Distraction Detection through Health Signals in Real-time

ZEYU CHANG* and SIDDHARTH DAS*, Georgia Institute of Technology, USA

Distraction detection is critical in various contexts, notably in driving scenarios, where it can prevent accidents and ensure safety. We present a novel approach to distraction detection utilizing a limited number of sensors placed on the hand for practicality. We conducted a comprehensive focus test, collecting electrocardiogram (ECG) and electrodermal activity (EDA) data from volunteers engaged in various distraction scenarios. Signal processing techniques, including filtering and peak detection, are applied to the ECG data, followed by classification using a simple neural network. Results demonstrate promising performance on same-subject data, with F1 scores exceeding 75%. Our study highlights the feasibility of hand-worn sensors for distraction detection, underscored by the need for personalized modeling and improved generalization techniques.

CCS Concepts: • **General and reference → General conference proceedings**.

Additional Key Words and Phrases: Signal Processing, Machine Learning, Neural Networks, ECG, EDA, Health Sensing

ACM Reference Format:
Zeyu Chang and Siddharth Das. 2024. Distraction Detection through Health Signals in Real-time. 1, 1 (May 2024), 9 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In contemporary society, the imperative to detect and mitigate distractions holds significance, particularly within the area of transportation safety, where split-second decisions can spell the difference between an uneventful journey and a catastrophic incident. Among the multifaceted approaches to this challenge, the integration of wearable technology and physiological signal analysis emerges as a compelling avenue.

The surge in deep learning and neural networks in the past decade has catalyzed a surge in research of using computer vision to approach this problem[5], [6], [15]. However, the computational demands inherent in vision-based approaches, particularly those reliant on convolutional neural networks (CNNs). This poses significant barriers to real-time implementation, compromising their utility in dynamic scenarios where quick feedback is required. As a result, wearable technologies, augmented by physiological signals such as electrocardiogram (ECG), electrodermal activity (EDA), electromyography (EMG), and electroencephalogram (EEG) [9], hold promise for unobtrusive, real-time distraction detection. The use of health signals circumvents the computational burdens of visual analysis while providing actionable insights into the cognitive states of individuals.

In response to these imperatives, this paper introduces a novel approach to distraction detection, centered on the utilization of a limited number of sensors strategically positioned on the hand for optimal wearability and unobtrusiveness. Our methodology is grounded in a comprehensive focus test, wherein ECG and EDA data were collected

*Both authors contributed equally to this research.

Authors' Contact Information: Zeyu Chang, chang.zeyu@gatech.edu; Siddharth Das, sdas337@gatech.edu, Georgia Institute of Technology, Atlanta, Georgia, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.
Manuscript submitted to ACM

from volunteers engaged in diverse distraction scenarios. Leveraging advanced signal processing techniques, including filtering and peak detection, in conjunction with a simple neural network architecture, we demonstrate the viability of hand-worn sensors for real-time distraction detection, achieving promising performance metrics with F1 scores exceeding 75% on same-subject data.

2 RELATED WORKS

Fatigue and distraction are two major causes of inefficiency in almost any context. In terms of driver fatigue or driver distraction, they can lead to severe road accidents, causing damage to not only the drivers but also their surrounding environments. Past research on this area can be divided into three major methods: Math Model-based approaches, Rule-based approaches – in particular Fuzzy Inference Systems (FIS) [1], and Machine Learning (ML) based approaches [13]. For ML-based approaches, they can be further categorized into learning using shallow models or using deep models.

With the surge of Deep Learning (DL) in the past decade, many ML-based distraction detection research looks into the capabilities of Convolutional Neural Networks (CNN) in detecting anomalies in images[5], [6], [15], [2]. These CNN-based approaches frequently make use of a camera system that captures facial features during driving and use CNN to perform binary classification on these image datasets. More recent research using DL models often combines CNN with additional networks, such as an LSTM network [10]. [11]. Another extension of a CNN-based approach is to combine its result with data from other modalities, such as body signals [3], [14], eye tracking data [14], road-view cameras [8], etc.

While DL-based approaches have given promising results, they can be difficult to be implemented in a real-time system, given the computational intensity of CNNs and image classification. Furthermore, the installation of cameras within vehicles can be challenging, and these models may perform badly when lighting conditions are not optimal. Research using health signals is free of this issue. In fact, using biological signals offers good indication of early signs of fatigue and distraction [13]. However, many researchers use EEG signals[7], [16], [9], which is highly intrusive due to the placement of electrodes. In addition, EEG signals are extremely susceptible to noise, making signal processing a major part of any distraction detection system [13].

[9] shows that in addition to EEG signals, other signals such as EMG, ECG, and EDA can also show differences when drivers are in a distracted state. This motivates us to use ECG and EDA signals to perform distraction detection. To further reduce the intrusiveness nature of signal collection, we come up with a compact way of placing sensors on the palms, simulating a pair of driving gloves. We then take on a shallow ML classification model, consisting of 3 layers of fully connected neural network, to classify processed signals.

3 DATA COLLECTION

To collect meaningful data that represent both a focus state and a distracted state, we decide to construct a custom-built focus test webpage. We use this test interface and combine it with other distractions common in driving, and collect ECG and EDA signals as subjects go through the focus test.

3.1 Custom Focus Test Interface

The custom focus test performed 100 tasks of 4 different types in rapid succession. The first type is a speed-based reaction test, which attempts to observe how quickly the testee clicks any button. The other three types involve the recognition of a matching pattern on the screen with subtle variance between the tasks. Clicking the matching word, the matching text color, or the matching background color are the 3 tasks. Between each task, the location of the clicked button shifts, to slightly vary the scenario each time. Observe Figure 1 for a sample task within the test interface.

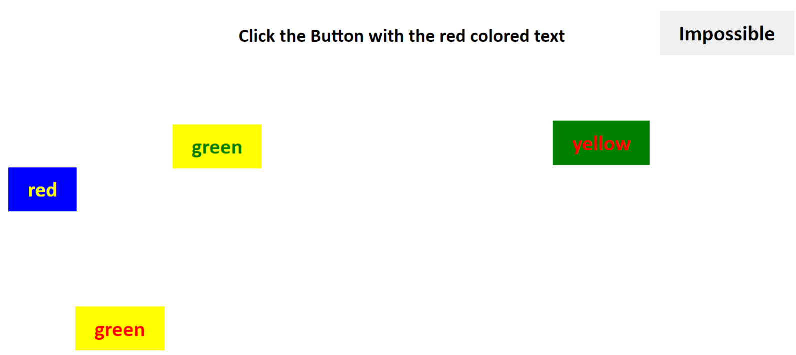


Fig. 1. Sample Task within the Testing Interface

3.2 Raw Signal Collection

To minimize the interference of sensors on human activities, we cluster the two sensors on the fingers of both hands: EDA sensor is placed on the left index and ring finger; ECG sensor is placed on the left index finger, right index, and middle finger. We believe such placement is conventional and less intrusive than EEG sensors in previous research [13], as these sensors can be integrated into a pair of driving gloves.

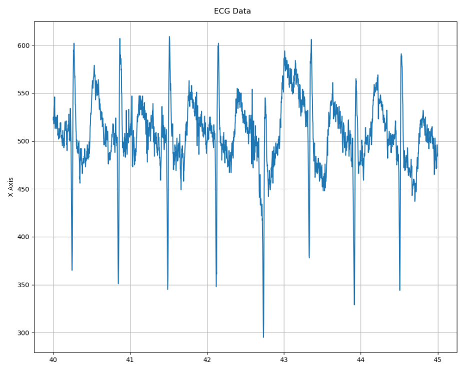
To record the signals, we conducted 2 trials, each consisting of 4 separate runs of the focus test. Specifically, we record one run without any outside interference – this will serve as the baseline where the test subject is focusing on the task. For distraction tasks, we took an incrementing approach. To start, the test subject will listen to a piece of music they like; in the next round, the subject will open up a video by themselves and watch it while doing the test; and finally, the subject will have a conversation with another person. We picked these distraction scenarios because we think they can best simulate the distractions a driver will face: listening to the radio, looking at the scenery/road accident, and having a phone call. The distractions are conducted after 1 minute into each run. Each run is conducted twice to prevent unexpected failure during data collection.

3.3 Signal Processing

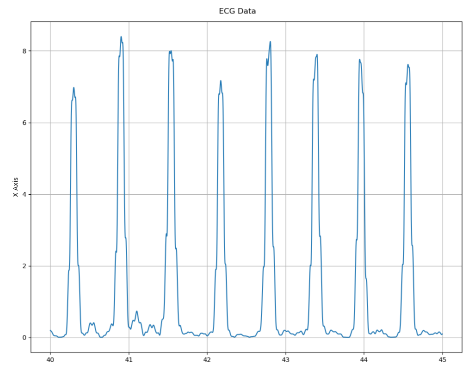
Raw signals contain a non-trivial amount of noise and need to be filtered and processed before performing any analysis. Luckily, past research has shown several ways to process these signals and retrieve useful information from them.

For EDA signal processing, the book titled "Advances in Electrodermal Activity Processing with Applications for Mental Health" provides both guidance and inspiration for our signal processing methods [4]. Initially, we utilize a moving average of 0.5 seconds to reduce the impact of movement-based artifacts and data disparities. After that, a second order band-pass filter is utilized on the resulting data, with an upper bound of 3 Hz. That creates our refined signal, which undergoes further manipulation later.

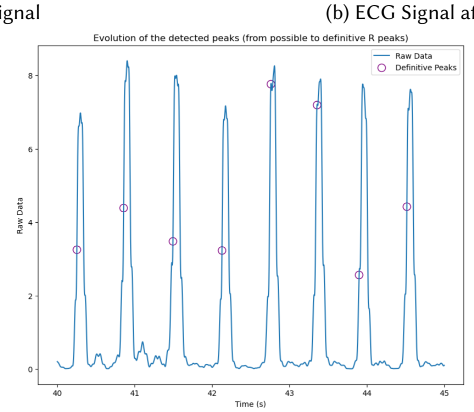
In terms of ECG signal, past research presented the Pan-Tompkins algorithm for processing raw ECG signal into smoothed signal suited for R peak detection [12]. Raw Signal like figure 2a is first processed using a combination of band-pass filter, derivative filter, squaring, and moving average to get the Integrated Signal in figure 2b. By using peak detection algorithm, we are able to estimate the R peaks in the integrated signal, represented in figure 2c. Finally, the ECG signal is broken down into R-R intervals for later classification tasks.



(a) Raw ECG Signal



(b) ECG Signal after Pan-Tompkins Algorithm



(c) Estimated R peaks marked

Fig. 2. ECG Signal, before and after signal processing

4 MODELS & EXPERIMENTS

4.1 ECG R-R Interval Study

With the R-R intervals extracted from raw signals, we wish to gain information on how distractions may cause slight differences in R-R interval signals. Since we want a fast classification system for future real-time implementation, we decide that a simple neural network will be the perfect model for this task.

Data normalization: To standardize the approach, all R-R interval data is padded to the same length of 1000 by edge values. This would work nicely as long as the subject's heart rate is above 60 bpm. The padded signals are then passed into a 3-layer Fully Connected Neural Network (FCNN) with a binary output, representing the signal from either the focus section or the distracted section.

Training set selection: There are several ways to train the model with the collected data. The most intuitive way is to train the model on the R-R intervals of a single person, since signals across different people can vary dramatically.

Apart from training the model custom to each of our volunteer, we also tried training the model on all the signals together, in hope of creating a generalized model. Finally, we attempt transfer learning by training the model base on the signals from a person, say Alice, and classifying the signals from another person, say Bob.

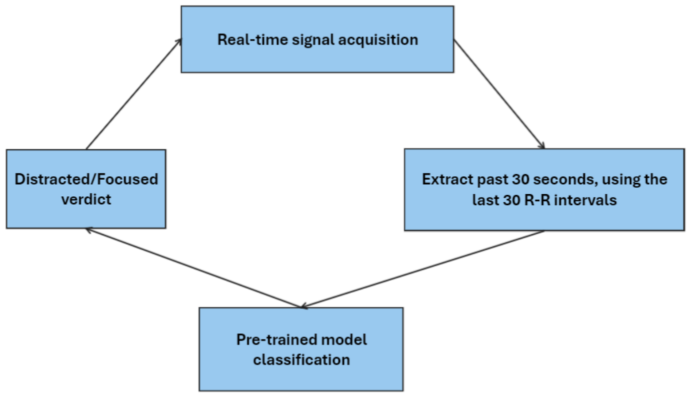


Fig. 3. Real time system diagram

Real-time system: Our real-time system works as follows: suppose the model is trained on Alice's data. After the model converges, we use the model to predict whether Alice is distracted or not for the past 30 seconds. This is implemented by signal processing the real-time acquired ECG signal and using the past 30 R-R intervals. If the number of intervals classified as "Distracted" exceeds a certain threshold, then we report that Alice is distracted for the past 30 seconds.

4.2 EDA Peak Count Study

With clean EDA signal data, we chose to experiment with rapid analysis methods for future real-time applications. From our test signals, we exclude the first minute of data, because the first minute of every trial was the same. The distractions always began after the first minute. The justification for this was to potentially observe a visible difference within the EDA after the distraction mechanism began.

No further time window narrowing occurs in order to mantain the support for real-time analysis. This system functions currently on 4 minute intervals, where signal data is ready for analysis every 4 minutes. From there, we perform a peak extraction algorithm in order to count the fluctuations within the EDA signal. Depending on how the peaks contrast with the baseline, the potential to detect a change in mental focus is there. Further discussion of this occurs in the results section.

5 RESULTS

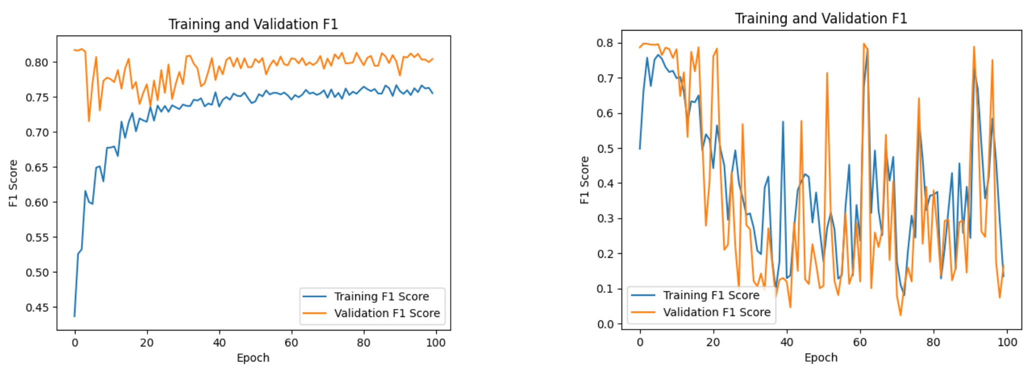
5.1 R-R Interval binary classification

F1-Score comparison: From table 1, we conclude that when using the same network architecture, custom fitted model should be used instead of generally fitted model. When trained on person-specific signals, a 3-layer FCNN can give some nice results where the F1 score is above 75%. On the other hand, fitting a model with combined data from different volunteers doesn't work, as we can see the extreme fluctuation in F1 score from figure 4b. Furthermore, transfer learning

Test Subject	Custom model	General model	Transfer learning (1)
Person 1	0.814	0.45	-
Person 2	0.763	0.45	0.354
Person 3	0.772	0.45	0.239

Table 1. F1 score table. Clearly, custom models work the best as they are fitted to each person, while transfer learning perform extremely bad. F1 score is selected as the metric because labeled data is imbalanced.

also gives very poor F1 scores. These results clearly indicates that health signals differs significantly across people, and models should learn on person-specific data.



(a) Custom fitted model

(b) Generalized model

Fig. 4. F1 plots. A custom fitted model converges, while the model trained on combined data has a lot of fluctuation

Real-time detection using trained model: To validate our real-time detection system design, we first pass in the recorded data (training and testing signals) in time order to see how the system perform. After that, test volunteers perform 2 runs of the focus test, one without interference and the other one with conversation interference. From figure 5, our system can give a F1 score of above 75%. When conducting new trials of the focus test, this system can indeed report distraction when test volunteers are having a conversation.

5.2 EDA Peak Count

The full analysis results for average EDA peak counts by activity are illustrated within figure 6. We can observe a noticeable increase in the quantity of peaks across every individual with the presence of distractions. Some individuals are displaying more peaks from activities which generate greater cognitive load (video & conversation). However, due to the number of trials performed, we are currently unable to provide deeper level analysis into detecting levels of distraction. Still, the insight of an increased peak count compared to the baseline is a useful insight.

Furthermore, performing this analysis in real-time creates the potential to generate a baseline EDA peak count for each individual. As observed though, the variance due to sensor attachment and the natural state between individuals prevents transferring comarpison baselines between individuals. Overall, EDA analysis is viable for distraction detection.

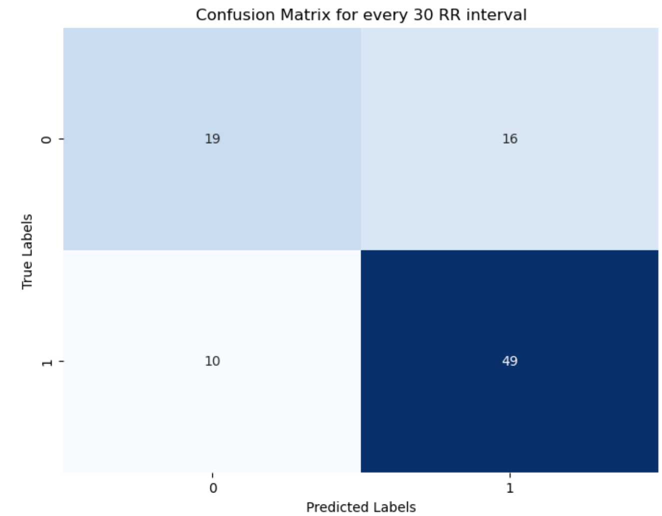


Fig. 5. Confusion matrix of real-time detection result

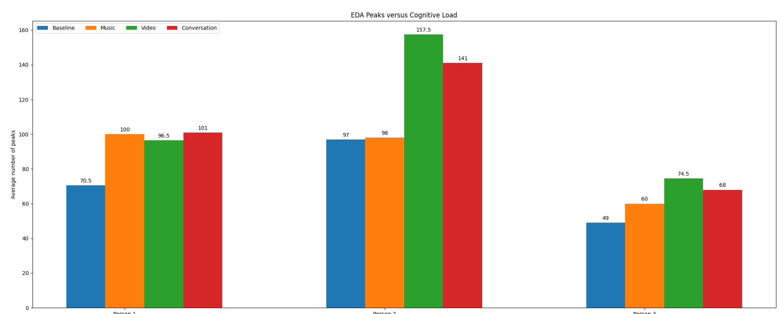


Fig. 6. EDA signal peaks results. When watching a video or having a conversation, there is significantly more peaks (green, red) in general compared to the baseline (blue).

6 DISCUSSION & CONCLUSION

We proposed a method of detecting distraction and reporting them in real-time, using only limited signal sensors placed on the fingers. While many recent research in the area involves extensive use of complex deep learning neural networks and computer vision, we particularly aimed to avoid the use of facial expressions and computer vision methods for simpler computation. We started by selecting the sensors and devising a custom focus test interface to simulate a scenario that requires high focus. After collecting the signals from multiple volunteers, we looked into ways of processing the signals and classifying detection using processed signals. Our results shows that with properly processed signals, a simple 3-layer FC network can classify ECG R-R intervals with a F1 score of over 75% when the model is fitted per person. Furthermore, EDA signals shows more frequent peaks when cognitive load increases, suggesting its usefulness in distraction detection.

Many previous works using CNNs certainly have better results, achieving above 85% F1 score [3]. Research using a combination of CNN and LSTM or using multi-modal data can achieve even better results, with accuracy above 95% [11],

[10]. Other research studying EEG signals also achieve nicer results than our approach [7], [16]. However, we struggle to find discussion on their models' real-time capabilities or their sensors' intrusiveness in the literature. Our method, despite having lower performance overall compared to advanced computer vision approaches and EEG approaches, shows consistent performance in real-time and is much less intrusive than wearing EEG electrodes. Furthermore, real-time testing shows that our system and model can be used on a consumer laptop CPU to examine 30 seconds of signals in less than 2 seconds, suggesting promising efficiency.

LIMITATIONS

Our real-time distraction detection system based on ECG R-R interval classification can give non-trivial results when trained on signals from the same person. However, this system is not ideal because while it has a high true-positive rate (distractions are detected), it also has a high false-positive rate since half of the sequence are marked as "distracted" even though they are not. Since the performance of our real-time system directly relates to the performance of our model itself, improvements should be done to increase our FCNN's accuracy. Future work could explore the combination of EDA analysis with this model, as it may reduce the false-positive rate, but a threshold of EDA peak needs to be customized for every test subject. Furthermore, machine learning model always works better when the amount of data is large, so training the model with more signals from the user could help the model to perform better.

REFERENCES

[1] Silvia Cateni, Valentina Colla, Marco Vannucci, Alice Borselli, et al. 2012. Fuzzy inference systems applied to image classification in the industrial field. *Fuzzy Inference System: Theory and Applications* (2012).

[2] Kartik Dwivedi, Kumar Biswaranjan, and Amit Sethi. 2014. Drowsy driver detection using representation learning. In *2014 IEEE international advance computing conference (IACC)*. IEEE, 995–999.

[3] Martin Gjoreski, Matja Ž Gams, Mitja Luštrek, Pelin Genc, Jens-U. Garbas, and Teena Hassan. 2020. Machine Learning and End-to-End Deep Learning for Monitoring Driver Distractions From Physiological and Visual Signals. *IEEE Access* 8 (2020), 70590–70603. https://doi.org/10.1109/ACCESS.2020.2986810

[4] Alberto Greco, Gaetano Valenza, and Enzo Pasquale Scilingo. 2018. *Advances in Electrodermal Activity Processing with Applications for Mental Health: From Heuristic Methods to Convex Optimization* (1st ed.). Springer Publishing Company, Incorporated.

[5] Whui Kim, Hyun-Kyun Choi, Byung-Tae Jang, and Jinsu Lim. 2017. Driver distraction detection using single convolutional neural network. In *2017 International Conference on Information and Communication Technology Convergence (ICTC)*. 1203–1205. https://doi.org/10.1109/ICTC.2017.8190898

[6] Matti Kutila, Maria Jokela, Gustav Markkula, and Maria Romera Rue. 2007. Driver Distraction Detection with a Camera Vision System. In *2007 IEEE International Conference on Image Processing*, Vol. 6. VI – 201–VI – 204. https://doi.org/10.1109/ICIP.2007.4379556

[7] Guofa Li, Weiquan Yan, Shen Li, Xingda Qu, Wenbo Chu, and Dongpu Cao. 2021. A temporal–spatial deep learning approach for driver distraction detection based on EEG signals. *IEEE Transactions on Automation Science and Engineering* 19, 4 (2021), 2665–2677.

[8] Nanxiang Li and Carlos Busso. 2014. Predicting perceived visual and cognitive distractions of drivers with multimodal features. *IEEE Transactions on Intelligent Transportation Systems* 16, 1 (2014), 51–65.

[9] Ying Li, Li Zhao, Kun Gao, Yisheng An, and Jelena Andric. 2022. Revealing driver psychophysiological response to emergency braking in distracted driving based on field experiments. *Journal of Intelligent and Connected Vehicles* 5, 3 (2022), 270–282. https://doi.org/10.1108/JICV-06-2022-0024

[10] Jimiama Mafeni Mase, Peter Chapman, Grazziela P. Figueredo, and Mercedes Torres Torres. 2020. A Hybrid Deep Learning Approach for Driver Distraction Detection. In *2020 International Conference on Information and Communication Technology Convergence (ICTC)*. 1–6. https://doi.org/10.1109/ICTC49870.2020.9289588

[11] Luntian Mou, Chao Zhou, Pengfei Zhao, Bahareh Nakisa, Mohammad Naim Rastgoo, Ramesh Jain, and Wen Gao. 2021. Driver stress detection via multimodal fusion using attention-based CNN-LSTM. *Expert Systems with Applications* 173 (2021), 114693.

[12] Jiapu Pan and Willis J. Tompkins. 1985. A Real-Time QRS Detection Algorithm. *IEEE Transactions on Biomedical Engineering* BME-32, 3 (1985), 230–236. https://doi.org/10.1109/TBME.1985.325532

[13] Gulbadan Sikander and Shahzad Anwar. 2019. Driver Fatigue Detection Systems: A Review. *IEEE Transactions on Intelligent Transportation Systems* 20, 6 (2019), 2339–2352. https://doi.org/10.1109/TITS.2018.2868499

[14] Salah Taamneh, Panagiotis Tsiamyrtzis, Malcolm Dcosta, Pradeep Buddharaju, Ashik Khatri, Michael Manser, Thomas Ferris, Robert Wunderlich, and Ioannis Pavlidis. 2017. A multimodal dataset for various forms of distracted driving. *Scientific data* 4, 1 (2017), 1–21.

[15] Martin Wollmer, Christoph Blaschke, Thomas Schindl, Björn Schuller, Berthold Farber, Stefan Mayer, and Benjamin Trefflich. 2011. Online Driver Distraction Detection Using Long Short-Term Memory. *IEEE Transactions on Intelligent Transportation Systems* 12, 2 (2011), 574–582. https://doi.org/10.1109/TITS.2011.2119483

[16] Xin Zuo, Chi Zhang, Fengyu Cong, Jian Zhao, and Timo Hämäläinen. 2022. Driver distraction detection using bidirectional long short-term network based on multiscale entropy of EEG. *IEEE Transactions on Intelligent Transportation Systems* 23, 10 (2022), 19309–19322.