proceedings

Proceedings

Detection of Drinking via Inertial Sensor †

Volkan Senyurek [1,*], Masudul Imtiaz [1] and Naeemul Hassan [2]

1 Department of Electrical and Computer Engineering, University of Alabama, Tuscaloosa, AL 35487, USA; masudul4145@gmail.com
2 Department of Computer and Information Science, University of Mississippi, Oxford University of Mississippi, Oxford, MS 38677, USA; nhassan@olemiss.edu
* Correspondence: vsenyurek@eng.ua.net

† Presented at the 6th International Electronic Conference on Sensors and Applications, 15–30 November 2019; Available online: https://ecsa-6.sciforum.net//.

Published: 14 November 2019

Abstract: Alcohol addiction is the third leading lifestyle-related cause of death in the United States. There are not enough support tools for alcoholics who want to quit alcohol consumption. The detection of drinking in a free-living environment may improve the just-in-time adaptive intervention to this behavior. Traditional methods to detect alcohol consumption suffer from long response time that hinders prompt intervention and prevention. This paper proposes to employ inertial sensors to automatically detect drinking of alcohol by leveraging the hand gesture characteristics that are specific to drinking. Due to the lack of publicly available sensor dataset of alcohol drinking, this paper focused on the detection of general beverage drinking by exploiting the hand gestures. A public dataset containing seven daily activities (including hand-related activities such as eating, drinking, smoking, etc.) collected from 11 subjects in a controlled environment was adopted for this study. The detection model was developed using deep neural networks containing both convolutional and recurrent neural networks. The proposed approach achieved an F1-score accuracy of 87% in the Leave-One-Subject-Out (LOSO) cross-validation. We argue that contributions of this paper would be useful when an alcohol specific dataset becomes available.

1. Introduction

According to the WHO, alcohol use disorders, including alcohol dependence (AD) are the widespread public health problems those have resulted in millions of deaths per year. Alcohol usage not only causes severe adverse effects on the brain, liver, heart, and pancreas but also causes productivity loss, health care expenditures, motor vehicle accidents, crime, and other related costs [1,2]. Internationally, approximately 16% of drinkers engage in heavy episodic drinking (defined as consuming 60 g or more of alcohol per occasion), and the prevalence of AD exceeds 5% in many countries [3]. Detection of drinking in free-living could improve just-in-time interventions aimed at the control of over-drinking, and reducing alcohol-related diseases and injury. The most accurate measure of alcohol intoxication is a direct evaluation of blood alcohol concentration (BAC) [4,5]. But this is an invasive method and may not be practical for continuous monitoring. The other approach is the use of breath-analyzers that correlate to BAC through breath alcohol concentration (BrAC). Although this method is available for portable devices, it is not practical for continuous monitoring, and the measurement could be inaccurate due to external and internal factors such as humidity and temperature [6,7]. Recently, on-body electrochemical alcohol measurement from biofluids such as sweat and interstitial fluid (ISF) have drawn significant attention. The studies show that there is a close correlation between sweat alcohol concentrations and metabolites in ISF, and blood alcohol level [8,9]. Transdermal alcohol concentration (TAC) monitoring systems use excreted alcohol ($\tilde{1}$% of consumed alcohol) through the skin as vapor state insensible sweat [10]. SCAM^{TM}, WrisTS^{TM}, BACtrack and Proof^{TM} are general alcohol monitoring platforms that utilize transdermal approach. However, the response time of the TAC monitoring systems is one or more hours slower than the response time of BrAC [11] based systems. Additionally, TAC-based measurements may generate false detection from external alcohol-containing vapors (i.e. bar scenario, paint, etc.). In recent studies [12–14], gait analysis based-methods were used to detect alcohol consumption employing inertial sensors (accelerometer and gyroscope) of smartphones. These studies show that approximately twenty minutes after the initial alcohol consumption, alcohol that penetrates the blood-brain barrier affects the psychomotor performance [15]. Sun et al. [16] has investigated the emotional and physiological (heart rate, breathing rate, skin temperature, and activity) data related to drinking. Bae et al. [17] have used smartphone sensor, time and usage data for real-time detection of drinking in young adults. Day of the week, time of day, activity, screen duration, call duration, and typing speed were used as features.

In summary, the common drawback of previous studies was the long response time as those approaches were based on the measurement of biological and physical effects of alcohol. Instead of relying on time-demanding biological and physical effects of alcohol consumption, this paper proposes to detect alcohol drinking from hand gestures specific characteristic that are related to drinking. The current success of wearable sensors (especially inertial sensors) in the detection of hand-related activities indicates an inertial sensor placed on the wrist may objectively detect drinking events. However, there is no public dataset available on this regards to validate this hypothesis. Before conducting a human study involving an alcoholic person and thereby collect data, this study intended to test the proposed method on the detection of drinking generic beverages. Hence, a publicly available dataset that has several daily activities including drinking was adopted and a deep learning based model was developed to distinguish drinking from other daily activities.

The main contributions of this paper are:

* We investigated to what extent drinking can be automatically discerned from other daily activities.
* We proposed a Convolutional neural network (CNN)-Long Term Short Memory(LSTM) model which can distinguish drinking from eating and smoking activities in different postures.
* We proposed a model which could detect drinking within 100 seconds of alcohol consumption.

2. Related Work

According to our best knowledge, there is no hand gesture dataset for specifically alcoholic beverage drinking. Hence, the literatue review was limited to generic beverage consumption. In [18], the authors recognized the drinking activities by using random forest (RF) and support vector machine (SVM). In the same dataset, they reached F1-score of 0.78 for person-independent validation. In [19], the authors utilized the inertial signal from one wrist device. By using an SVM classifier, drinking with glass and cup were classified with an F1-score of 0.57 and 0.67 in LOSO validation, respectively. By adding a smart ring on the index finger, these scores reached to 0.93 and 0.92 respectively. In [20], five inertial sensors (attached to the wrists, upper arms and on the upper torso) were used for activity recognition. With a Hidden Markov Model (HMM), the drinking activity was classified with an F1-score of 0.85 in person-dependent validation.However, these two dataset does not contain any smoking activity which a critical hand related activity involves both hand and mouth.

3. Methodology

3.1. Dataset

We used a public dataset which was presented in [18]. This dataset contains seven different activities (standing, sitting, drinking while sitting, drinking while standing, eating, smoking while sitting and smoking while standing) performed by eleven participants. The data is collected by using a smartwatch application (LG Watch R, LG Watch Urbane, Sony Watch 3). 6D IMU data (accelerometer and gyroscope) were recorded at a sample rate of 50Hz. During the drinking activity, participants had a cup of coffee or tea while sitting or standing while been in a group conversation. The dataset contains approximately 45 hours of data. The start and end of each activity were labeled by participants who made a waving gesture.

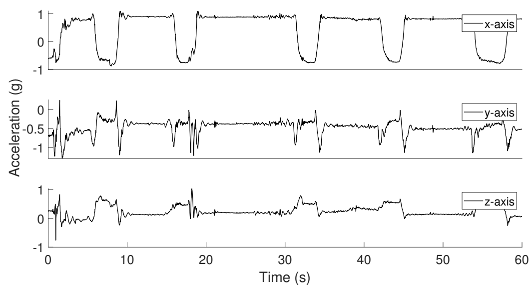

Figure 1. An example of the acceleration signals (x,y,z) during a drinking event.

3.2. Recognition Algorithm

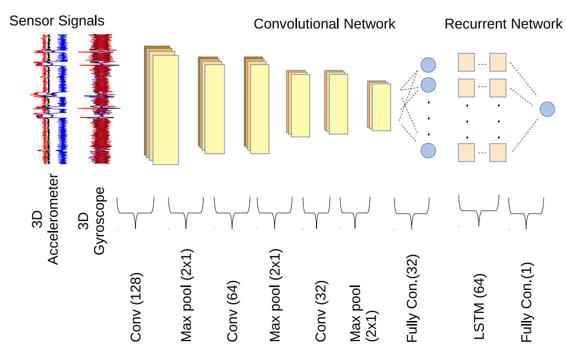

Figure 2. The overall pipeline of the proposed CNN LSTM model for drinking activity.

In this paper, we aim to detect drinking cycle in daily activities. Each drinking cycle can be modeled as a sequence of specific hand movement. Typically a drinking activity starts with the holding a cup or bottle, raising of the container to the mouth, then tilting the wrist up to drink. It continues withholding the cup at the mouth for a while. This cycle finishes with a downward motion of the hand from the mouth. Figure 1 illustrates a characteristic pattern in the responses of the accelerometers while a person drunk with the dominant hand.

Recently, deep learning gained popularity in many areas [21]. CNNs are a type of deep neural network with the ability to act as feature extractors. Such networks are able to learn multiple layers of feature hierarchy automatically. However, CNNs are not able to learn sequential correlations. On the other hand, LSTM recurrent neural networks are well suited to model temporal dynamics. The combination of

CNNs and LSTMs has offered state-of-art solutions for time series problems such as speech recognition and human activity recognition [22,23]. The idea in this work is to use the temporal-sequential dependency of hand gestures for drinking detection. The method presented in this study has two main neural network stages. In the first stage, a CNN act as a feature extractor. CNN layers can estimate the abstracted features of the hand gestures that take place during fixed-size overlapping windows of raw sensor streams. CNN part of the proposed model consists of 3 one-dimensional convolutional layers and a fully connected layer with 32 units. Each of the convolutional layers is followed by a batch normalization layer, rectified linear units, and max-pooling operation layer with a decimation factor of 2. To prevent overfitting during training, a dropout layer was used after fully connected layer with 50% dropout change. We used 128, 64, and 32 filters in convolutional layers. The filter size of all convolutional layers is half of their input data size, which corresponds to 10.24 seconds at a sampling rate of 50 Hz. The second step deals with the classification of window sequences as drinking or not, via the LSTM network. The proposed recurrent network consists of two consecutive LSTM layers with 64 hidden cells and one fully connected layer with a single neuron. The LSTM layer takes the CNN output as an input, with a sequence length of 10 samples and one sample step. 10 samples windows length correspond to 100 seconds. The network was trained by the dataset using the sliding window approach. The length of the window was 512 samples, with 50% overlap. In the training process, stochastic gradient descent with momentum (SGDM) optimizer was used with a learning rate of 10^{-3}, three epochs and a mini-batch size of 16. Figure 2 shows the overall architecture of proposed method.

The training was performed on a MATLAB 2018b environment installed on a windows 10 computer equipped with an Intel core i9 9th generation CPU with 32GB DDR4 RAM, and an NVIDIA GeForce RTX 2080Ti GPU with 12GB memory.

3.3. Evaluation

Leave-out-one-subject cross-validation was performed to evaluate the performance of the detection model. Having a dataset from 11 subjects, 10 were used as the training of the models, and the remaining one subject's data were used for testing. This procedure was repeated 11 times. Precision, recall, F1-score, accuracy, true positives (TPs), false positives (FPs), true negatives (TNs) and false negatives (FNs) were computed as performance metrics.

4. Results

Table 1 shows the confusion matrix of the proposed algorithm for drinking detection for each segment. Overall performance of the proposed method is summarized in the table 2. Figure 3 shows the false positive percentage of each activity in the total false positives.

Table 1. Confusion matrix of drinking detection algorithm

		Prediction	
		Drinking	**Non-drinking**
Activity	Drinking+Standing	6200	1130
	Drinking+Sitting	6375	957
	Eating	227	3443
	Sitting	64	3608
	Smoking+Standing	242	3433
	Smoking+Sitting	951	2718
	Standing	15	3652

Table 2. Overall performance metrics of drinking detection algorithm

Accuracy	Recall	Precision	F1-score
0.89	0.85	0.89	0.87

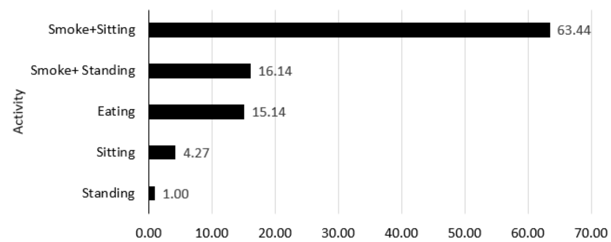

Figure 3. False positive percentage of each non-drinking activity.

5. Discussion

The performance scores of the proposed model show that drinking activity is easily detectable by using hand movements. According to the Figure 3, smoking while sitting is the most challenging task of the developed algorithm to separate from drinking. Drinking and smoking while sitting can be more similar to each other than other activities. Although eating activity contains similar hand gestures, the proposed algorithm differentiated drinking easily from eating. This study has some limitations that can be addressed in future studies. The dataset included only drinking of coffee and tea. In order to improve assessments and comparison, a more realistic dataset obtained from the drinker's natural environment representing a broader range of alcohol drinking activities should be used. Further, this study was based on data extracted exclusively from IMU sensors. An extended analysis using a combination of multiple sensors, for example, heart rate, Electromyogram for physiological signals during smoking activity, may enhance the precision and accuracy of the system. Not only individual sensor signals but also additional information such as time of day and day of the week may improve the detection performance.

In the study, the effect of the model parameters, the number of convolutional layers, feature map size, and kernel filter size were not investigated. In the future study, these parameters can be examined for improving the performance of the model. In addition, a fixed-size of sliding windows (512 samples) was used in the study. As a future study, the various size of the window would be investigated for optimized performance. The proposed algorithm can be implemented in a smartphone, smartwatch or a cloud-based server to continuously analyze real-time sensor signals for Just-In-Time Adaptive Intervention systems.

6. CONCLUSIONS

In this paper, we studied the detection of drinking activities by using inertial sensor signals as a preliminary study of alcohol beverage drinking recognition. Drinking activities were detected by using a CNN-LSTM based deep neural network with an average of 89% accuracy. This result is encouraging for an initial effort to aim alcohol drinking detection. In the future, we would like to gather more realistic data from drinkers who consume different type of alcohols in the real-life environment.

Funding: This work was not supported by any organization.

References

1. Alexis C Edwards, Sara L Lönn, Katherine J Karriker-Jaffe, Jan Sundquist, Kenneth S Kendler, and Kristina Sundquist. Time-specific and cumulative effects of exposure to parental externalizing behavior on risk for young adult alcohol use disorder. *Addictive behaviors* **2017**, *72*, 8–13.
2. Michael Roerecke and Jürgen Rehm. Alcohol use disorders and mortality: A systematic review and meta-analysis. *Addiction* **2013**, *108*, 1562–1578.
3. World Health Organization. *Management of Substance Abuse Unit.* Global status report on alcohol and health. World Health Organization, 2014.
4. Raye Z Litten and John P Allen. *Measuring alcohol consumption: Psychosocial and biochemical methods*. Springer Science & Business Media, 2012.
5. Phurpa Dema Thungon, Ankana Kakoti, Lightson Ngashangva, and Pranab Goswami. Advances in developing rapid, reliable and portable detection systems for alcohol. *Biosensors and Bioelectronics* **2017**, *97*, 83–99.
6. TM Worner and J Prabakaran. The accuracy of breath alcohol analysis using the breathalyzer. *Alcohol and Alcoholism* **1985**, *20*, 349–350.
7. Syed S Ali, Michael P Wilson, Edward M Castillo, Peter Witucki, Todd T Simmons, and Gary M Vilke. Common hand sanitizer may distort readings of breathalyzer tests in the absence of acute intoxication. *Academic Emergency Medicine* **2013**, *20*, 212–215.
8. Michael J Buono. Sweat ethanol concentrations are highly correlated with co-existing blood values in humans. *Experimental physiology* **1999**, *84*, 401–404.
9. U Fischer, K Rebrin, TV Woedtke, and P Abel. Clinical usefulness of the glucose concentration in the subcutaneous tissue-properties and pitfalls of electrochemical biosensors. *Hormone and metabolic research* **1994**, *26* 515–522.
10. Donald M Dougherty, Nora E Charles, Ashley Acheson, Samantha John, R Michael Furr, and Nathalie Hill-Kapturczak. Comparing the detection of transdermal and breath alcohol concentrations during periods of alcohol consumption ranging from moderate drinking to binge drinking. *Experimental and clinical psychopharmacology* **2012**, *20*, 373.
11. Tara E Karns-Wright, John D Roache, Nathalie Hill-Kapturczak, Yuanyuan Liang, Jillian Mullen, and Donald M Dougherty. Time delays in transdermal alcohol concentrations relative to breath alcohol concentrations. *Alcohol and alcoholism* **2016**, *52*, 35–41.
12. Brian Suffoletto, Pedram Gharani, Tammy Chung, and Hassan Karimi. Using phone sensors and an artificial neural network to detect gait changes during drinking episodes in the natural environment. *Gait & posture* **2018**, *60*, 116–121.
13. Muxi Qi. A comprehensive comparative performance evaluation of signal processing features in detecting alcohol consumption from gait data. 2016.
14. Zachary Arnold, Danielle LaRose, and Emmanuel Agu. Smartphone inference of alcohol consumption levels from gait. In Proceedings of the 2015 International Conference on Healthcare Informatics, pp. 417–426.
15. Shinichi Demura and Masanobu Uchiyama. Influence of moderate alcohol ingestion on gait. *Sport Sciences for Health* **2008**, *4*, 21–26.
16. P. Sun, N. M. Wergeles, C. Zhang, L. M. Guerdan, T. Trull, and Y. Shang. Ada - automatic detection of alcohol usage for mobile ambulatory assessment. In Proceedings of the 2016 IEEE International Conference on Smart Computing (SMARTCOMP), pp. 1–5.
17. Sangwon Bae, Tammy Chung, Denzil Ferreira, Anind K. Dey, and Brian Suffoletto. Mobile phone sensors and supervised machine learning to identify alcohol use events in young adults: Implications for just-in-time adaptive interventions. *Addictive Behaviors* **2018**, *83*, 42–47.
18. Muhammad Shoaib, Hans Scholten, Paul JM Havinga, and Ozlem Durmaz Incel. A hierarchical lazy smoking detection algorithm using smartwatch sensors. In Proceedings of the 2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom), pp. 1–6.
19. Alessandra Moschetti, Laura Fiorini, Dario Esposito, Paolo Dario, and Filippo Cavallo. Recognition of daily gestures with wearable inertial rings and bracelets. *Sensors* **2016**, *16*, 1341.

20. Holger Junker, Oliver Amft, Paul Lukowicz, and Gerhard Tröster. Gesture spotting with body-worn inertial sensors to detect user activities. *Pattern Recognition* **2008**, *41*, 2010–2024.
21. Li Deng. A tutorial survey of architectures, algorithms, and applications for deep learning. *APSIPA Transactions on Signal and Information Processing* **2014**, *3*.
22. Francisco Ordóñez and Daniel Roggen. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors* **2016**, *16*, 115.
23. Yu Zhang, William Chan, and Navdeep Jaitly. Very deep convolutional networks for end-to-end speech recognition. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 4845–4849.

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).