International Journal of Computer Trends and Technology
ISSN: 2231–2803 / https://doi.org/10.14445/22312803/IJCTT-V71I10P112

Volume 71 Issue 10, 107-113, October 2023

Original Article

Enhancing Road Safety: Detecting Texting Distracted Driving with Eye-Tracking and Machine Learning

Rohan Singh Rajput[1], Shantanu Neema[2]

[1]*Headspace, Los Angeles, CA, USA.*
[2]*Syntelli Solutions, Austin, TX, USA.*

[2]*Corresponding Author : shantanu.neema@syntelli.com*

Received: 29 August 2023 Revised: 01 October 2023 Accepted: 14 October 2023 Published: 31 October 2023

Abstract *- Driving while texting is risky as it diverts the driver's attention from the road and requires them to shift between handling their phone and the vehicle. Despite its dangers, many still engage in this behaviour. To address this, some companies have implemented features to detect when someone is driving and restrict calls and messages unless confirmed otherwise by the user. This research introduces a method that combines eye-tracking technology with machine learning to identify when a driver is texting. A driving simulator was utilized to evaluate 26 participants under various conditions: normal driving, emotional distraction, cognitive distraction, and texting. Features were extracted from the eye movement data, encompassing fixation count and duration. After processing this data through machine learning models, an impressive accuracy rate of over 90% for identifying texting while driving. These findings are promising and hint at the potential for a realtime system that can detect and warn drivers when they are engaged in texting.*

Keywords *- Texting distraction, Driving performance, Eye tracking, Machine learning, Statistical analysis.*

1. Introduction

Texting while driving is one of the common distractions that can potentially result in fatal accidents and near-death experiences. Many studies have documented that distracted driving, including texting, significantly increases the chances of accidents [1][3][7][8]. This finding has prompted the prohibition of cell phone usage while driving in many states across the United States. A recent report from the National Highway Traffic Safety Administration (NHTSA) reported that using a cell phone while driving creates enormous potential for deaths and injuries on the roads in the United States [7]; in 2020 alone, a total of 3,142 people were killed in motor crashes involving distracted drivers.

One potential solution to mitigate the risk of distracted driving is to identify such behaviour as early as possible in real time. Different machine learning algorithms have been used to predict driver distraction. A few are mentioned here: Kircher & Ahlstrom used logistic regression to predict driver's distraction based on the driving performance indicators [4]. Ragab et al. used random forest to detect the driving distraction [9]. Liang et al. proposed a realtime system for driver's cognitive distraction using support vector machines [5]. The same group proposed a Bayesian networks-based distraction detection system [6]. These approaches mainly focused on detecting visual distractions, such as adjusting in-vehicle devices (e.g., tuning radio) or detecting cognitive distractions like cell phone conversations while driving. However, less attention has been given to the automatic detection of texting distraction, primarily a combination of visual distraction (eyes off the road) and sensorimotor distraction (one hand moving between the car controls and the smartphone).

This paper addresses the research gap in detecting texting distractions automatically through eye-tracking data. Most research work that reviewed the eye-tracking approach focused on various driving distractions but not specifically texting distractions [11][12][13][14]. Recarte and Nunes reported the impact of verbal and spatial imagery tasks, such as repeating words and mental image rotation, on eye fixations while driving [11]. Strayer et al. used eye-tracking data to study the effect of hands-free cell phone conversation on visual attention [12]. Sodhi and Reimer analysed fixation durations for radio-tuning and rear-view mirror-checking tasks [13]. Victor et al. studied eye movement data for visual and auditory in-vehicle tasks [14]. Bitkina at el. suggested that the driver's fatigue and workload is based on gaze behaviour and eye tracking metric could be utilized for predictions [16]

Thus, detecting texting distracted driving in real time remains an open challenge. This paper aims to contribute to this area of research. A machine learning-based approach is proposed in this work, which first extracts multiple features from the drivers' eye-tracking data, examines the features for statistical significance, and finally feeds the features that pass the significance test to a set of machine learning algorithms. The following sections explain the approach taken to answer this.

2. Experimental Design

The participants were recruited from Bryan and College Station, TX communities. All the participants had at least 1.5 years of driving experience and a valid driver's license and were not on medications which might affect their ability to drive safely. Participants were in 2 age groups, young cohorts 18-27 years and old cohorts above 60 years of age. The personality type A/B using Jenkins Activity Survey [15] is scored for each participant. Additionally, trait anxiety inventory (TAI) is also noted for each participant to see the effects of long-term stress on driving behaviour or eye movements.

The driving performance data is collected using a high-fidelity simulator manufactured by Realtime- Technologies, Inc. Eye tracking data was collected unobstructed using two components, a light source and a camera. The Institutional Review Boards (IRB) of the University of Houston and Texas A&M University approved the experimental procedures. The light source is directed towards the eye, and the camera tracks the reflection of the light source along with ocular features such as the pupil.

Fig. 1 Simulated driving setup

In the simulated driving setup, each participant drove along a 10.9 km long track with two lanes in each direction. The simulated environment was designed in a daylight setting where there were no cars to follow, only oncoming traffic (>12 vehicles per km), no traffic lights or stop signs during each session (except at the beginning), the posted speed limit of 45 mph on which participants drove for about 12-14 min per session with realistic highway view including construction zones, lane marks, construction cones etc. The simulator used 3 screens, one on each side and the middle screen about ~ 3.5 ft away from the participant. All screens were displayed at 1920px. × 1080px. resolution. The eye gaze locations were recorded at 60 Hz from the top-right corner of the middle screen, which approximates about 90° visual field for a participant on that screen. Driving parameters like speed, acceleration, steering angle, lane position and braking were recorded at 42 Hz.

Using the driving environment, as explained earlier, the participants were asked to drive. Each participant went for four different sessions: normal, cognitive, emotional, and driving while texting. The order of these four driving sessions was randomized. The secondary activity was triggered by questionnaires by an experimenter for cognitive and emotional distractions. A sensorimotor was used for texting by sending back words to participants smart phones. There was a 2-minute break between each session of the drive. Each session of the drive was designed in the following pattern.

N1: No distraction (about 80 sec)

D1: Driving distraction (about 160 sec)

N2: No distraction (240 sec)

D2: Driving distraction (about 160 sec)

N3: No distraction (about 120 sec)

Out of these 5 phases, phases D1 and D2 were used for analysis. D1 and D2 of each distracted driving were compared against corresponding driving periods in the baseline driving session.

2.1. Texting Distraction

During each distracted driving phase (D1 and D2), the participants received one word at a time on the smartphone. The participants were asked to type each word in the backward order and send as they received.

2.2. Cognitive Distraction

The participants were asked a set of mathematical questions in one of the two distracted driving phases (D1 and D2) and a set of analytical questions in the other distracted driving phase. The experimenter asked these questions orally. The order of the phases was randomized.

An example of the mathematical question is: What is the sum of 24 + 58? An example of the analytical question is My grandfather's daughter hit her daughter. How do the daughters relate to each other? The participants were required to answer these questions to the best of their abilities.

2.3. Emotional Distraction

During each of the two distracted driving phases (D1 and D2), the participants were asked a set of emotionally charged questions by the experimenter, such as Give an example of a time when you were angry with someone in the past and realized it was not their fault? The participants were required to answer these questions to the best of their abilities.

3. Data Collection and Feature Extraction

Eye-tracking data from 26 participants were used in the analysis. The participants were in two age groups: 15 young cohorts (9 female) of ages ranging from 21 to 27 years (mean = 22.67, std. dev. = 1.80), and 11 old cohorts (6 females) of ages ranging from 61 to 73 years (mean = 67.18, std. dev. = 4.45). These participants collected 4.74 million eye gaze data points (60 points per second × 760 seconds per participant × 26 participants × 4 drives). From these raw data, we derived a set of eye movement features. A statistical significance test was conducted on these features, and finally, the parts that passed the statistical significance test to machine learning classifiers.

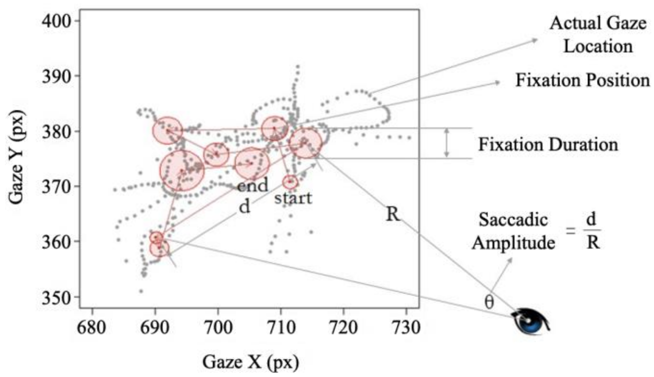

Fig. 2 The plot shows 10 seconds of the eye gaze data in the light gray coloured dots and the fixations in red circles. The fixations are computed from the dispersion-threshold algorithm [2]. The size of the red circles represents fixation duration. Fixation count represents the number of these red circles. Saccade duration is the time duration between two successive red circles

Dispersion-threshold-based algorithm extracted three features: fixation count, fixation duration, and saccadic duration [2]. For a given eye gaze location (x, y), the algorithm grouped successive gaze locations if the participant was within 20 20-pixel radius from the given point. To qualify this group of gazes as fixation, participants had to meet a minimum duration of 200 milliseconds (ms), about 12 successive gaze locations [2]. Otherwise, the group was discarded.

The algorithm was applied iteratively to the entire eye gaze dataset for fixation computation. The saccades were calculated as a time duration between two successive fixations. The data was noise-cleaned for spurious fixations (fixation duration > 2.5s) and spurious saccades (saccadic duration > 800ms) [2]. The thresholds were chosen according to the suggestions given in [2]. Figure 1 provides a visual illustration of the algorithm for 10 seconds of the eye gaze data.

4. Results and Discussion

4.1. Distracted Driving vs Baseline Driving

Next, the D1 portion of distracted driving was compared to the corresponding portion in normal driving (ND1). A similar comparison with the D2 portion against the corresponding portion in normal driving (say ND2). Each distracted driving phase (D1 and D2) had a slightly different corresponding normal driving portion (ND1 and ND2) because of the variability in each participant's driving speed. In the rest of the paper, the emotionally distracted driving phases D1 and D2 were noted as E1 and E2, respectively, and the corresponding typical driving phases ND1 and ND2 as NE1 and NE2, respectively. Similar notations for cognitive and texting distractions were used. For cognitively distracted driving, phases D1 and D2 are C1 and C2, and the corresponding typical driving phases ND1 and ND2 are NC1 and NC2, respectively. Likewise, the texting distracted driving phases D1 and D2 are denoted as T1 and T2, and the corresponding typical driving phases ND1 and ND2 are distinguished as NT1 and NT2.

For each driving phase, *fixation count* was calculated, representing the total number of fixations per participant in each driving segment and *fixation duration*, which represents the mean of each participant's fixation durations for each driving segment. These two features are visualized in Fig 1, in which the number of red circles represents *fixation count*, and the size of the circle represents *fixation duration*. From the saccades, *saccadic duration* representing the mean of each participant's time duration was calculated between two successive fixations for each driving segment.

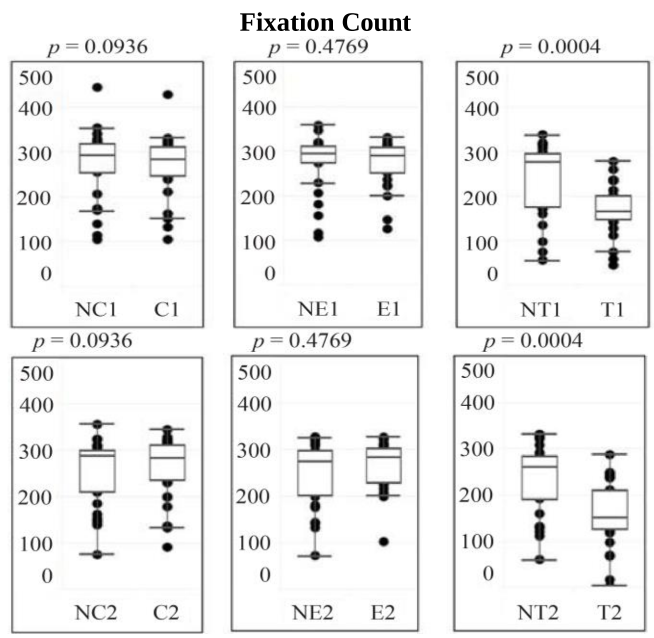

Fig. 3 Boxplots illustrate fixation count for normal driving and driving with distractions. n = 26 participants. The p-values for the paired t-tests are shown at the top of the box- plots. The differences are statistically significant, with 99% confidence interval for texting distractions.

4.2. Distracted Driving vs. Baseline Driving

A paired t-test was performed to examine whether any of the three eye movement features could reveal any statistically significant difference between distracted driving and typical driving phases. Specifically, emotionally distracted driving phases E1 and E2 were compared against the corresponding distinct driving phases NE1 and NE2. Similarly, cognitively distracted driving phases C1 and C2 were compared against the corresponding normal driving phases NC1, and NC2, respectively. Likewise, texting distracted driving phases T1 and T2 were compared against the corresponding normal driving phases NT1 and NT2, respectively.

Fig 3 illustrates the distributions of the participant's *fixation count* for the normal driving and the three distracted driving sessions. The p-values of the statistical tests are reported on the top of the plots. The results reveal that in comparison to normal driving, the participants had their eyes on the road significantly a smaller number of times ($p < 0.01$) during the texting distracted driving (T1, T2). This confirms a typical behaviour: drivers look away from the road while texting. For the other two distracted drives (i.e., cognitive C1, C2 and emotional E1, E2), the fixation count distributions are like that of the normal drives ($p > 0.05$), indicating that the participants don't look away from the road under the cognitive and emotional distractions.

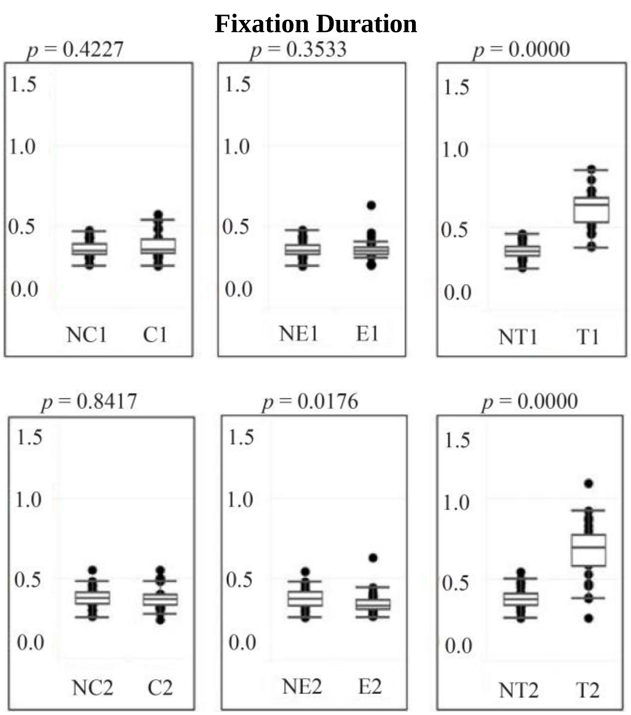

Fig. 4 Boxplots illustrate *fixation duration* for normal driving and driving with distractions. *n* = 26 participants. The *p*-values for the paired t-tests are shown at the top of the box plots. The differences are statistically significant, with 99% confidence interval for the texting

Fig 4 illustrates the distributions of the participant's *fixation duration* for the distracted and normal drives. The p-values of the statistical tests are reported on the top of the plots. The results reveal an interesting pattern; compared to normal driving, the participants fixated on the road longer while texting but not on the other two distractions. A possible justification for this behaviour is that, by looking on the road for a more extended period, the participants were compensating for the missing information about the surroundings they were supposed to acquire while looking at the cellphone screen for texting. It is interesting to find out how this behaviour affected their driving performance. Fig 6 compares participants' driving performance for the regular drive and distracted drives. The p-values of the statistical tests are reported on the top of the plots. It shows that the participants had significantly more significant lane departures ($p < 0.01$) while texting (T1, T2). Furthermore, the participants lower the acceleration (T1: $p < 0.1$; T2: $p < 0.01$) and drive slower ($p < 0.01$) while texting. These observations indicate that texting creates dangerous driving behaviour (lane departure), while drivers either consciously or subconsciously compensate for this riskier behaviour by slowing down (acceleration and speed).

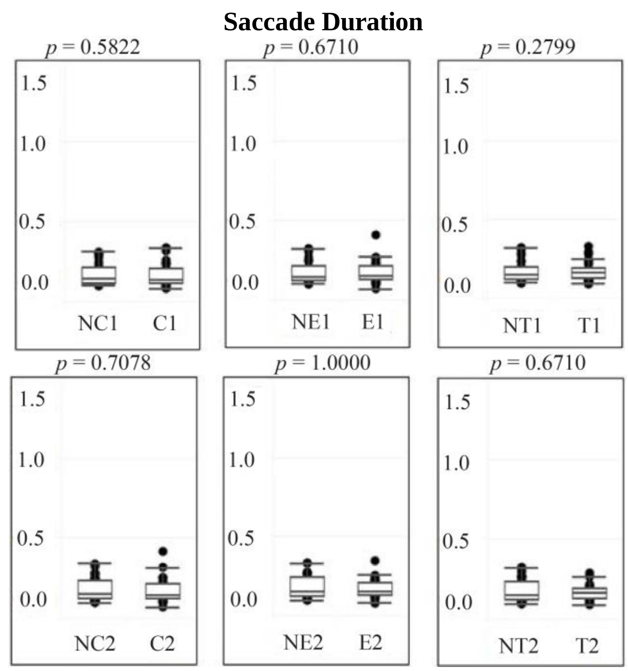

Fig. 5 Boxplots illustrate *saccade duration* for normal driving and driving with distractions. *n* = 26 participants. The *p*-values for the paired t-tests are shown at the top of the box plots. None of the differences are statistically significant

Fig 5 illustrates the distributions of the saccadic durations for all drives. The p-values of the statistical tests are reported on the top of the plots. It shows no statistically significant difference ($p > 0.1$) in saccadic durations between the drives, indicating that the time gap between two successive fixations of distracted drives was like that of the standard drives. Therefore, *saccade duration* may not be a suitable feature for classification.

Young Cohort (Fixation Duration)

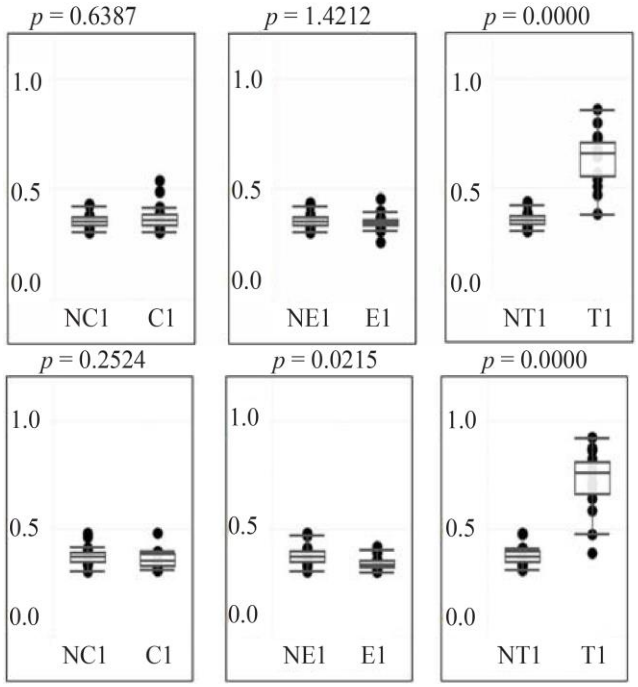

Fig. 6 Boxplots illustrate *fixation duration* for the young cohort. The *p*-values for the paired t-tests are shown at the top of the box plots. The differences are statistically significant, with 99% confidence interval for the texting distractions

Old Cohort (Fixation Duration)

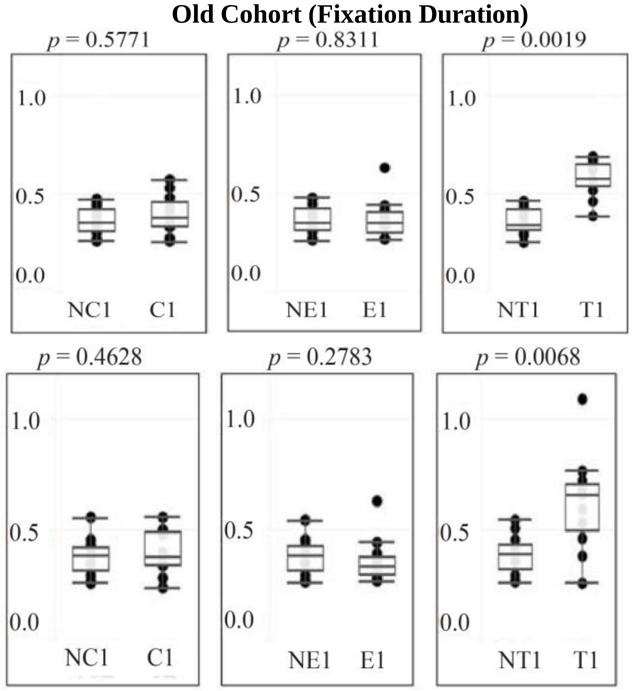

Fig. 7 Boxplots illustrate *fixation duration* for the old cohort. The *p*-values for the paired t-tests are shown at the top of the box plots. The differences are statistically significant, with 99% confidence interval for the texting distractions

Fig 6 and Fig 7 illustrate boxplots of fixation durations grouped by the young and old cohorts. Overall, the distributions of each cohort are like those observed for the entire group, meaning fixation duration for the cognitive and emotional distractions are not significantly different ($p > 0.01$) compared to the baseline driving. Still, it is pretty different for the texting distraction ($p < 0.01$). Interestingly, the young cohort exhibits different texting behaviour than the old cohort. The young cohort has a higher *fixation duration* than the old cohort, indicating that compared to the old group, the young group looks on the road for longer, resulting in better driving performance.

4.3. Machine Learning Classifiers

None of the three eye-gaze features revealed any statistically significant difference between the cognitively distracted drives (C1 and C2) and the usual drive and between the emotionally distracted drives (E1 and E2) and the standard drive. Therefore, these two distracted drives were dropped from the classification analysis. In the texting distracted drives, the *saccade duration* feature was dropped because it failed the statistical significance test. Therefore, the *fixation count* and *duration* feature were fed to machine learning (ML) algorithms to evaluate the features' predictive power in detecting texting distracted drives.

The input data consisted of two sets: One to classify T1 from NT1 and the other to classify T2 from NT2. Each dataset consisted of two features (fixation count and fixation duration) and a target feature of binary class (texting distracted drive and regular drive). Each dataset had 52 samples (26 participants * 2 driving phases). Each dataset was divided into training and test sets with a 60/40 split. The model was trained on a training set using the 5-fold cross-validation. Five-fold was chosen in place of the standard 10-fold to minimize the overfitting issue of the smaller dataset.

A total of four machine learning algorithms were explored as part of this study. Specifically, the decision tree (DT) classifier was used from the tree-based classifiers. The Naïve Bayes (NB) classifier was used from the Bayesian classifiers. From the SVM (Support Vector Machine) classifiers, both linear and radial kernels were selected to use SVM-L (with the linear kernel) and SVM-R (with the radial kernel). The default parameters that the R tool provides for each classifier were used in the analysis.

Table 1. Classification accuracies for detecting texting (T1) for distracted driving. The kappa values are shown within parentheses.

ML Algorithm	Fixation Count (A)	Fixation Duration (B)	(A) + (B)
DT	80% (60%)	95% (90%)	95% (90%)
NB	70% (40%)	95% (90%)	95% (90%)
SVM (L)	75% (50%)	95% (90%)	95% (90%)
SVM (R)	75% (50%)	95% (90%)	95% (90%)

Table 2. Classification accuracies for detecting texting (T2) for distracted driving. The kappa values are shown within parentheses

ML Algorithm	Fixation Count (A)	Fixation Duration (B)	(A) + (B)
DT	75% (50%)	75% (50%)	85% (70%)
NB	70% (40%)	85% (70%)	90% (80%)
SVM (L)	65% (40%)	85% (70%)	90% (80%)
SVM (R)	65% (50%)	85% (80%)	85% (70%)

Table 1 summarizes the classification accuracies. It reports higher accuracy and kappa values for the *fixation duration* feature, indicating that it offers discriminatory solid power between the two drives. For the T1 drive, all four algorithms perform equally well for the *fixation duration* feature but perform suboptimal for the *fixation count* feature.

The same observations can be made from the analysis of the T2 drive, in which the algorithms perform better for the *fixation duration* feature than for the *fixation count* feature.

However, combining *fixation count* with *fixation duration* ((A) + (B)) does improve overall accuracy and the kappa value of the algorithms, specifically for the T2 drive, suggesting that both features should be used in the classification.

Overall, accuracy and the kappa values are higher for the T1 drive than for the T2 drive. The reason for the higher accuracy of T1 is that the T1 distributions for fixation count and fixation duration were more concentrated than that of the T2 distributions, and they were a bit less overlapping with the NT1 distributions than that of T2 overlapping with NT2.

It is interesting to find out what made the participants change their behaviour in the T2 drive. This, however, is out of the scope of this research work and will be considered in future analysis.

5. Conclusion

This paper presents an eye movement analysis for detecting texting and distracted driving. Three types of eye movement parameters (fixation count, fixation duration, and saccadic duration) were analysed. The parameters were compared against everyday driving (without texting); the participants significantly altered their fixation patterns while texting ($p < 0.01$). The cognitive and emotional distractions do not exhibit such a dramatic change in the fixation patterns. The texting distraction affects driving performance, too. The study results reveal that participants significantly altered their driving by lowering driving speed and acceleration to compensate for more considerable lane departure, thus decreasing the risk of an accident.

Furthermore, machine learning algorithms were explored to classify texting distracted drives with 90% and above accuracy. This allows the development of a feedback system to alert drivers in realtime about the possible danger of texting and take away some driving controls if texting continues despite the alert. Of course, more research is necessary to handle this alteration of machine-human interactions in realtime.

Since eye fixation data offers discriminatory solid power, the next step is to extend this offline analysis to a realtime analysis. The development of a machine learning classifier that can predict texting distractions as early as possible could be facilitated by using a mere few seconds of eye gaze data from the start of texting.

Acknowledgments

The authors sincerely thank the Texas Department of Transportation for collecting this data and the Computational Physiology Laboratory at the University of Houston for generously providing the data used in this research. Their support has been invaluable in enabling us to conduct this study, contributing to its depth and quality.

References

[1] Simone Benedetto et al., "Driver Workload and Eye Blink Duration," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 14, no. 3, pp. 199-208, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[2] Tanja Blascheck et al., "State-of-Theart of Visualization for Eye Tracking Data," *Eurographics Conference on Visualization*, pp. 1-20, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[3] Jaeik Jo et al., "Vision-Based Method for Detecting Driver Drowsiness and Distraction in Driver Monitoring System," *Optical Engineering*, vol. 50, no. 12, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[4] Katja Kircher, and Christer Ahlstrom, "Predicting Visual Distraction Using Driving Performance Data," *Annals of Advances in Automotive Medicine/Annual Scientific Conference*, vol. 54, pp. 1-10, 2010. [Google Scholar] [Publisher Link]

[5] Yulan Liang, Michelle L. Reyes, and John D. Lee, "Real-Time Detection of Driver Cognitive Distraction Using Support Vector Machines," *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, no. 2, pp. 340-350, 2007. [CrossRef] [Google Scholar] [Publisher Link]

[6] Yulan Liang, John D. Lee, and Michelle L. Reyes, "Nonintrusive Detection of Driver Cognitive Distraction in Real Time Using Bayesian Networks," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2018, no. 1, pp. 1-8, 2007. [CrossRef] [Google Scholar] [Publisher Link]

[7] Distracted Driving, National Highway Traffic Safety Administration, 2020. [Online]. Available: https://www.nhtsa.gov/risky-driving/distracted-driving

[8] Oskar Palinko et al., "Estimating Cognitive Load Using Remote Eye Tracking in a Driving Simulator," *Proceedings of the 2010 Symposium on Eye-Tracking Research and Applications*, pp. 141-144, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[9] Amira Ragab et al., "A Visual-Based Driver Distraction Recognition and Detection Using Random Forest," *International Conference Image Analysis and Recognition*, pp. 256-265, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[10] Francisco Vicente et al., "Driver Gaze Tracking and Eyes off the Road Detection System," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 4, pp. 2014-2027, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[11] A. Recarte Miguel, and M. Nunes Luis, "Effects of Verbal and Spatial-Imagery Tasks on Eye Fixations While Driving," *Journal of Experimental Psychology: Applied*, vol. 6, no. 1, pp. 31-43, 2000. [CrossRef] [Google Scholar] [Publisher Link]

[12] David L. Strayer, Frank A. Drews, and A. Johnston William, "Cell Phone-Induced Failures of Visual Attention during Simulated Driving," *Journal of Experimental Psychology: Applied*, vol. 9, no. 1, pp. 23-32, 2003. [CrossRef] [Google Scholar] [Publisher Link]

[13] Manbir Sodhi, Bryan Reimer, and Ignacio Llamazares, "Glance Analysis of Driver Eye Movements to Evaluate Distraction," *Behavior Research Methods, Instruments, and Computers*, vol. 34, pp. 529-538, 2002. [CrossRef] [Google Scholar] [Publisher Link]

[14] Trent W. Victor, Joanne L. Harbluk, and Johan A. Engström, "Sensitivity of Eye-Movement Measures to In-Vehicle Task Difficulty," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 8, no. 2, pp. 167-190, 2005. [CrossRef] [Google Scholar] [Publisher Link]

[15] C. David Jenkins, Stephen Zyzanski, and Ray Rosenmen, "Jenkins Activity Survey," *Psychological Corporation*, pp. 1-2, 1979. [Google Scholar] [Publisher Link]

[16] Olga V.L. Bitkina, Jaehyun Park, and Hyun K. Kim, "The Ability of Eye-Tracking Metrics to Classify and Predict the Perceived Driving Workload," *International Journal of Industrial Ergonomics*, vol. 86, 2021. [CrossRef] [Google Scholar] [Publisher Link]