Detecting Blood Alcohol Level from Smartphone-Based Gait Data using Neural Networks

Ruojun Li [1] Jiaming Nie [1] Ganesh Prasanna [1] Emmanuel O. Agu [1] Michael Stein, MD[2]
Ana Abrantes, PhD[2] Debra Herman [2] Kristin Grimone [2]

[1]Computer Science, Worcester Polytechnic Institute

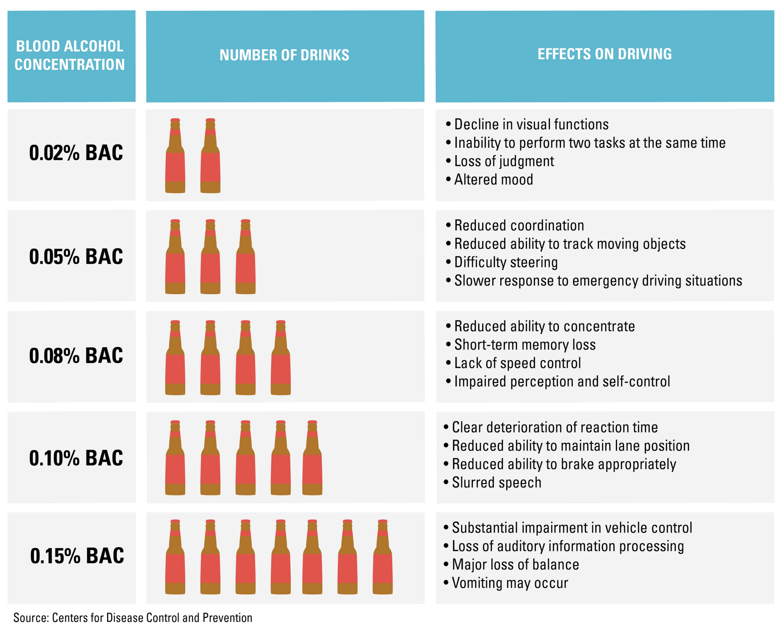
[2]Butler Hospital, Brown University

Objective

This study aims to detect BAC using deep learning algorithms from human motion data obtained from smartphone sensors (accelerometer and gyroscope). It is envisioned that the deep learning model will be deployed in a mobile application to notify drivers who are intoxicated over the legal drinking limit, preventing DUIs.

Motivation

Excessive alcohol consumption and Driving under the influence (DUIs) cause over 80,000 deaths annually. Traditional methods of detecting DUIs include using a breathalyzer to determine a person's Blood Alcohol Content(BAC) and computing the BAC using Widmark's equation. Passive methods to unobtrusively detect the user's BAC and notify them when they are over the limit are attractive. We investigate using deep learning algorithms to analyze smartphone sensor data to predict the BAC of smartphone users from their gait data. Specifically, we compare several deep learning architectures (MLP, LSTM, CNN) with traditional machine learning methods (Random Forest) and report on our findings.



BLOOD ALCOHOL CONCENTRATION	NUMBER OF DRINKS	EFFECTS ON DRIVING
0.02% BAC		• Decline in visual functions • Inability to perform two tasks at the same time • Loss of judgment • Altered mood
0.05% BAC		• Reduced coordination • Reduced ability to track moving objects • Difficulty steering • Slower response to emergency driving situations
0.08% BAC		• Reduced ability to concentrate • Short-term memory loss • Lack of speed control • Impaired perception and self-control
0.10% BAC		• Clear deterioration of reaction time • Reduced ability to maintain lane position • Reduced ability to brake appropriately • Slurred speech
0.15% BAC		• Substantial impairment in vehicle control • Loss of auditory information processing • Major loss of balance • Vomiting may occur

Source: Centers for Disease Control and Prevention

Figure 1: BAC Level and Effects.

Related Work

Kao et al.[1] explored alcohol-induced gait anomaly detection by analyzing smartphone accelerometer signals. Arnold *et al* [2] found that Random Forest classifier types outperform Naive Bayes, decision trees and SVMs for this task. Most prior work uses traditional machine learning using hand-crafted time, frequency and energy features that have performed well on the related Human Activity Recognition task[3]. However, such methods do not always generalize well [4] and deep learning has outperformed machine learning for many similar tasks.

We investigate various deep learning approaches including Multi-Layer Perceptron (MLP), Long Short-Term Memory-based models (LSTM) and Convolutional Neural Networks (CNN) on raw sensor data.

Methods

Data Collection: Smartphone gait (accelerometer and gyroscope) data was gathered from 70 subjects at Butler Hospital, while they walked at various intoxication levels after being breathalyzed to gather ground truth BAC data. [5]

Data Preprocessing: The accelerometer and gyroscope data were pre-processed in several steps: filter smoothing, overlapping and segmenting. A visual representation of the overlapping is shown in figure 2.

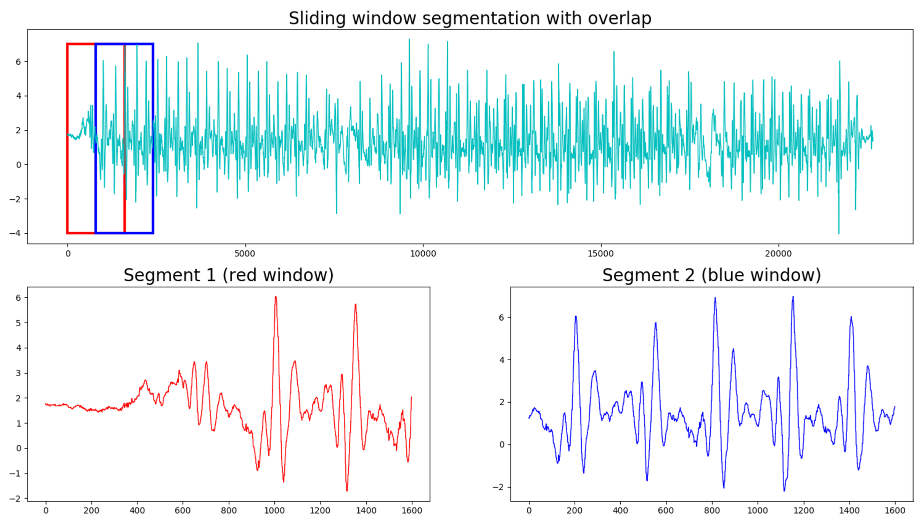


Figure 2: Overlapping of Data Segments

Following figure 3 shows a summary of our experimental methodology.

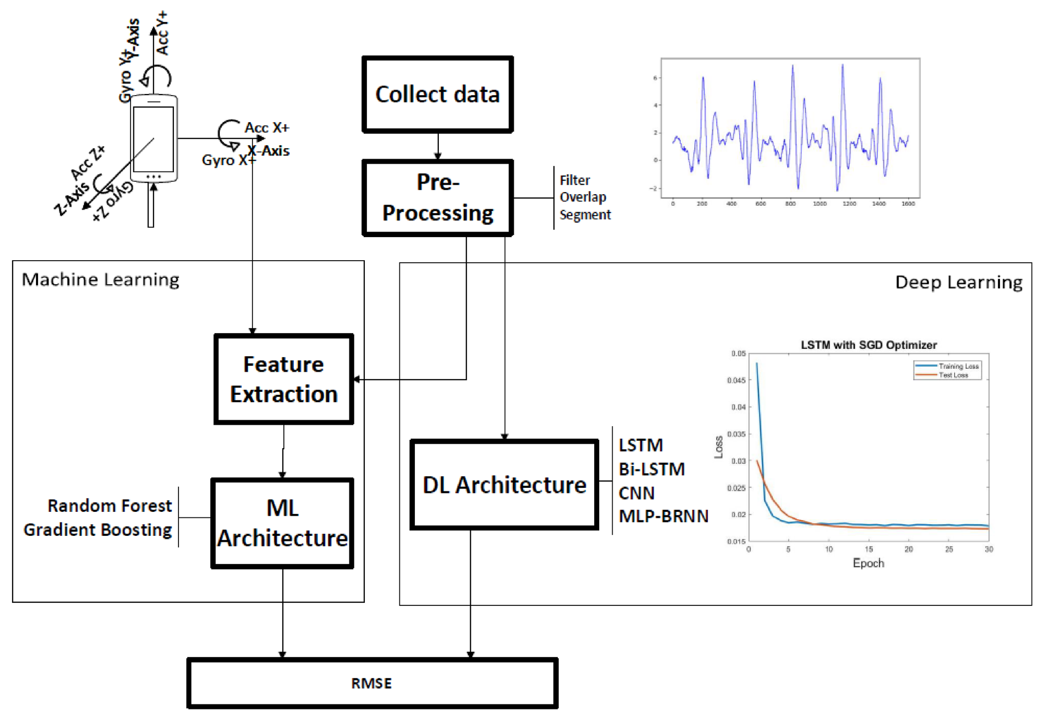


Figure 3: Flowchart of our experimental methodology.

LSTM: LSTM has been demonstrated to be especially useful for sequence long term data.

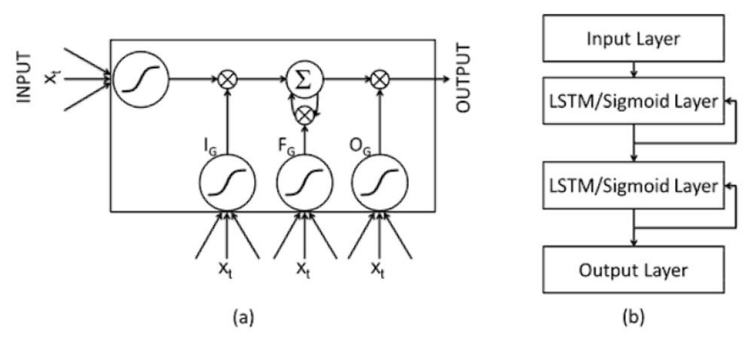


Figure 4: (a)LSTM cell (b)Stacked Architecture.

Bi-LSTM: We used a Bi-LSTM neural network implemented in the MXNet and Tensorflow Frameworks.

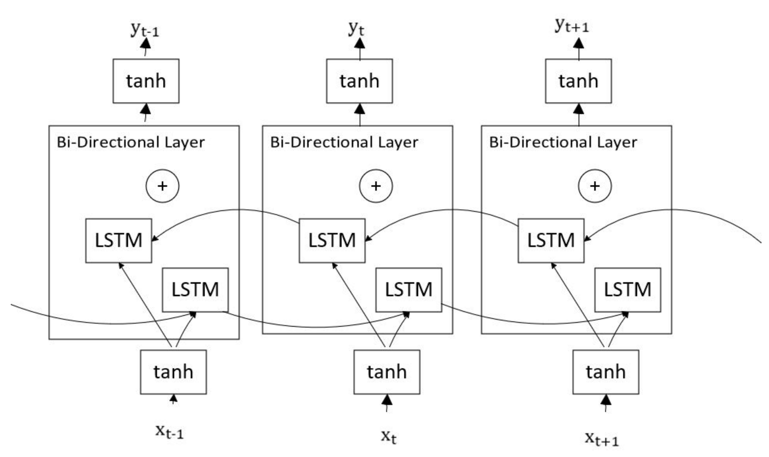


Figure 5: Unfolded Bi-directional LSTM Structure.

GAF: Gramian Angular Field(GAF) encode raw time-series data into an image format that produces highly competitive classification results using CNNs [6]. We believe it will allow us to improve regression performance.

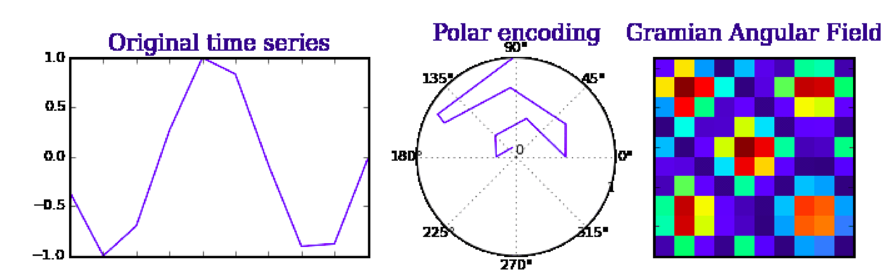


Figure 6: Imaging Time series by Gramian Angular Field.

Results

In figure 7, We selected random forest and gradient boosting ML algorithms on hand-crafted features as baselines. LSTM, CNN, and MLP on raw time-series data are compared with the baseline algorithms.

In figure 8, the performance of 3 different optimizers' are compared (Adam, Adagrad, and SGD). During the preprocessing procedure, overlapping and non-overlapping segmentation approaches are compared. All experiments were repeated 20 times, and the average performance is computed.

Root mean square error(RMSE) is the metric of a models regression performance.

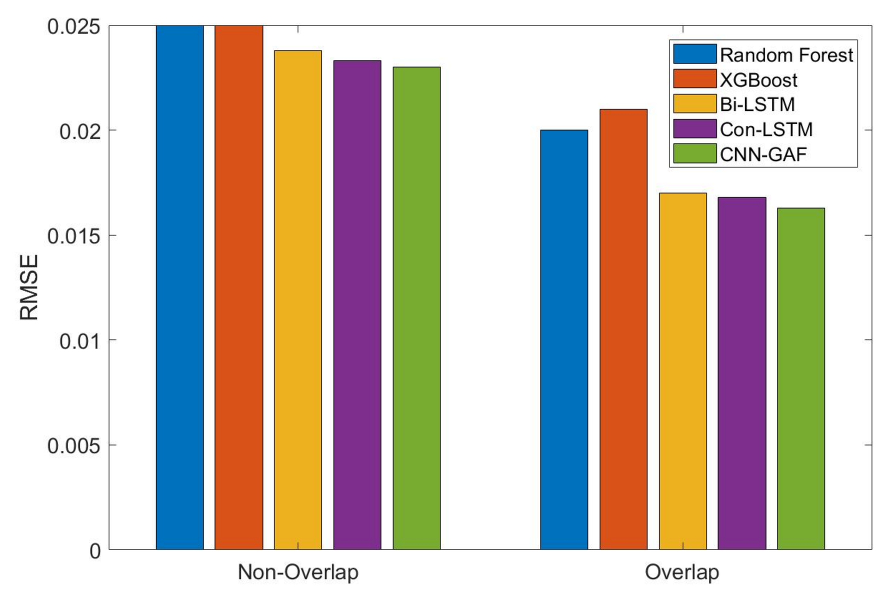


Figure 7: Comparison of Machine Learning with features and LSTM without features.

Figure 7 shows that overlapped segments have a smaller RMSE than non-overlapped segments. Also, LSTM and CNN deep learnign approaches outperform the Random Forest and gradient boosting ML approaches.

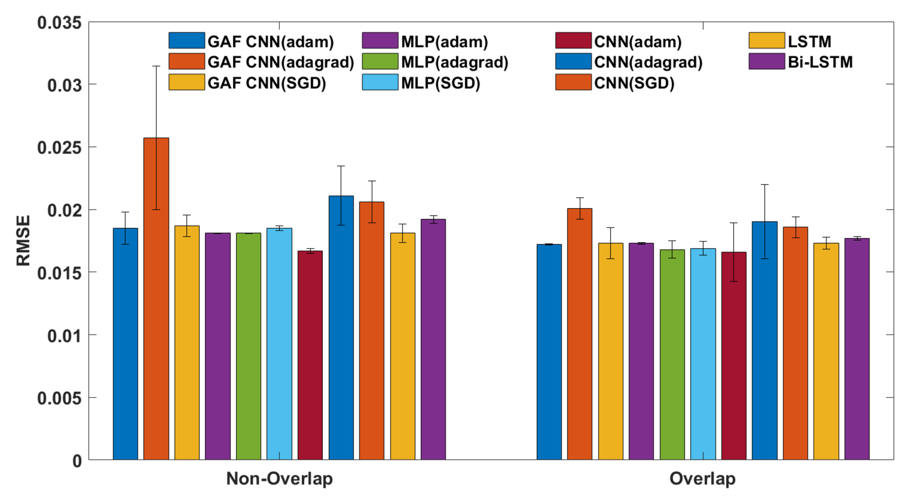


Figure 8: Comparison of LSTM,CNN and MLP with different optimizers

Figure 8 illustrates the difference between deep learning algorithms and optimizers.

Conclusion

- This research found that overlapping data segments improves results of intoxicated gait level analysis.
- Deep learning methods (LSTM and CNN) on raw data outperformed ML (random forest, gradient boosting) methods on hand-crafted features.
- The Adam optimizer performs better than Adagrad and SGD for most DL algorithms explored.

Detecting Blood Alcohol Level from Smartphone-Based Gait Data using Neural Networks

Ruojun Li [1] Jiaming Nie [1] Ganesh Prasanna [1] Emmanuel O. Agu [1] Michael Stein, MD[2]
Ana Abrantes, PhD[2] Debra Herman [2] Kristin Grimone [2]

[1]Computer Science, Worcester Polytechnic Institute

[2]Butler Hospital, Brown University

References

[1] H.-L. C. Kao, B.-J. Ho, A. C. Lin, and H.-H. Chu, "Phone-based gait analysis to detect alcohol usage," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, ser. UbiComp '12, 2012, pp. 661–662.

[2] Z. Arnold, D. Larose, and E. Agu, "Smartphone inference of alcohol consumption levels from gait," in *2015 International Conference on Healthcare Informatics (ICHI)*, vol. 00, 2015, pp. 417–426.

[3] H. F. Nweke, Y. W. Teh, M. A. Al-garadi, and U. R. Alo, "Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges," *Expert Systems with Applications*, vol. 105, 2018.

[4] O. Dehzangi, M. Taherisadr, and R. ChangalVala, "Imu-based gait recognition using convolutional neural networks and multi-sensor fusion," *Sensors*, vol. 17, no. 12, 2017.

[5] M. A. Gutierrez, M. L. Fast, A. H. Ngu, and B. J. Gao, "Real-time prediction of blood alcohol content using smartwatch sensor data," in *Smart Health*, X. Zheng, D. D. Zeng, H. Chen, and S. J. Leischow, Eds. Cham: Springer International Publishing, 2016, pp. 175–186.

[6] Z. Wang and T. Oates, "Imaging time-series to improve classification and imputation," in *Proceedings of the 24th International Conference on Artificial Intelligence*, ser. IJCAI'15, 2015, pp. 3939–3945.