On Smartphone Sensability of Bi-Phasic User Intoxication Levels from Diverse Walk Types in Standardized Field Sobriety Tests

Ruojun Li, Ganesh Prasanna Balakrishnan, Jiaming Nie, Yu Li, Emmanuel Agu,
Michael Stein, Ana Abrantes, Debra Herman and Kristin Grimone.

Abstract— **Intoxicated driving causes 10,000 deaths annually. Smartphone sensing of user gait (walk) to identify intoxicated users in order to prevent drunk driving, have recently emerged. Such systems gather motion sensor (accelerometer and gyroscope) data from the users' smartphone as they walk and classify them using machine or deep learning. Standard Field Sobriety Tests (SFSTs) involve various types of walks designed to cause an intoxicated person to lose their balance. However, SFSTs were designed to make intoxication apparent to a trained law enforcement officer who manually proctors them. No prior work has explored which types of walk yields the most accurate results when assessed autonomously by a smartphone intoxicated gait assessment system. In this paper, we compare how accurately Long Short Term Memory (LSTM), Convolution Neural Network (CNN), Random Forest, Gradient Boosted Machines (GBM) and neural network classifiers are able to detect intoxication levels of drunk subjects who performed normal, walk-and-turn and standing on one foot SFST walks. We also compared the accuracy of intoxication detection on the ascending (increasing intoxication) vs descending (decreasing intoxication) limbs of drinking sessions (bi-phasic). We found smartphone intoxication sensing more accurate on the descending limb of the drinking episode and that intoxication detection on the normal walks of subjects were just as accurate as the SFSTs.**

Index Terms— **Blood Alcohol Concentration (BAC), Standard Field Sobriety Tests (SFST), Long Short Term Memory (LSTM), Convolution Neural Network (CNN), Gramian Angular Field (GAF)**

I. INTRODUCTION

About 11 percent of all pedestrians who die are hit by a drunk driver [1]. Drunk driving causes over 10,000 deaths (31 percent of overall driving fatalities) annually [1], making it important to identify drivers who are intoxicated over the legal Blood Alcohol Content (BAC) limit of 0.08%. While breathalyzers are the most accurate method of determining the intoxication levels of drivers, walk (gait) tests are also reliable and frequently used by law enforcement officers.

In about two-thirds of DUI incidents, when a law enforcement officer stops a suspected drunk driver, Standardized Field Sobriety Tests (SFST) are administered by the roadside to determine if the driver is too drunk to drive [2]. SFSTs involve various types of walks and tests that can be performed by sober people, but which drunk people find challenging or impossible to carry out (See figure 1). SFSTs are effective in identifying people who are too impaired to drive because alcohol affects the cerebellum, which coordinates voluntary movements such as walking, posture, balance, and coordination, causing them to lose their balance [3]. However, even in a very inebriated state (BAC $> 0.15\%$), most people's normal walking gait is only marginally altered by alcohol. SFSTs are specifically modified walks (e.g. narrowing subject's stance), which trigger a loss of balance at intoxication levels that are close to the legal driving limit (0.08%). Specific types of walks in the SFSTs include the walk-and-turn (tandem) gait test (See figure 1) and the standing on one foot tests that require balance and coordinated motion, making them challenging in an inebriated state.

Fig. 1: DUI Suspect performing the walk and turn test

The National Highway and Traffic Safety Association and a team of psychologists created the SFSTs [2] to identify drivers who are intoxicated over the legal limit. Subjects who fail these tests are usually arrested. Traditionally, SFSTs have been administered by a trained professional, who supervises the subject, corrects their mistakes in performing the tests and judges if the subject is over the limit. However, manual SFST administration and assessment is subjective and error prone [4]. To overcome these drawbacks and provide objective intoxication assessment, smartphone methods of sensing intoxicated users from their gait, have recently emerged. Such systems gather motion sensor (accelerometer and gyroscope) data from the users' smartphone as they walk and classify them using machine or deep learning.

The efficacy of SFSTs is well studied when manually administered by a human proctor as shown in figure 1. However, their efficacy in the context of autonomous sensing using a smartphone has never been previously studied. Specifically, it is currently unknown what accuracy can be achieved using smartphone sensing when subjects perform the different SFST walk types. In this paper, we compare the accuracy of smartphone intoxication sensing for subjects performing 1) a Normal/natural walk 2) the walk-and-turn SFST test, and 3) the standing on one foot SFST test. We believe a smartphone intoxication sensing system can detect sudden maneuvers by an intoxicated subject to regain imbalance (hopping, stretching out arms) triggered by SFSTs. Smartphone sensing of intoxication is possible because contemporary smartphones, which are now owned by 77% of Americans, have motion sensors (accelerometers and gyroscopes) that can be used to detect and classify subjects' movements as sober or intoxicated. They also have powerful, programmable processors for analyzing the motion data and running powerful machine or deep learning algorithms. Smartphone sensors are so precise that they can detect subtle gait changes due to intoxication, which may not be observable by a human proctor.

We also investigate a second research question related to intoxication detection in this paper. Alcohol drinking sessions (e.g. at a bar or party) typically consists of two phases (bi-phasic): *1) Ascending limb (increasing intoxication):* during which the subject's BAC increases up to a peak value for that session, and *2) Descending limb (decreasing intoxication):* during which the subject's BAC decreases from its peak value for the session till they become sober (0 BAC). We investigate whether smartphone intoxication sensing from gait is more accurate (or consistent) using gait data gathered during the ascending or descending drinking limbs. This question explores the possibly differential effects of alcohol on motor skills during the ascending and descending limbs.

Our Contributions: Using data gathered from motion sensors on a smartphone carried by intoxicated subjects in an IRB-approved controlled study, we investigate the following research questions in the context of smartphone sensing of intoxication from gait during SFSTs:

- We compare the accuracy of BAC regression for normal, SFSTs - walk-and-turn and standing on one foot using the LSTM, CNN with GAF encoding, random forest and Gradient Boosted Machines (GBM) machine learning regression algorithms.
- We compare the accuracy of BAC regression for the ascending and descending limbs for the normal, walk-and-turn and standing on one foot SFSTs using the LSTM, CNN with GAF encoding, random forest and GBM machine learning regression algorithms.

We found the intoxication detection more accurate on the descending limb of the drinking session, which is encouraging as most DUIs occur when subjects drive right after they have stopped drinking. We also found that the normal walks of subjects were just as accurate or slightly more accurate than other SFST walks, which implies that accurate, pervasive intoxication sensing while a drinker socializes at a party or bar could be possible.

II. RELATED WORK

Most smartphone intoxication sensing approaches involve sampling the motion sensors (accelerometer and gyroscope) on the users smartphone as they walk and classifying this data using machine or deep learning algorithms [5], [6], [7]. Traditional machine learning approaches classify hand-crafted features engineered to capture the dynamics of intoxicated human gait [5] [8]. Subjects in all prior work of this type only performed their natural/normal walk and no SFSTs (Walk-and-turn or standing on one foot) walks. Differences in classification on the up and down limbs of drinking were also not investigated.

Aiello and Agu [5] extracted statistical, time and frequency domain features including sway area, sway volume, kurtosis and skew from smartphone accelerometer and gyroscope data and classified the users intoxication into four BAC ranges: [0.00-0.08), [0.08-0.15), [0.15-0.25), [0.25+). Kao *et al* [6] proposed a smartphone sensing system that classifies users as drunk or sober (2 classes) using accelerometer features as inputs. Arnold et al. [9] classified sensor data from smartphone accelerometers and gyroscopes to predict the number of drinks a person had consumed but did not explicitly infer a BAC level. In Virtual Breathalyzer, Nassi *et al* [8] used FFT, statistical, histogram and gait features to train machine learning models such as Lasso, Gradient Boosted Machines, Regression Trees and AdaBoost, which were then used to regress the BrAC (Breath Alcohol Concentration) level of subjects. Gharani *et al* [7] utilized Artificial Neural Networks to detect the smartphone users BAC level. They extract features such as energy, mean and standard deviation from the raw accelerometer and gyroscope data, which they used to train a Multi-layer Perceptron (MLP) Bayesian Regularized Neural Network (BRNN) for regressing the users BAC level.

Numerous experiments have previously been conducted to measure the effect of alcohol consumption on motor performance[10] [11]. For a given BAC, the impairment is more significant in the ascending limb than the descending limb. In studies where drinkers performed a motor skill and an information-processing task at the same BAC, drinkers displayed impairment on both tasks during the ascending limb, but only motor skills are significantly impaired on the descending limb [12]. Smartphone intoxication sensing systems currently do not distinguish whether the subject is on the ascending or descending limb, often using a single machine learning model to predict intoxication on both limbs, which might be yielding wrong BAC estimates.

In summary, our work is novel in two main ways. First, prior smartphone intoxication sensing using machine or deep learning, utilized data gathered from participants who performed only their natural (normal) walk. Secondly, prior work also did not consider the bi-phasic nature (ascending and descending limbs) of alcohol consumption. We perform BAC regression on data gathered from intoxicated subjects who performed their natural/normal walking gait as well as the SFSTs (standing on one foot stand and the walk-and-turn), and compared accuracies. We also investigated how accurate intoxication sensing is on the ascending and descending limbs of alcohol consumption, determining whether a single machine learning model can predict BAC accurately on both limbs.

III. BACKGROUND

Human walking: involves shifting a person's Center of Mass (CM) back to front and also from side to side. When humans walk, the base of the support (distance between the feet) in the sagittal (back to front) and the lateral planes are wide. Consequently, the person's weight has a greater moment arm, which provides greater resistance to their shifting CM and stops them from toppling. When the base of the support is narrower, toppling is easier and the human tries to align their CM with the base of their support.

A. Standardized Field Sobriety Tests (SFST):

Ataxia, a degenerative disease of the muscular disease and alcohol both affect the cerebellum causing uncoordinated movements, slurred speech and difficulty with eye movements. Doctors have created tests to uncover subtle forms of ataxia and alcohol intoxication. These tests try to disrupt a person's natural balance. While a healthy person who has control over their muscle movement can overcome these disruptions and maintain their balance, an ataxia patient or intoxicated person fails the tests.

The Standardized Field Sobriety Tests (SFSTs) were developed by the National Highway Traffic Safety Administration (NHTSA) in collaboration with the Southern California Research Institute to provide a protocol that police officers can use to objectively identify drivers under the influence of drugs or alcohol. Specific SFST tests include the *walk-and-turn* test and the *standing on one foot* test. These are both divided attention tests, which emulate a motor cognition task similar to driving.

1) Walk and turn: The *walk-and-turn test* requires the person to concentrate, paying attention to oral instructions issued by the officers while also performing physical tasks prescribed in those instructions. For proper assessment during the walk and turn test, the subject is made to walk nine heel-to-toe steps along a line with their arms at their sides. After completing the nine steps, the subject turns and repeats nine heel-to-toe steps in the opposite direction.

The aerial footprint of the tandem gait is shown in figure 2 (left). L and R denote the left and right feet respectively and the numbers denote the steps. Figure 1 shows a driver performing the walk and turn test, under the supervision of law enforcement officers. The drastic reduction in the subjects's lateral support makes maintaining their balance during the walk and turn test challenging. When subjects are about to lose their balance, to decrease the moment of inertia they naturally spread their arms perpendicular to their body, failing the test. As such, only sober people or those without ataxia have adequate muscle coordination (controlled by the cerebellum) to successfully perform this walk, making it a simple, effective method of detecting intoxication and ataxia.

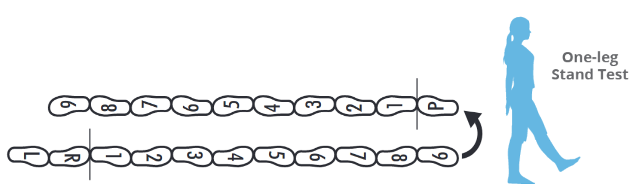


Fig. 2: Walk and turn test (left), One-leg stand (right) test

2) Standing on one foot: The *one-leg balancing test* was designed to take advantage of the fact that a persons ability to balance on one leg provides insights into their brains functional capability. During the one-leg stand test, starting from a standing position with their feet close together and their arms by their side, the subject is then required to raise one of their legs six inches off the ground, point their foot and then extend the raised foot till both arms are legs are straight. With their arms still by their sides, the subjects are required to count out loud from 1001 to 1030, or until they lose their balance (fail the test). Figure 2 (right) shows a subject performing the one-leg stand with their legs straight and their arms close to their body.

With just a single leg on the ground for support, the lateral width of the subject's base reduces to the width of a foot. As subjects are instructed to keep their arms by their sides, their body's moment of inertia is reduced, making them less able to resist forces/imbalances about their body's lateral or sagittal axis. Inebriated subjects tend to lose their balance. In order to regain their balance, subjects hop, spread out their arms or put their foot down. These compensatory movements are sudden and jerky maneuvers, which we believe can be detected by their smartphone's accelerometer and gyroscopes if placed on their body.

B. Biphasic nature of Alcohol Consumption:

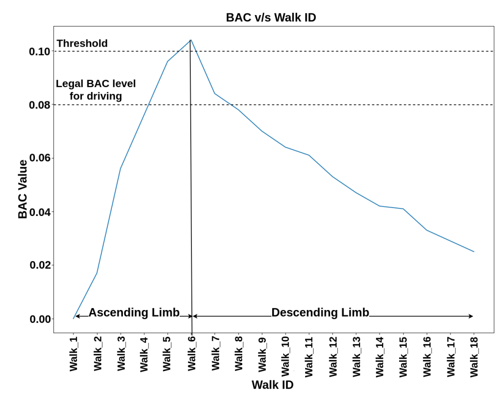


Fig. 3: BAC vs Walk ID.

The *ascending limb* of the Blood Alcohol Concentration (BAC) curve is the part where the subject's BAC increases from the sober state up till the peak BAC reached during the drinking session. The peak BAC value reached and the number of drinks required to reach this peak depends on the the subject's sex, how much blood the person has, and various other factors including medication, food consumption, rate of consumption (e.g. gulping vs. sipping). IRB-approved alcohol studies usually exclude subjects taking strong medications and study protocols try to ensure that subjects consume alcohol at the same rate and at the same intervals of time.

The *descending limb* starts from the peak BAC value reached during a drinking session, after the person stops consuming alcohol. The subjects BAC value decreases from this peak BAC value down to the sober state during the descending limb. The liver breaks down alcohol, a process that is regulated by an enzyme called alcohol dehydrogenase (ADH), which is secreted by the liver [13]. The descending limb is nearly linear with a slope that depends on the subject's metabolism and how much ADH they secrete.

IV. METHODOLOGY

A. Data Collection

Screening Recruitment for this study included Facebook ads and local posters. During the initial phone contact, callers completed the eligibility screen. Eligible individuals were scheduled for an individual 4-5 hour study session and given a list of instructions in advance (no water at least 2 hours prior to arrival, no food or other drinks at least 4 hours prior to arrival, no alcohol or marijuana at least 24 hours prior to arrival). A total of 65 participants provided consent and completed all study activities. The details of subjects who were a part of the study are listed below in table I.

TABLE I: Subjects Information

Category	Value
Gender Ratio (female:male)	8:5
Weight	74.646 kg $\pm$ 12.290 kg
Height	172.07 cm $\pm$ 7.10 cm
Age	30.787 years $\pm$ 12.010 years

Upon arrival at the study office, individuals completed Informed Consent and a brief eligibility check to confirm their age (21-65 y.o determined by checking a valid id), weight(85-230 lbs), walking disabilities, pregnancy status, absence of recent alcohol-related medical treatment, absence of recent drug use determined by urine toxicology and breath alcohol concentration (BAC) = 0. Urine toxicology was assessed using a self-contained test cup (Screeners Dip Drug Test with the Integrated Screeners Autosplit KO12B); pregnancy (for women) was assessed with an hCG dipstick test (Alere, San Diego, CA); BAC was assessed using a handheld Alcosensor IV Breathalyzer (Intoximeters Inc., St. Louis, MO). If found ineligible, participants were compensated for their time and did not continue with the study session. All participants received Hurricane Beer, 8.1 % alcohol content, starting with either 3 or 4 ounces of alcohol, depending on weight (men over 150 lbs and women over 160 lbs received 4 ounces initially) and subsequently varying the amount based on the rate of BAC ascent.

Ascending Limb: All eligible subjects started the study in a sober state and performed all the walk tests (normal walking gait, and the standing on one foot and the walk-and-turn SFST tests) at baseline. After a baseline walk, the participant was administered their first drink. The subjects waited 15 minutes after completing every drink before the next BAC was taken. Immediately after, and at approximately 5 and 15 minutes after completing the drink, the subject was instructed to rinse their mouth with water and spit, to purge any remaining mouth alcohol. After each breathalyzing (BAC recorded) the subject is instructed to perform all walk types until their BAC reached a peak value of 0.1 (ascending limb). The walks were performed along a corridor at Butler Hospital, where a 75-foot tape line was placed in a straight line. For the normal/natural walk, participants started their walk in the middle of the line, and were instructed to walk at their normal pace to the end of the line, turn, walk to the other end, turn, and return to their starting point. Each walk lasted 45-60 seconds. At the end of this walk, the SFSTs were administered. The walk-and-turn SFST was administered at all the walks; the standing on one foot was administered when the BAC was .02, .04, .06, .08 and .10. Study staff were trained in the field sobriety tests by watching YouTube videos for each task, having a local policeman meet with and validate this training, and having staff observe each other to develop administration consistency.

As different subjects get drunk at different rates, the actual number of drinks required to reach a BAC of 0.1 varied from subject to subject, however all participants were administered alcohol in the same manner (3-4 ounces at a time, a BAC check 15 minutes later, and another 3-4 ounces until a BAC of 0.1 was reached). During all walks, a Google Pixel smartphone running an Android sensor data gathering app was used to collect accelerometer, gyroscope and compass data continuously. Before the subjects walked, the data gathering app was turned on and the smartphone was placed in a harness strapped to the subject's hip, close to their center of mass. The sensors have a sampling rate of 200 Hz, with a resolution of 0.0012 rad/s and 0.0024 m/s^2 for the gyroscope and accelerometer respectively. More details regarding the sensor characteristics can be found in [14]

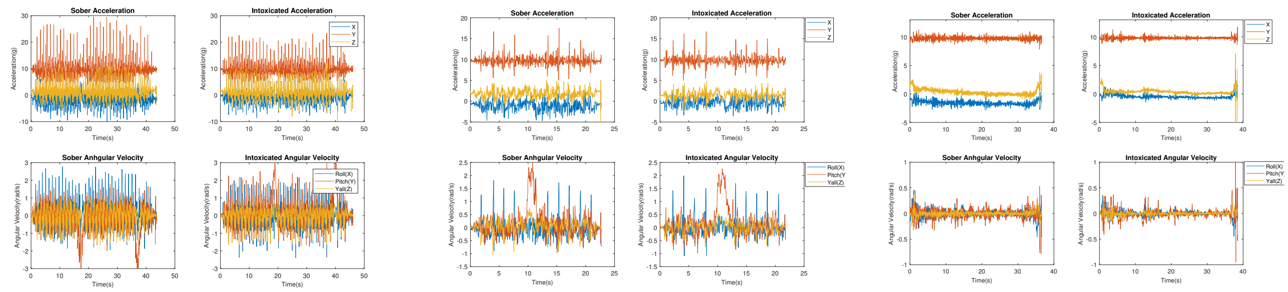
Descending Limb: Once subjects reached the peak BAC value of 0.1, participants were allowed to eat as they wanted. Every 7 minutes they were breathalyzed and instructed to perform all walk types, but no more alcohol was given (descending limb). Data for the descending limb was collected until the subject reached a level of sobriety that does not hinder their cognitive and motor skills (approximately 0.2). This point was determined by the battery of SFSTs (standing on one foot, walk-and-turn) that are conducted every seven minutes and the BAC value recorded.

As waiting for the subject BAC to return to 0 may take between 4 hours to a day, subjects were discharged when their motor cognition was back to normal. All participants were provided cab rides to and from the study session; no study participant was allowed to drive themselves. The data corresponding to a single subject when sober and when inebriated is shown in figure 4. Figure 4a is a plot of data collected when a subject executed their normal walking gait, figure 4b corresponds to the walk-and-turn test and figure 4c corresponds to the standing on one foot test. The graphs show a relatively greater difference in the signals corresponding to the One-foot stand test.

B. Machine/Deep Learning Intoxication Regression Models

We explored both machine learning models using hand-crafted features as well as neural networks (deep learning) on raw sensor data.

1) Machine Learning (Random Forest and Gradient Boosted Machine (GBM): We extend Aiello and Agu's [5] work on classifying manually engineered accelerometer and gyroscope features that capture gait dynamics and imbalance caused by alcohol. We include additional features utilized by Nassi *et al* [8] in their virtual breathalyzer research.



(a) Normal walking gait

(b) Walk-and-turn test

(c) One-foot Stand test

Fig. 4: Accelerometer and Gyroscope signals

The features we extracted included postural sway features, zero crossing rate, signal magnitude area, energy, skewness, kurtosis, range, standard deviation and frequency peaks.

We computed a total of 95 features for each 8-second segment using data from 65 subjects for each walk type. We then trained Random forest and Gradient Boosted Machine (GBM) regressors to predict the BAC value for each walk. We utilized the XGBoost implementation of GBMs [15]. Overfitting is a common problem when using decision trees, which random forests and GBMs can reduce. To further mitigate overfitting we performed feature selection and dimensionality reduction on our 95 features using the backward elimination and PCA algorithms. Thus, we were able to select and use only the most relevant features for BAC regression from smartphone sensor data.

Optimal model parameters that yield the lowest Root Mean Square regression Error (RMSE) across folds were selected using an exhaustive grid search. We trained the models about 30 times on different folds of the train and test data splits to ensure consistency.

2) CNN + GAF: Convolutional Neural Networks (CNNs) have performed well on many object detection and image analysis tasks and are invariant to translation, rotation, and scale. We converted the raw sensor data into an image-based representation, which we analyzed using a CNN. The most common way to visualize a time-series signal as an image is to create a color-coded spectogram by converting the signal into the Fourier domain. We instead encoded segments of the accelerometer and gyroscope signals as a two-dimensional image using the Gramian Angular field (GAF) [16][17]. GAF is a technique used to encode time-series data as images, which has been found to yield more accurate results than Fourier transforms when using CNNs [16][17]. The time-series data is normalized and the cosine angles are calculated. The radius of a given angle is the corresponding time stamp of the data. The time-series data is converted into a polar representation and then into a Markov transition matrix as discussed in [16][17]. Fig.5 shows the polar and corresponding GAF representations of a signal.

We utilize a shallow CNN architecture (figure 7)to classify the GAF-encoded accelerometer and gyroscope data in order to preserve spatial information of the small GAF feature map while obtaining a good receptive field for contextual understanding of the segment.

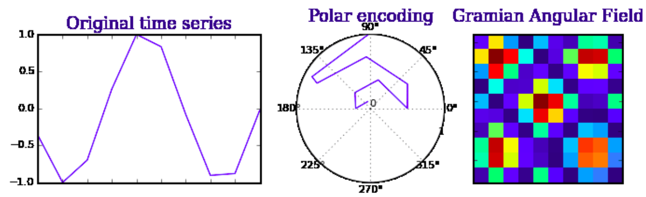


Fig. 5: The GAF representation of a signal.

3) LSTM: Learning long-term data dependencies is intractable using Recurrent Neural Networks (RNN) [18] [19] . Hochreiter *et al* developed Long Short-term Memory (LSTM) networks [20] that can learn long term dependencies quite effectively. Similar to RNNs, LSTM networks have recurring LSTM cells that pass information through time and have a memory about previous states. LSTM has gates that can control the amount of information that each cell adds (input gate) to the memory (cell state) and also the amount of information each cell's outputs (output gate).

In this paper we use a variant of the classic LSTM cell developed by Zaremba *et al* [21] that has a dropout mechanism controlled by a forget gate that controls how much information to forget. The architecture of our LSTM is shown in figure 8. The green cells shown in the network are LSTM cells developed by Zaremba *et al* [21]. X_t denotes the input vector to the LSTM cell at timestep t, C_t is the cell state at timestep t that acts like a conveyor belt (from which gates add and remove information) and h_t is the output of the cell at timestep t that is passed on to the next cell. FC denotes dense layers, and ReLU is the Rectified Linear Unit activation.

We trained the deep learning networks to minimize the objective function L shown below:

$$L = \sqrt{\sum_i (\hat{y}_i - y_i)^2 / N} \tag{1}$$

V. EVALUATION

We used the Root Mean Square Error (RMSE) as our metric for determining the performance of the machine and deep learning regression algorithms evaluated. First we divided all sensor data collected such that each walk type was a separate dataset. We further sub-divide each of the walk types into three datasets - up (ascending limb data only),

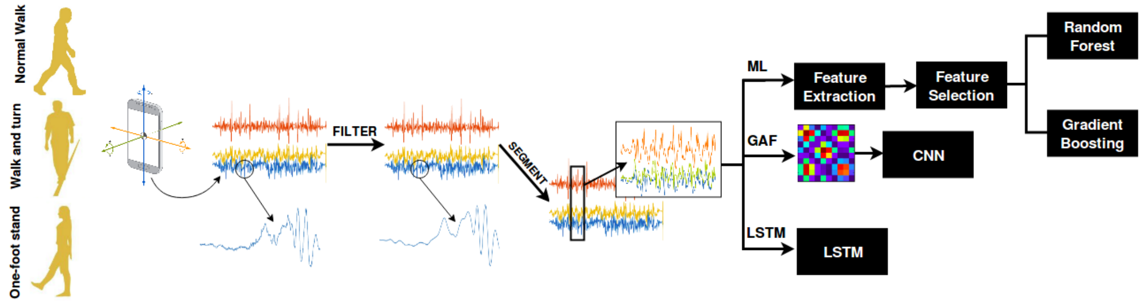


Fig. 6: The data collection and testing framework

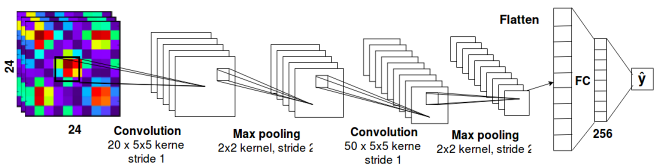


Fig. 7: The CNN architecture.

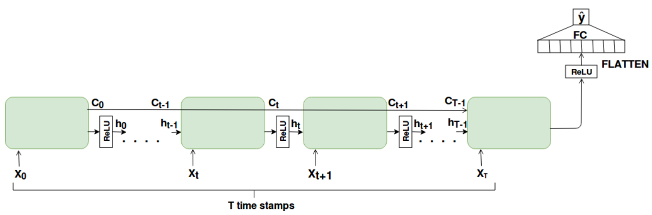


Fig. 8: The LSTM architecture.

down (descending limb data only) and all (a combination of the ascending and descending limbs).

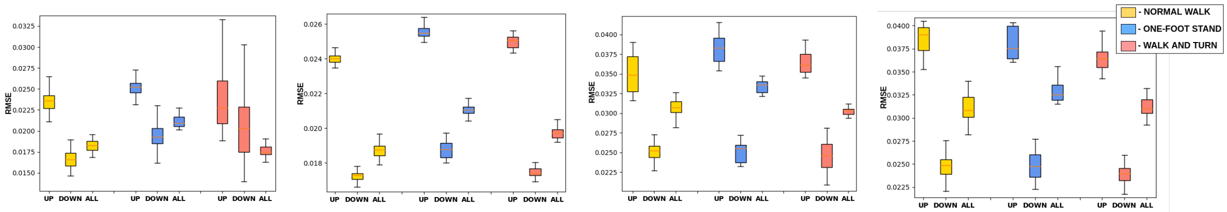
For the Random Forest (RF) and Gradient Boosted Machines (GBM) we segmented the each signal into 8-second segments with 50% of each segment overlapping the next segment. We then extracted features for each segment. We used backward elimination to eliminate features with low importance ($pvalue < 0.05$), normalized the data and trained RF and GBM on the normalized data.

For the deep learning models, to obtain adequate data we segmented the raw sensor data into 4-second windows with 50% overlap and converted the segments into the GAF matrix for every axis of the sensors. We used this feature map to regress the BAC using the CNN architecture shown in figure 7. We also trained the LSTM shown in figure 8, using the raw segments to reduce the L2 loss between the predicted and the true BAC. The results of our models on the individual datasets can be found in table II.

VI. DISCUSSION

Table II shows our test RMSE obtained for each of the limbs for each walk type. We used subject-level splitting such that all entire data gathered for a given subject was part of either the train or the test set but not both. We make the following observations from our results:

1) *BAC Regression on the Descending Limb is more accurate:* the regression models gave a lower RMSE on the descending limb when compared to the ascending limb or all the data. This is very important when using smartphone sensing to detect intoxicated drivers in order to prevent Driving Under the Influence (DUI) incidents. In most cases, people usually decide to drive after they have terminated their alcohol consumption session. Marczinski and Fillmore [22] examined the willingness of subjects to drive in the inebriated state in the ascending and descending limb and found out that for a given BAC, subjects in the descending limb are significantly more willing to drive. Amlung et al. [23] attributed this to the reduction in perceived danger in the descending limb compared to the ascending limb.
2) *Intoxication sensing using normal walk is as accurate as the SFST walks:* From table II we can conclude that BAC sensing using data gathered during the Normal Walk is as good or slightly better than the results obtained from data gathered during the SFSTs. Even though SFSTs are preferred and very effective when manually proctored by a trained law enforcing officer, normal walks are just as good for smartphone sensing. This implies that pervasive, continuous, smartphone intoxication sensing can be done as a drinker walks about their drinking venue (e.g. to the bathrooms, to interact socially or get more drinks). The precision of accelerometers and gyroscopes on modern smartphones combined with the power of deep learning yield results on normal walks that are just as precise and robust as SFSTs. We envision that future smartphone intoxication sensing systems can possibly run as a background process on the phone as the person performs their normal daily activities or interacts socially.
3) *LSTM performs the best:* Random Forest and Gradient Boosting algorithms require manual feature engineering and for a large dataset, engineering features that are able to capture variances in the data corresponding to change in BAC is challenging. Thus using handcrafted features that have been proven to work well for alcohol intoxication, gives us a very high error as shown in II. Using CNNs to predict based on the Markovian Transition Matrices of the data gives more accurate results as the features extracted are tuned using the error between the predicted BAC and the true BAC. The LSTM performs the best as the network learns robust representations of motion sensor data and temporal transitions that are predictive of BAC.



(a) LSTM (b) CNN with GAF (c) Gradient Boost (d) Random Forest

Fig. 9: Experimental Results on the different walk-types and limbs

Model	Dataset								
	Normal Walk			One-foot Stand			Walk and turn		
	Up	Down	All	Up	Down	All	Up	Down	All
LSTM	0.0237	0.0170	0.0180	0.0257	0.0200	0.0211	0.0256	0.0210	0.0177
CNN + GAF	0.0240	0.0173	0.0188	0.0250	0.0175	0.0198	0.0225	0.0187	0.0211
Random Forest	0.0385	0.0249	0.0311	0.0380	0.0248	0.0329	0.0367	0.0239	0.0311
Gradient Boosting	0.0350	0.0249	0.0306	0.0381	0.0250	0.0333	0.0365	0.0245	0.0302

TABLE II: RMSE of models on different datasets (regression)

VII. CONCLUSION AND FUTURE WORK

In this paper, we compared smartphone sensing of intoxicated subjects from data gathered as they performed their normal walk as well as the walk and turn and standing on one foot SFST tests. We also compared the accuracy of intoxication detection on the ascending and descending limbs of drinking episodes. We found the intoxication detection more accurate on the descending limb of the drinking episode, which is encouraging as most DUIs occur after subjects drive right after they have terminated drinking. We also found that the normal walks of subjects were just as accurate or slightly more accurate than the SFST walks, which implies that accurate, pervasive intoxication sensing while a drinker socializes at a party or bar could be possible. In future work, we hope to conduct larger scale evaluations in real life social drinking settings. Prior work found that the impairment of motor skills due to intoxication were generally more pronounced than impairment of cognitive skills in both the ascending and descending limbs [12]. This implies that even though an intoxicated person is able to walk steadily and is predicted to have a low BAC value by an intoxication sensing system, their mind may not be prepared to coordinate a dynamic, cognitive activity such as driving. Future work could also explore differences in the impairment of motor and cognitive skills based on BAC values sensed by an smartphone intoxication sensing system.

ACKNOWLEDGMENT

This work was supported in part by NIH/NIAAA grant number 1R21AA025193-01.

REFERENCES

[1] National Center for Statistics and Analysis, “Alcohol impaired driving (traffic safety facts 2017 data),” 2018. [Online]. Available: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812630

[2] “DWI detection and standardized field sobriety testing (sfst) refresher,” 10 2015. [Online]. Available: https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/sfst_ig_refresher_manual.pdf

[3] E. V. Sullivan, M. J. Rosenbloom, A. Deshmukh, J. E. Desmond, and A. Pfefferbaum, “Alcohol and the cerebellum : Effects on balance, motor coordination, and cognition.” [Online]. Available: https://pubs.niaaa.nih.gov/publications/ahrw19-2/138-141.pdf

[4] S. Cole and R. H. Nowaczyk, “Field sobriety tests: Are they designed for failure?” *Perceptual and Motor Skills*, vol. 79, no. 1, 1994.

[5] C. Aiello and E. Agu, “Investigating postural sway features, normalization and personalization in detecting blood alcohol levels of smartphone users,” in *In Proc Wireless Health*, 2016, pp. 1–8.

[6] C. H.-L. Kao, B.-J. Ho, A. C. Lin, and H.-H. Chu, “Phone-based gait analysis to detect alcohol usage,” pp. 661–662, 09 2012.

[7] P. Gharani, B. Suffoletto, T. Chung, and H. A. Karimi, “An artificial neural network for movement pattern analysis to estimate blood alcohol content level,” in *Sensors*, 2017.

[8] B. Nassi, L. Rokach, and Y. Elovici, “Virtual breathalyzer,” *CoRR*, vol. abs/1612.05083, 12 2016.

[9] Z. Arnold, D. Larose, and E. Agu, “Smartphone inference of alcohol consumption levels from gait,” in *ICHI*, 2015, pp. 417–426.

[10] E. Mellanby, “Alcohol : its absorption into and disappearance from the blood under different conditions.”

[11] M. Vogel-Sprott and M. T. Fillmore, “Impairment and recovery under repeated doses of alcohol: Effects of response-outcomes,” *Pharmacology Biochemistry and Behavior*, vol. 45, pp. 59–63, 1993.

[12] J. N Fogarty and M. Vogel-Sprott, “Cognitive processes and motor skills differ in sensitivity to alcohol impairment,” *J. Alcohol*, 2002.

[13] A. Paton, “Alcohol in the body,” vol. 330, 2005.

[14] *LSM6DSM - iNEMO inertial module: always-on 3D accelerometer and 3D gyroscope*, STMicroelectronics, 2017, rev. 7. [Online]. Available: https://www.st.com/resource/en/datasheet/lsm6dsm.pdf

[15] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” in *Proc ACM SIGKDD*. ACM, 2016, pp. 785–794.

[16] Z. Wang and T. Oates, “Encoding time series as images for visual inspection and classification using tiled conv. neural networks,” 2014.

[17] Z. Wang1 and T. Oates, “Imaging time-series to improve classification and imputation,” 2015.

[18] S. Hochreiter, Y. Bengio, P. Frasconi, and J. Schmidhuber, “Gradient flow in recurrent nets: Difficulty learning long-term depend.” 2001.

[19] R. Pascanu, T. Mikolov, and Y. Bengio, “On the difficulty of training recurrent neural networks,” *Intl. Conf Machine Learning*, 2013.

[20] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, pp. 1735–80, 12 1997.

[21] W. Zaremba, “Recurrent neural network regularization,” 2014.

[22] C. A Marczinski and M. Fillmore, “Acute alcohol tolerance on subjective intoxication and simulated driving performance in binge drinkers,” *Psychology of Addictive Behaviors*, vol. 23, 2009.

[23] M. T. Amlung, D. H. Morris, and D. M. McCarthy, “Effects of acute alcohol tolerance on perceptions of danger and willingness to drive after drinking,” *Psychopharmacology*, pp. 4271–4279, Nov 2014.