed, variance of speed, net roll velocity, median roll velocity, and max roll velocity. Multiple classification algorithms, including Naïve Bayesian Network, J48 Decision Tree, and Random Forest, were used to predict motions based on these feature sets. On average, Random Forest performs best with accuracy rate of 63.7%. The result shows their approach cannot distinguish in many cases between smoking and using phone due to a failed feature extraction. We notice there is no rotation features extracted in their experiment while most features are based on acceleration which concretes our assumption. In an obvious contrast to above study, the authors in [12] applied accelerometer, gyroscope and magnetometer at wrist to recognize smoking puffs. 32 features were extracted from the raw data which included 4 duration features, 6 velocity features, 6 displacement features and 18 angle features. They detected smoking behavior by dividing the activity into sub-activities: smoking while standing alone, smoking while walking, and smoking in a group. They achieved a recall of 81% and precision of 91% for recognizing smoking activity from above sub-activities by using conditional random fields (CRF) classifier. This work is important to us as it shows the effectiveness and importance of feature extraction from multi-inertial sensors on the improvement of the evaluation's accuracy.

One of the recent studies on smoking detection was conducted in 2016 [13]. The authors made a more comprehensive evaluation for smoking and other activities recognition such as smoking while standing, smoking while sitting, smoking while in a group conversation, smoking while walking, standing, sitting, walking, eating, drinking. They extracted six time- domain features from both accelerometer and gyroscope which are mean, standard deviation, minimum, maximum, kurtosis, and skew of the sample. They designed a two-layer hierarchical smoking detection algorithm (HLSDA) and achieved a very high precision as well as recall for smoking (90-97% F-measure). In the first layer, three commonly used classifier, decision tree, SVM and random forest, are used to identify the target activity. In the second layer, a single misclassified activity will be replaced with its immediate neighbor if it is different than its two neighbors on its left and right and they are same, or two consecutive misclassified activities will be replaced with their immediate neighbor if it's different than its neighbors and there are at least three same neighbors on their left and right. However, as the authors said by themselves, this algorithm may not be accurate if smoking and its confounding gestures being performed at the same time with the same hand.

According to our literature study, the previous works paid more attention on the concurrent activity while smoking (smoke while performing a concurrent activity such as smoking while standing, smoking while sitting, smoking while in a group conversation, and smoking while walking) rather than confounding gestures which may confused with smoking gestures. Therefore, in this study, we focus on recognizing smoking gesture with other confounding gestures including drinking, eating, scratching head, and biting nails. Moreover, similar to those areas where feature engineering has proved its merits in improving classification results (vehicle road condition recognition [19] and large-scale E-Commerce lucky money determination [20]), we explore an extended set of features for confounding gestures identification, and construct a solid relevant feature subset by using our ensemble feature selection machine. And then we analyze the influence on the evaluation performance by using different sensor solutions.

3 SYSTEM OVERVIEW

3.1 Data Collection

We collected a dataset for smoking and other similar activities including drinking, eating, scratching head, scratching nose, using phone and biting nails. 5 participants (age range: 30–35) took part in our data collection experiments and each activity was performed for 5–6 min. All participants performed these activities alone and in a controlled environment. The raw sensor data was collected by using on-board six axis accelerometer and gyroscope from Arduino 101 development board (figure 1) in the form of a wrist band. These six axis accelerometer and gyroscope are ranged to [-2G, 2G] and [-250 degrees, 250 degrees] with a calibration before data collection. All the raw data sampled at 20 Hz with minor variations based on onboard clock accuracy.

Figure 1: Arduino 101

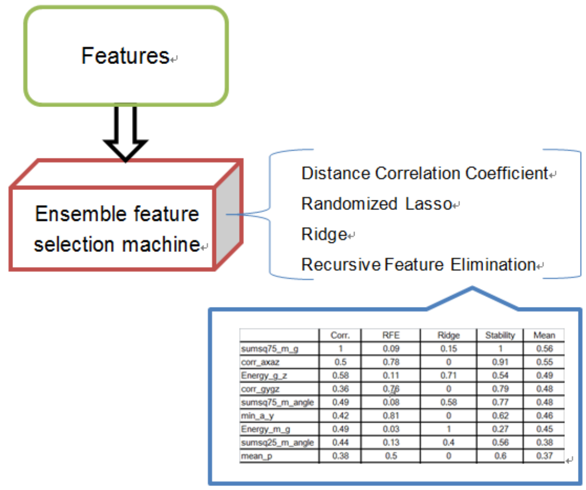
3.2 Feature Extraction and Selection

In this study, we extracted three groups of features which are orientation features-Acc, orientation features-Gyro and rotation features, and then analyzed the classification performance using different feature groups. The orientation features are computed using the Arduino 101's on-board 6-axis accelerometer and gyroscope measurements, in which the accelerometer is used to measure the orientation with respect to gravity and the gyroscope is used to measure the angular velocity. In addition, we calculated the quaternions from the accelerometer and gyroscope values by using the Madgwick filter algorithm which was developed by Sebastian Madgwick in 2010 [14]. These quaternions were then used to calculate the Euler angles pitch, yaw, and roll. We called the features computed from the Euler angles pitch, yaw, and roll as rotation features. Moreover, the magnitude of acceleration and the magnitude of angular velocity were utilized for the extraction of orientation features, and the magnitude of Euler angles was utilized for the extraction of rotation features. They were calculated as square-root of the sum of the squares of readings in each accelerometer axis, gyroscope axis, and pitch, yaw, and roll, respectively. Table 1 summarizes the features we extracted in this work, a total of 165 features were calculated for the orientation and rotation features. All these features were extracted over a range of window sizes of 10, 15, 20, 25, and 30s with 50% overlap. We choose a window size of 30 seconds, corresponding to 600 samples of raw data collection as it was one of the optimal window size in terms for the smoking detection in our experiment and this window size was also used in the previous works [13] [15].

A total of 165 features were extracted in this paper (55 orientation accelerometer features, 55 orientation gyroscope features and 55 rotation features), however not all of them are useful for the smoking detection due to the fact that these features can be relevant (the features that have influence on the evaluation) or irrelevant (the features have no effect on the evaluation). In order to find the relevant feature subset from the entire feature set, in order to increases the efficiency of the machine learning algorithm (prediction), we borrowed the idea from ensemble learning algorithm, and then designed an ensemble feature selection mechanism for the feature selection as shown in figure 2. Four commonly-used feature selection approaches were used in the ensemble feature selection procedure to reinforce the result, which are Distance Correlation Coefficient, Randomized Lasso, Ridge, and Recursive Feature Elimination. After the entire feature sets are fed to the ensemble feature selection process, each feature was graded by each of the above four approaches. We scaled each feature score to a range from 0 to 1 and then calculated the mean score for each feature. We sorted the mean scores and the unique features corresponding to the highest 15 mean scores were selected as the relevant features. The number of selected features may be larger than 15 as different features can have the same score.

Table 1: Features Used for Smoking Activity Classification

Name	Definition
mean value	$\left(\frac{1}{n}\sum_{1}^{n} a_i\right)$
root-mean-square	$\sqrt{\left(\frac{1}{n}\sum_{1}^{n} a_i{}^2\right)}$
standard deviation	$\sqrt{\frac{1}{n}\sum_{1}^{n}(a_i - mean)^2}$
max	$\max(a_i)$
min	$\min(a_i)$
max-min	$\max(a_i)\text{-}\min(a_i)$
signal magnitude area	$\frac{1}{t}\left(\int_0^t \lvert a(t)\rvert dt\right)$
correlation	$\frac{cov(x,y)}{\sigma_x\sigma_y}$
energy	$t\sum_{n=0}^{N-1} a^2[n]$
percentiles 25	$prctile(a_i, 25)$
percentiles 75	$prctile(a_i, 75)$
sum of data below percentile 25	$\sum (a_i)^2 \ ; a_i < prctile(ai, 25)$
sum of data below percentile 75	$\sum (a_i)^2 \ ; a_i < prctile(a_i, 75)$
interquartile range	$prctile(a_i, 75) - prctile(a_i, 25)$



	Corr.	RFE	Ridge	Stability	Mean
sumsq75_m_g	1	0.09	0.15	1	0.56
corr_axaz	0.5	0.78	0	0.91	0.55
Energy_g_z	0.58	0.11	0.71	0.54	0.49
corr_gygz	0.36	0.76	0	0.79	0.48
sumsq75_m_angle	0.49	0.08	0.58	0.77	0.48
min_a_y	0.42	0.81	0	0.62	0.46
Energy_m_g	0.49	0.03	1	0.27	0.45
sumsq25_m_angle	0.44	0.13	0.4	0.56	0.38
mean_p	0.38	0.5	0	0.6	0.37

Figure 2: ensemble feature selection machine

3.3 Data Interpretation

We implemented our algorithm in python 2.71, and we used Scikit-learn 0.17.1, which is an open source Python library to implement machine learning algorithms. We selected two classifiers in our evaluation: SVM and Random Forest, which are widely used in smoking detection research in the literature [12] [16] [17] [18]. We compared these classifiers' performance with their hyper-parameters optimization mode using Scikit-learn library [8]. For performance evaluation, a stratified 10 fold cross-validation without shuffling approach was applied. We computed the accuracy, precision, recall and F1-measure for each estimator. Here we present the F1-measure score, as it can be interpreted as a weighted average of the precision and recall [8].

4. Performance Evaluation

4.1 Recognition with Accelerometer

In this section, we present the results obtained using only the accelerometer measurements. A total number of 55 orientation features were extracted from the accelerometer reading. Figure 3 shows the feature selection result using our feature selection procedure. 18 features were selected from 55 features according to their mean score (unique features of top 15 scores were selected). In Table 2, results with SVM classifier are presented. All activities except eating activity showed low accuracies. The average F1-score for all activities is 63% while F1-scores for each activity are 48% for smoking, 53% for drinking, 86% for eating, 56% for biting nail and 77% for scratching head. Eating activity achieved the highest score in all of the 5 activities. Table 3 presents the results using random forest. Compared with SVM results given in Table 2, accuracies for all the activities are improved. The average F1-score for all activities using RF classifier is 84% while F1-scores for each activity are 78% for smoking, 63% for drinking, 96% for eating, 87% for biting nail and 93% for scratching head. In particular, the accuracy for the smoking activity and biting nail activity achieved a significant improvement; up to 30% increase using RF classifier. Again, the performance of eating activity recognition is still the best among these 5 similar activities. According to the results obtained from SVM and RF, although the SVM classifier performed relatively poor, the performance with only accelerometer measurements is still acceptable for these 5 similar activities recognition with an average of 84% accuracy using the random forest classifier.

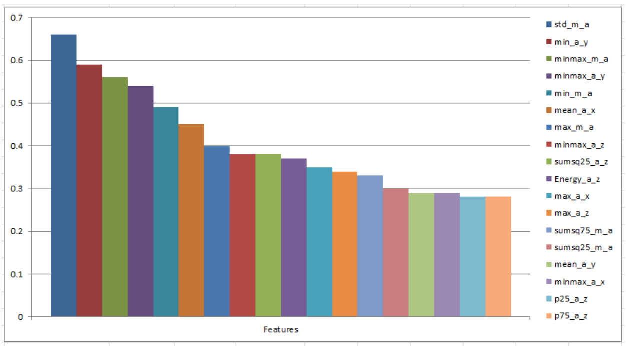


Figure 3: Selected Features with mean scores (Accelerometer)

Table 2: Accuracy Performance for Each Activity By SVM (Accelerometer)

Accuracy Performance By SVM			
Activities	Precision	Recall	F1-score
Smoking	0.47	0.5	0.48
Drinking	0.42	0.71	0.53
Eating	0.9	0.82	0.86
Biting Nail	0.62	0.5	0.56
Scratching Head	1	0.62	0.77
Avg/total	0.67	0.62	0.63

Table 3: Accuracy Performance for Each Activity By RF (Accelerometer)

Accuracy Performance By RF			
Activities	Precision	Recall	F1-score
Smoking	1	0.64	0.78
Drinking	0.56	0.71	0.63
Eating	0.92	1	0.96
Biting Nail	0.77	1	0.87
Scratching Head	1	0.88	0.93
Avg/total	0.87	0.84	0.84

4.2 Recognition with Gyroscope

In this section, we present the results obtained using only the gyroscope measurements. A total of 55 orientation features were extracted from the gyroscope readings, same as the number of features used with accelerometer. Figure 4 shows the feature selection result using our feature selection procedure. 25 features were selected from 55 features according to their mean score (unique features of top 15 scores were selected). In Table 4, results with SVM classifier are presented. All the activities recognized by SVM showed a less accuracy compared with the SVM results obtained using the accelerometer. The average F1-score for all activities is 33% while F1-scores for each activity are 56% for smoking, 38% for drinking, 58% for eating, 0% for biting nail and 0% for scratching head. Note that none of the biting nail and scratching head activity is recognized by using SVM classifier. Table 5 present the results with random forest. Compared with SVM results given in Table 4, accuracies for all the activities are improved. The average F1-score for all activities using RF classifier is 65% while F1-score for each activity is 69% for smoking, 67% for drinking, 70% for eating, 44% for biting nail and 78% for scratching head. Particularly, the accuracy for the biting nails and scratching head activity achieved a significant improvement using RF classifier from 0% to 44% and 78% respectively. According to the results represented from SVM and RF, using the gyroscope reading alone, the SVM classifier still performed poorly for these 5 similar activities. Although the Random Forest classifier did much better than SVM in this section, the performance with only gyroscope solution is significantly low, achieving an average accuracy of 65%.

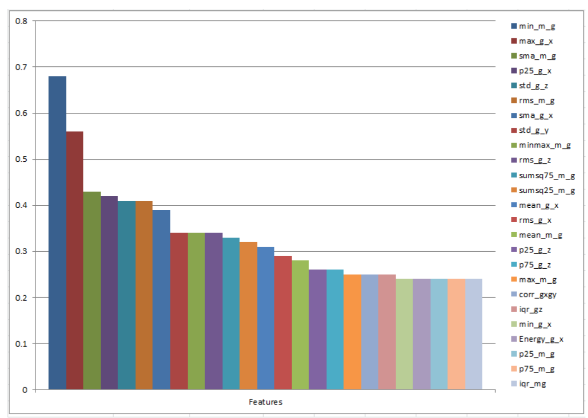


Figure 4: Selected Features with mean scores (Gyroscope)

Table 4: Accuracy Performance for Each Activity By SVM (Gyroscope)

Accuracy Performance By SVM			
Activities	Precision	Recall	F1-score
Smoking	0.43	0.64	0.56
Drinking	0.26	0.57	0.38
Eating	0.83	0.64	0.58
Biting Nail	0	0	0
Scratching Head	0	0	0
Avg/total	0.39	0.4	0.33

Table 5: Accuracy Performance for Each Activity By RF (Gyroscope)

Accuracy Performance By RF			
Activities	Precision	Recall	F1-score
Smoking	0.75	0.64	0.69
Drinking	0.55	0.86	0.67
Eating	0.78	0.64	0.7
Biting Nail	0.5	0.4	0.44
Scratching Head	0.7	0.88	0.78
Avg/total	0.67	0.66	0.65

4.3 Recognition with Both Accelerometer and Gyroscope

4.3.1 Recognition with rotation features only. In this section, we present the results obtained using only rotation features (features computed from the Euler angles pitch, yaw, and roll). A total of 55 rotation features were extracted from the Euler angles pitch, yaw, and roll. Figure 5 shows the feature selection result using our feature selection procedure. 21 features were selected from 55 features according to their mean score (unique features of top 15 scores were selected). In Table 6, results with SVM classifier are presented. All the activities recognized by SVM showed worse accuracy compared with the SVM results obtained using orientation features from accelerometer, but better than the SVM results obtained using the orientation features from gyroscope. The average F1-score for all activities is 55% while F1-scores for each activity are 64% for smoking, 37% for drinking, 86% for eating, 67% for biting nail and 0% for scratching head. Note that the scratching head activity was note recognized by the SVM classifier. Table 7 present the results with random forest. Compared with SVM results given in Table 6, accuracies for all the activities are improved. The average F1-score for all activities using RF classifier is 80% while F1-scores for each activity are 71% for smoking, 93% for drinking, 100% for eating, 95% for biting nail and 38% for scratching head. Particularly, the recognition of drinking, eating and biting nails activities achieved a high accuracy using RF classifier. According to the results represented from SVM and RF using rotation features alone, the SVM classifier still showed a poor performance in recognizing these 5 similar activities; however the performance with only using rotation features can still be acceptable for these 5 similar activities with average accuracy of 80% using random forest classifier.

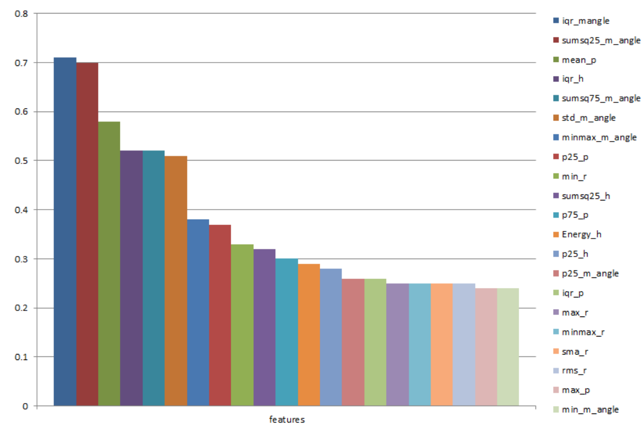


Figure 5: Selected Features with mean scores (rotation features)

Table 6: Accuracy Performance for Each Activity By SVM (rotation features)

Accuracy Performance By SVM			
Activities	Precision	Recall	F1-score
Smoking	0.73	0.57	0.64
Drinking	0.25	0.71	0.37
Eating	0.9	0.82	0.86
Biting Nail	1	0.5	0.67
Scratching Head	1	0	0
Avg/total	0.64	0.9	0.55

Table 7: Accuracy Performance for Each Activity By RF (rotation features)

Accuracy Performance By RF			
Activities	Precision	Recall	F1-score
Smoking	0.71	0.71	0.71
Drinking	0.88	1	0.93
Eating	1	1	1
Biting Nail	1	0.9	0.95
Scratching Head	0.38	0.38	0.38
Avg/total	0.8	0.8	0.8

4.3.2 Recognition with orientation and rotation features. In this section, we present the results obtained by using both orientation and rotation features. A total of 165 features were extracted from the accelerometer and gyroscope readings (55 orientation accelerometer features, 55 orientation gyroscope features and 55 rotation features). Figure 6 shows the feature selection result using our feature selection procedure. 19 features were selected from 165 features according to their mean score (unique features of top 15 scores were selected). In Table 8, results with SVM classifier are presented. In this section, the SVM classifier did a superior performance compared with the performance in seen in section 4.1, section 4.2 and section 4.3.1. The average F1-score for all activities achieved 78% while F1-scores for each activity are 90% for smoking, 56% for drinking, 80% for eating, 75% for biting nail and 80% for scratching head. Smoking activity achieved the highest score in all of the 5 activities. Table 9 presents the results with random forest (RF). Compared with SVM results given in Table 8, accuracies for all the activities have improved. The average F1-score for all activities using RF classifier is 88% while F1-scores for each activity are 97% for smoking, 75% for drinking, 78% for eating, 95% for biting nail and 88% for scratching head. Again, the performance of smoking activity recognition is the best among these 5 similar activities. Furthermore, for the RF classifier, the average F1-score of 88% is the highest score compared with the 84% in section 4.1, 65% in section 4.2 and 80% in section 4.3.1. According to the results represented from SVM and RF, although the average accuracy of 78% using SVM classifier is not as remarkable as the average accuracy of 88% achieved using RF, the performance with orientation and rotation features are reasonably high for these 5 similar activities using both the classifiers.

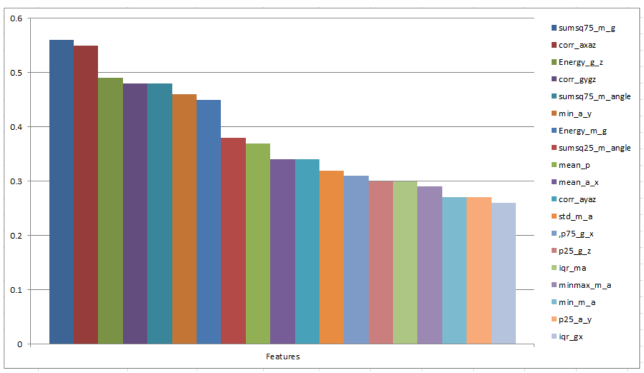


Figure 6: Selected Features with mean scores (Orientation and Rotation features)

Table 8: Accuracy Performance for Each Activity By SVM (Orientation and Rotation features)

Accuracy Performance By SVM			
Activities	Precision	Recall	F1-score
Smoking	0.82	1	0.9
Drinking	0.45	0.71	0.56
Eating	0.89	0.73	0.8
Biting Nail	1	0.6	0.75
Scratching Head	0.86	0.75	0.8
Avg/total	0.83	0.78	0.78

Table 9: Accuracy Performance for Each Activity By RF (Orientation and Rotation features)

Accuracy Performance By RF			
Activities	Precision	Recall	F1-score
Smoking	0.93	1	0.97
Drinking	0.67	0.86	0.75
Eating	1	0.64	0.78
Biting Nail	0.91	1	0.95
Scratching Head	0.88	0.88	0.88
Avg/total	0.9	0.88	0.88

4.4 Discussion

The first thing that can be noted from this work is, the solution of using only an accelerometer sensor to recognize the 5 highly related activities, namely smoking, drinking, eating, biting nail, and scratching head, can perform in a similar way to using both an accelerometer and gyroscope with random forest classifier. However, adopting the solution of using both the accelerometer and gyroscope features can improve the prediction result slightly using RF classifier, or strongly using SVM (see the F1-score in figure 7). In contrast, the solution of using the gyroscope only features without any other sensor measurements cannot provide good performance in detecting these activities.

The second thing that can been observed is, except using orientation features alone extracted from the gyroscope, both orientation features and rotation features are feasible to use in combination to recognize these 5 confounding gestures. However, it does not mean the orientation features extracted from the gyroscope are not important.

The third thing that we observed from our research is, although the performance of using the gyroscope features alone to predict these 5 highly related activities is not satisfactory, however, as shown in figure 8, these features extracted from the gyroscope could be highly relevant to the prediction. Figure 8 presents 19 features selected in section 4.3.2, in which 7 out of 19 features are contributed by the gyroscope, accounts for 36.80% of the whole feature set, and 3 features are contributed by the accelerometer and gyroscope combination, accounts for 15.8%. More than 50% of the features are directly or indirectly related to gyroscope features.

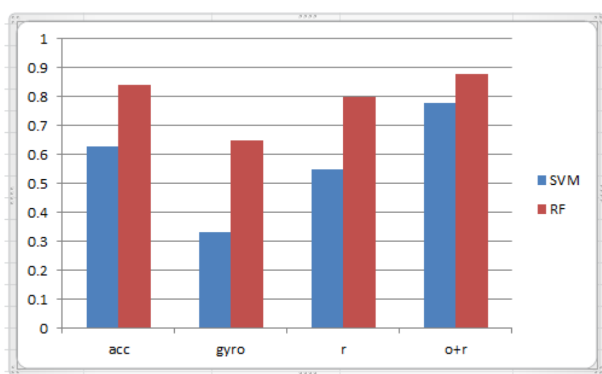


Figure 7: F1-score Summary

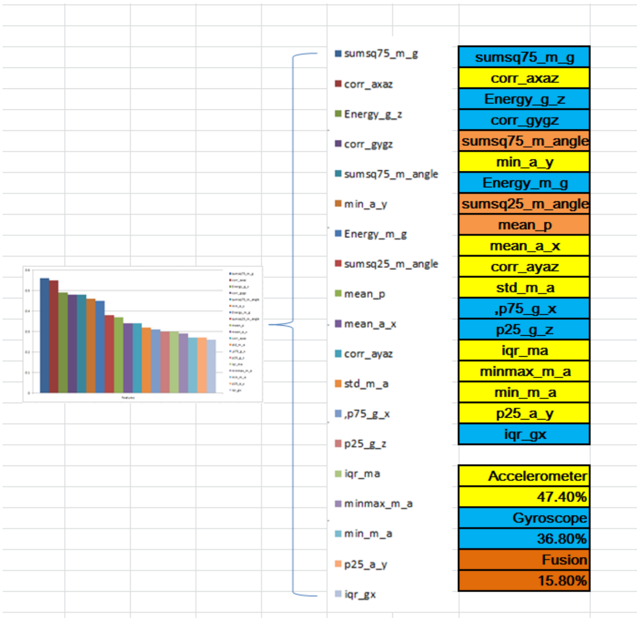


Figure 8: Percentage of Features occupied from an Accelerometer and Gyroscope

5. Conclusion and Future Work

In this paper, we evaluated the performance of confounding activity recognition using inertial sensors and particularly analyzed the performance with different sensor solutions. We categorized the feature set into two classes: orientation features and rotation features and we analyzed the performance using orientation and rotation features both alone and in combination. We collected our dataset from 5 participants with five confounding activity (smoking, drinking, eating, biting nail, and scratching head) and used SVM and random forest classifiers in our experiment. The results show that adopting the solution of using both an accelerometer and gyroscope achieve the highest accuracies than the solution with rotation features alone. The solution using accelerometer sensor alone are acceptable in some scenarios (such as recognition using RF), however, the solution of using the gyroscope alone is not acceptable due to poor performance. According to the average f1-score, random forest classifier achieved the highest performance in our experiment. As a future work, we plan to extend data collection to more participants and with more activities. In particular, we aim to include both concurrent activities and confounding activities together in our experiments.

6. References

[1] Okuyemi K S, Nollen N L, Ahluwalia J S. Interventions to facilitate smoking cessation[J]. American Family Physician, 2006, 74(2).

[2] B. Channel. Smoking - effects on your body. https://www.betterhealth.vic.gov.au/health/healthyliving/smoking-effects-on-your-body?viewAsPdf=true/, 2016. [Online; accessed 29-Dec-2016]

[3] W. H. Organization. Tobacco Fact Sheet. http://www.who.int/mediacentre/factsheets/fs339/en/, 2016. [Online; accessed 29-Dec-2016].

[4] D. of Health Australia. Tobacco Control. http://www.health.gov.au/tobacco/, 2016. [Online; accessed 29-Dec-2016].

[5] E. Sazonov, P. Lopez-Meyer, and S. Tiffany. A wearable sensor system for monitoring cigarette smoking. Journal of studies on alcohol and drugs, 74(6):956–964, 2013.

[6] J. L. Obermayer, W. T. Riley, O. Asif, and J. Jean-Mary. College smoking-cessation using cell phone text messaging. Journal of American College Health, 53(2):71–78, 2004.

[7] Bhandari B, Lu J C, Zheng X, et al. Non-invasive sensor based automated smoking activity detection[C]//Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE. IEEE, 2017: 845-848.

[8] Stitt J P, Kozlowski L T. A System for automatic quantification of cigarette smoking behavior[C]//Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE. 2006: 4771-4774.

[9] Ali A A, Hossain S M, Hovsepian K, et al. mPuff: automated detection of cigarette smoking puffs from respiration measurements[C]//Proceedings of the 11th international conference on Information Processing in Sensor Networks. ACM, 2012: 269-280.

[10] P. [10]M. Scholl and K. Van Laerhoven, "A feasibility study of wrist worn accelerometer based detection of smoking habits", in Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2012 Sixth International Conference on IEEE, 2012.

[11] A. Romeo, "Smoking Gesture Detection in Natural Settings", PhD thesis, WORCESTER POLYTECHNIC INSTITUTE, 2016.

[12] A. Parate, M.-C. Chiu, C. Chadowitz, D. Ganesan, and E. Kalogerakis,"Risq: Recognizing smoking gestures with inertial sensors on a wristband ," in Proceedings of the 12th annual international conference on Mobile systems, applications, and services ACM, 2014.

[13] M. Shoaib, H. Scholten, P. J. M. Havinga, O. D. Incel, "A hierarchical lazy smoking detection algorithm using smartwatch sensors", in 18th International Conference on e-Health Networking, Applications and Services (Healthcom),IEEE,2016

[14] Madgwick S. An efficient orientation filter for inertial and inertial/magnetic sensor arrays[J]. Report x-io and University of Bristol (UK), 2010, 25: 113-118.

[15] M. Shoaib, S. Bosch, H. Scholten, P. J. Havinga, and O. D. Incel,"Towards detection of bad habits by fusing smartphone and smartwatch sensors," in Pervasive Computing and Communication Workshops (Per-Com Workshops), 2015 IEEE International Conference on. IEEE, 2015, pp. 591–596.

[16] J. P. Varkey, D. Pompili, and T. A. Walls, "Human motion recognition using a wireless sensor-based wearable system," Personal and Ubiquitous Computing, vol. 16, no. 7, pp. 897–910, 2012

[17] N. Saleheen, A. A. Ali, S. M. Hossain, H. Sarker, S. Chatterjee, B. Marlin, E. Ertin, M. al'Absi, and S. Kumar, "puffmarker: a multi-sensor approach for pinpointing the timing of first lapse in smoking cessation," in Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. ACM, 2015, pp. 999–1010.

[18] Q. Tang, D. J. Vidrine, E. Crowder, and S. S. Intille, "Automated detection of puffing and smoking with wrist accelerometers," in Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2014.

[19] H.M. Xie, G.Y. Tian, G.Q. Du, Y. Huang, H.X. Chen, X. Zheng, and T. H. Luan. "A Hybrid Method Combining Markov Prediction and Fuzzy Classification For Driving Condition Recognition." *IEEE Transactions on Vehicular Technology*, 2018

[20] F., Min, C.M. Wong, H. Zhu, Y.J. Huang, Y.P. Li, X. Zheng, J. Wu, J. Yang, and C.M. Vong. "DAliM: Machine Learning Based Intelligent Lucky Money Determination for Large-Scale E-Commerce Businesses.", International Conference on Service-Oriented Computing (ICSOC), 2018