This is the author's accepted manuscript (AAM) of:

Herbers, E., Miller, M., Neurauter, L., Walters, J., & Glaser, D. (2023). Exploratory Development of Algorithms for Determining Driver Attention Status. Human Factors, 0(0). https://doi.org/10.1177/00187208231198932

Exploratory Development of Algorithms for Determining Driver Attention Status

Eileen Herbers[1,2]

Marty Miller[1]

Luke Neurauter[1]

Jacob Walters[1]

Daniel Glaser[3]

[1]Virginia Tech Transportation Institute, Virginia Tech, Blacksburg, VA, 24060, USA

[2]Biomedical Engineering and Mechanics, Virginia Tech, Blacksburg, VA, 24060, USA

[3]General Motors, Detroit, MI, 48243, USA

Acknowledgements: The research team would like to thank General Motors for allowing us to use the database that enabled our research. This project was funded by the Safety through Disruption (Safe-D) National University Transportation Center (UTC), a grant from the U.S. Department of Transportation – Office of the Assistant Secretary for Research and Technology, University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof. A special thanks goes to Eric Glenn, Safe-D UTC coordinator, for his assistance throughout the project in managing our Safe-D deliverables and facilitating Safe-D seminars, which were instrumental in our research.

Corresponding Author: Eileen Herbers, eherbers@vtti.vt.edu

Abstract

Objective: Varying driver distraction algorithms were developed using vehicle kinematics and driver gaze data obtained from a camera-based driver monitoring system (DMS).

Background: Distracted driving characteristics can be difficult to accurately detect due to wide variation in driver behavior across driving environments. The growing availability of information about drivers and their involvement in the driving task increases the opportunity for accurately recognizing attention state.

Method: A baseline for driver distraction levels was developed using a video feed of 24 separate drivers in varying naturalistic-driving conditions. This initial assessment was used to develop four buffer-based algorithms that aimed to determine a driver's real-time attentiveness, via a variety of metrics and combinations thereof.

Results: Of those tested, the optimal algorithm included ungrouped glance locations and speed. Notably, as an algorithm's performance of detecting very distracted drivers improved, its accuracy for correctly identifying attentive drivers decreased.

Conclusion: At a minimum, drivers' gaze position and vehicle speed should be included when designing driver distraction algorithms to delineate between glance patterns observed at high and low speeds. Distraction algorithms should be designed with an understanding of their limitations, including instances in which they may fail to detect distracted drivers, or falsely notify attentive drivers.

Application: This research adds to the body of knowledge related to driver distraction and contributes to available methods to potentially address and reduce occurrences. Machine learning algorithms can build on the data elements discussed to increase distraction detection accuracy using robust artificial intelligence.

Key Words

AUTOMATION, EXPERT SYSTEMS: Trust in automation

COGNITION: Distraction and interruptions

MOTOR BEHAVIOR: Eye movements, tracking

SURFACE TRANSPORTATION: Autonomous Driving; Distraction; Driver Behavior; Vehicle Automation

Précis

Varying driver distraction algorithms were developed using information obtained from a camera-based driver monitoring system (DMS) and vehicle kinematics in real time. This information could be used to inform drivers of their attention to the driving task, as well as to improve performance and benefits associated with driver assistance systems.

Introduction

In 2021, the National Highway Traffic Safety Administration (NHTSA) “project[ed] that an estimated 42,915 people died in motor vehicle crashes” in the US, a 10.5% increase from 2020 (NHTSA, 2022). Additionally, it’s estimated that 3,142 people (8.1%) were killed in motor vehicle crashes involving distracted drivers in 2020 (Stewart, 2022). In 2016, data from police-reported traffic crashes revealed that approximately 9% of all fatal crashes involved some degree of driver distraction (NHTSA, 2018). However, when analyzing events captured via the Second Strategic Highway Research Program (SHRP 2) Naturalistic Driving Study (NDS), distracted driving alone was observed during 68.3% of crashes (Dingus et al., 2016). This difference is important to note since many crashes involving distraction may be underreported due to a driver’s self-interest. This serves as the impetus for this project: to design better algorithms that detect driver distraction using information captured in real time, such as eye tracking and vehicle metrics.

To mitigate the effects of distraction, the Alliance for Automotive Innovation and federal lawmakers are calling for research regarding the use of Driver Monitoring Systems (DMSs) to minimize or eliminate motor vehicle distraction via the SAFE Act of 2021. This is designed to eventually facilitate the requirement of DMSs in vehicles equipped with advanced driver assistance systems (ADAS) (Preston, 2021). Since these systems are still currently being developed, the definition of which functions are included within DMSs changes between researchers, stakeholders, and developers. Generally, a DMS contains (1) data acquisition systems to monitor driver and vehicle behavior, (2) diagnostic functions to determine a driver’s current state, and (3) an assessment and implementation of effective countermeasures. This project focuses on these first two functions to determine the data elements necessary to accurately diagnose a driver’s attention state. So, although DMSs have the potential to alleviate the impacts of distracted driving, the technology does not automatically come with accurate distraction detection and warnings. Even though this project does not focus on assessing and implementing effective countermeasures, it is imperative that the human-machine interface (HMI) that alerts inattentive drivers is just as effective as the distraction detection algorithm. System distrust will result if the DMS does not detect when drivers are distracted, issues too many warnings when the drivers are in fact paying attention, or issues alerts that are too startling, confusing, or do no orient the driver to the driving task (Mehrotra et al., 2022).

The goal of this research, which was part of a larger Safety through Disruption (Safe-D) University Transportation Center (UTC) project (Miller et al., 2022), was to leverage a previous internally developed algorithm to distinguish different levels of driver attentiveness and to determine the key DMS and kinematic variables needed to accurately predict a driver’s attention level.

Background

Distracted driving is difficult to confidently detect due to wide variation in driver behavior and drivers themselves across driving environments. It is important to note that this project uses the terms *inattention* and *distraction* synonymously, even though distraction is a “specific type of inattention that occurs when drivers divert their attention away from the driving task” (NHTSA, 2010). There are occasions where drivers may be looking at the forward roadway but are inattentive or cognitively distracted (i.e., drowsy or daydreaming) (Masala & Grosso, 2014). However, the algorithms in this project are designed to determine when a driver is visually distracted, which is any task or situation that takes the driver’s eyes off the road (NHTSA, 2020). A study using data from the 100-Car NDS found that glances totaling more than 2 seconds off-road increased the crash/near-crash risk by at least two times. Scanning the driving environment (i.e., looking at the rearview mirror and out the side window) is also necessary for safe driving, as long as the driver's eyes return to the forward roadway within 2 seconds (Klauer et al., 2006).

DMSs are becoming a more prevalent source for detecting driver distraction, but the optimal approach for identifying distracted driving has yet to be decided. This project intends to add to the body of knowledge around driver distraction algorithms and provide insight into key variables that support the accurate detection of distracted drivers. The federally funded SAVE-IT program was designed to enhance the effectiveness of safety warning systems and mitigate distraction with effective countermeasures (NHTSA, 2004). This served as the catalyst for many distraction and driver workload algorithms, such as those that used vision-based systems and combined different glance patterns like glance location, glance frequency, glance duration, and eyelid movement (Ahlström & Kircher, 2010; Fernández et al., 2016; Kircher & Ahlstrom, 2018; Kircher et al., 2009; J. D. Lee et al., 2013; Liang et al., 2012; Victor et al., 2008; Yekhshatyan, 2010; Yilu et al., 2004; Zhang et al., 2006). Other research used driving simulator data to determine a driver's attention level and/or their interaction with distraction feedback displays (Donmez et al., 2007; Engström & Mårdh, 2007; J. Lee et al., 2013; Liang & Lee, 2008; Liang et al., 2007; Victor et al., 2005). The AttenD algorithm uses a time buffer that is depleted when the driver looks away from the road and is reset when the driver looks back at the road for a set amount of time (Kircher et al., 2009). Previous research has used this algorithm to determine the cognitive load of different tasks, study secondary behaviors of drivers, and delve further into other in-vehicle distractions (Ahlström et al., 2022; Lee et al., 2017; Seaman et al., 2017; B. Seppelt et al., 2017; Seppelt et al., 2018; B. D. Seppelt et al., 2017; Zhang et al., 2017). The AttenD algorithm was used as the basis for the previous algorithm that this project builds upon, which is described in detail later as V0.

Additionally, including other vehicle parameters and driver behavior measures can create a more robust algorithm when determining a driver's attention state. Smart Eye is one of the leading vision-based technologies used in DMSs, and focuses on expression and emotion analysis, activity and body pose tracking, as well as head, eye, and gaze tracking (SmartEye, 2023). Driver facial affects, lane keeping, steering movements, time headway, and pedal movement were all used in a simulation study to determine distraction, and basic vehicle kinematics can also be used non-invasively to detect driver distraction (Liu et al., 2021; McCall & Trivedi, 2004). It is theorized that distraction algorithms can continually be improved by using a combination of driver and vehicle data to predict driver's intended actions.

There has been extensive research into detecting driver inattention, but the necessary metrics for accurately determining driver distraction levels are still unknown. Prior studies show that including DMS data (glance behavior, etc.) and other vehicle data gives reasonable insight into predicting driver inattention. However, there seems to be a trade-off between accurate true distraction detection and high false positive rates (J. Lee et al., 2013). Therefore, this project investigated how to measure the levels of driver distraction, as well as which methods and data elements are essential toward determining a driver's attention level.

Methods

Naturalistic Database

Naturalistic driving is a research method in which vehicles are instrumented to record time series data under real-world driving conditions. The naturalistic data used for this research were collected through a privately funded (proprietary) field operational test of mid-size Crossovers and featured full time video consisting of external and in-cabin views, kinematic variables (vehicle speed and acceleration), driver input variables (steering angle, brake pressure, and throttle pressure), and driver monitoring data for 24 drivers (12F, 12M, mean age: 44.2 ± 7.3 SD) over an exposure of 8 days each. This research complied with the American Psychological Association Code of Ethics and was approved by the Institutional Review Board at Virginia Tech. Informed consent was obtained from each participant. Glance data were collected through a DMS, which provided 12 coded locations where the driver was looking for each timestamp; these are shown in detail later in Table 2. The DMS consisted of one infrared camera mounted on the steering column with infrared illuminators to ensure that drivers' faces were visible to the camera. In the trips used for this project, no DMS alerts or countermeasures were issued to the drivers, and no L1 or L2 systems were engaged. As part of the original purpose for this privately funded data, the research team developed an initial V0 algorithm for detecting inattentive drivers, which served as the basis for the further development of algorithms in this research and was used to select a subset of the validation data for the ground truth assessment.

Ground Truth Assessment

To assess the accuracy of the attention algorithms developed in this research, the ground truth level was first determined for a subset of the data. Ten-second, non-overlapping epochs were selected, and the driver's attention level was subjectively evaluated at the end of the epoch. This was done by reviewing four camera views simultaneously for the entirety of the event: vehicle driver view, driver lap, forward roadway, and rear-view camera. Ten second epochs were chosen to inform reviewers of the driving environment context, as well as the driver's glance behavior and apparent attentiveness to the driving task. This assessment provided a ground truth distraction level that could be used to compare with the algorithm-based distraction assessment.

Distraction Assessment Methodology

The assessment of distraction levels within the reviewed epochs was the key component of the review process. Table 1 shows the four levels of distraction that were used in the epoch review process and the supporting descriptions used to accurately judge the state of the driver at any given moment. Although subjective in nature, these descriptions included key indicators which facilitated consistent categorization. The most important indicator was the eyes-off-road time (EORT), since heavily distracted drivers generally have a higher proportion of EORT (Dunn et al., 2019).

Table 1. Description of Levels of Attention

Distraction level	Descriptor
Not Distracted	Driver is clearly engaged in the driving task, characterized by glances off-road to locations relevant for safe driving.
Slightly Distracted	Driver is looking away from the roadway more than strictly necessary.
Moderately Distracted	Driver is making more extended glances off-road, sometimes with phone use or longer uses of the center console.
Very Distracted	Typically, driver has combined sources of distraction with prolonged glances off-road to a cell phone and center console.

Choosing Subset of Data

A subset of the full dataset was reviewed to establish this ground truth attention. An initial 960, 10-second epochs were randomly selected equally across the 24 participants included in the dataset and equally across 0–20 mph, 20–40 mph, 40–60 mph, and 60–80 mph speed bins, as speed was a predicted metric of interest. After these events were reviewed per the above criteria, 406 events were added to increase the number of events featuring inattentive drivers (since having very distracted drivers was a rare event). These events were sampled from suspected distracted driving cases detected by the V0 algorithm developed in the previous research effort, and part of the 10-seconds of each epoch were allowed to overlap the original subset. V0 served as a baseline attempt at determining distraction but was determined to falsely predict a high number of distracted cases at low speeds when the driver was in fact attentive to the driving task. Through this review, these events provided an insight into instances in which the V0 algorithm could be improved, which formed the basis for the more refined algorithms developed in this project.

Interrater Reliability

Initially, two researchers each reviewed and rated the same 200 events of the initial 960 events using the definitions in Table 1. These definitions used explicit glance patterns and behaviors to reduce personal bias in determining an overall attention level (since each driver may have a different definition of what levels of distraction look like). This process was completed for an additional 100 events deemed to be distracted via the initial V0 algorithm. A calibration review followed the independent review to ensure that the events were categorized consistently and found that there was a 91% interrater reliability between the two researchers. Therefore, the remaining 1,066 events were split between the researchers to review independently.

Algorithm Development

Buffer-based Algorithms

This project began by building off an initial attempt at a distraction algorithm (V0), which was based in part on methods utilized for the Atten-D algorithm mentioned earlier. In keeping with the similar design of V0, the subsequent algorithms utilized two buffer values: AttentionDuration, and InAttentionDuration. The AttentionDuration buffer increased when the driver was looking at the forward roadway, while the InAttentionDuration buffer increased when the driver was looking away from the forward roadway. The amount that each buffer increased was dependent on the DMS output and differed between algorithms. Figure 1 illustrates the overall flow of information in which the DMS output was used to create the buffer values, which was then used to output the driver's attention status. The following algorithms took this general approach and used a varied combination of the variables listed.

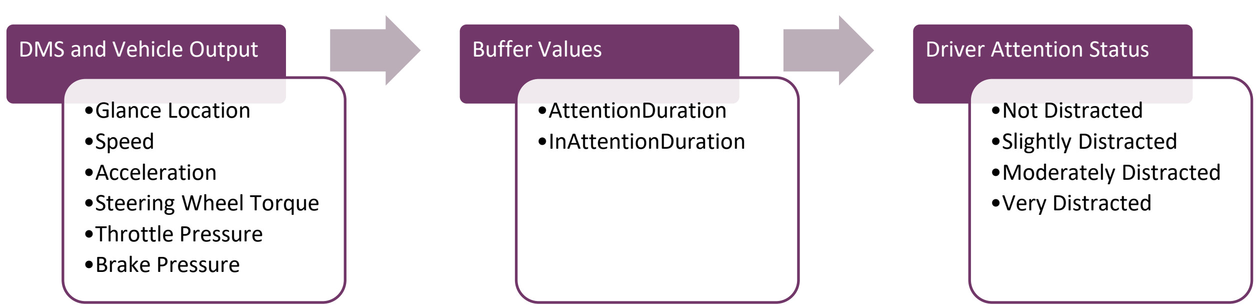

Figure 1. Algorithm flow chart

Previous Algorithm: V0

As mentioned previously, V0 formed the basis of the buffer-based algorithms developed later in this project. This algorithm did not use any vehicle kinematics or driver input data, and only used glance data provided by the DMS. This included the time series of the drivers' glance locations, which were sorted into four main categories: Attentive, Driving Related (inner), Driving Related (outer), and Inattentive. Any unknown glances were categorized as such. As shown in Table 2, each glance type had an associated weighting that impacted its effect on the buffer values; the starred* value denotes an increase in the

AttentionDuration, while all other values increase the InAttentionDuration. As an example, if the driver was looking out the forward roadway for 1 second, the AttentionDuration would increase by 1 and the InAttentionDuration would stay at its current value. If the driver was looking at the right window for 1 second, the InAttentionDuration would increase by 2/3, and the AttentionDuration would stay at its current value. Locations such as the instrument cluster or rearview mirror do not increase the InAttentionDuration as quickly as an off-road glance, since they generally relate to driving. However, longer glances to these locations could classify a driver as "inattentive."

Table 2. Glance Classification and Factor by Glance Location

Glance	Glance Type	Factor
Forward	Attentive	1*
Instrument Cluster Rearview Mirror	Driving Related (inner)	1/3
Right Window Left Window Right Mirror Left Mirror	Driving Related (outer)	2/3
Off Road Center Stack Driver Lap Passenger Footwell	Inattentive	1
Unknown	Unknown	0

Once the AttentionDuration reached a value of 2 (effectively 2 seconds), the driver was classified as "attentive", and the InAttentionDuration buffer was set to 0. Conversely, once the InAttentionDuration reached a value of 0.5, the driver was no longer classified as "attentive" and the AttentionDuration was set to 0. Once the InAttentionDuration value increased over 2.5, the driver was classified as "inattentive." These numbers were based on the 100-car NDS, which found that near-crash/crash risk was doubled when a driver's eyes were off the forward roadway greater than 2 seconds during the 5-second period prior to a crash (starting from the time at which a precipitating event occurred and continuing through to 1-second after the crash) (Klauer et al., 2006). An initial limitation with V0 was the fact that it gave a binary output of either "attentive" or "inattentive." To be able to compare the algorithm output to the ground truth attention assessment, further classifications were done by setting thresholds for the InAttentionDuration buffer: 0.5 seconds classified the driver as slightly distracted, 1.5 seconds was classified as moderately distracted, and 2.5 seconds was classified as very distracted.

Once the output from V0 was adjusted using these thresholds, then it could be directly compared to the ground truth assessment. The distraction levels were assigned four values: Not Distracted (0), Slightly Distracted (1), Moderately Distracted (2), and Very Distracted (3). The mean-squared error (MSE) of distraction levels was calculated individually for each event (n = 1366) and aggregated within each ground-truth assessment category. The first notable finding was that the amount of error made by the algorithms differed based on speed for epochs in which driver was "not distracted". This is visible in Figure 2, which shows the distribution of MSE of distraction levels between the algorithm's attention assessment and the ground truth distraction level of the driver at low speed (< 25 mph) and high speed (> 25 mph). The V0 algorithm resulted in a notable number of false alarms (i.e., the drivers were not distracted, but the algorithm determined that they were). Inspection of these mis-classified epochs revealed that many of these situations took place in low-speed driving situations where glance patterns deviated significantly from those observed at higher speeds. This finding was not surprising given the different glance demands typical between low and high-speed driving environments and indicates a clear area to improve upon in subsequent algorithms.

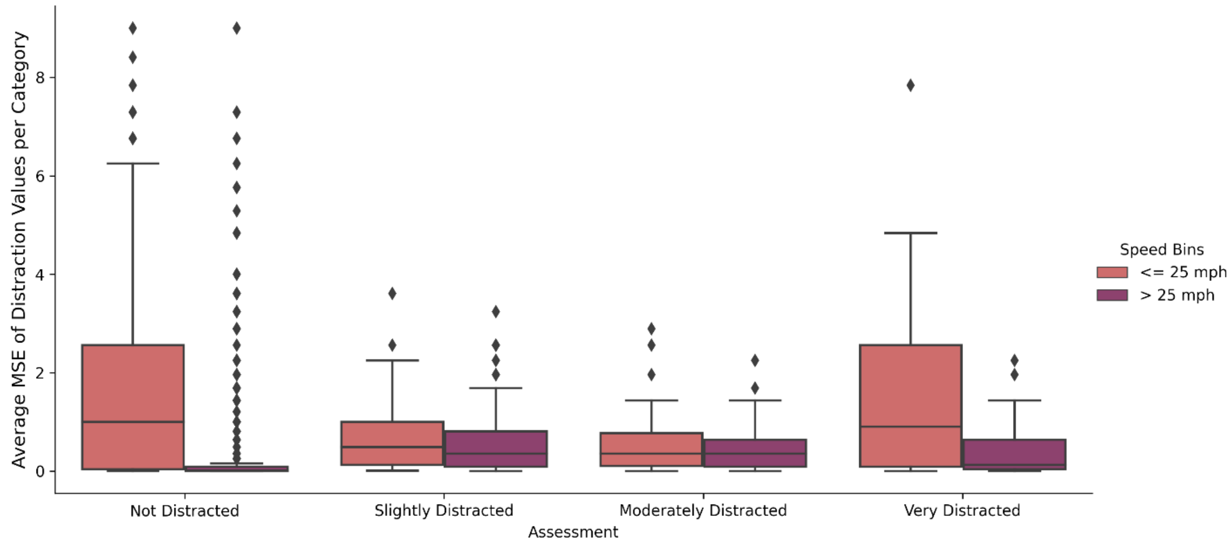

Figure 2. Initial algorithm (V0) MSE of Distraction Levels separated by speed

Incorporating Speed: V1

From the initial assessment of V0, it was apparent that incorporating speed into the algorithm could reduce the number of false positive attention assessments. During the initial video review, it was also found that at extremely low speeds, drivers may be looking at places other than the forward roadway (e.g., pedestrians at crosswalks, or perpendicular traffic) to assess driving risks around the vehicle, which contribute to attentive driving. Thus, the next two algorithms (V1 and V2) incorporated vehicle speed so that the buffer values increased only if the driver was going at least 5 mph. Therefore, if the driver was stopping or going below 5 mph, any off-road glances did not change the attention status of the driver. Additionally, the InAttentionDuration buffer was increased by a factor based on the vehicle speed; at set vehicle speed intervals, the original factor from Table 2 was multiplied by another value between 1 and 2 depending on the speed; then the InAttentionDuration buffer increased more quickly at higher speeds based on inattentive glances. Since an increased driving speed inherently increases a driver's risk of crashing, as well as increases the crash severity, this approach essentially "penalizes" riskier glances away from the road (Elvik, 2005).

Linear Speed Increases: V2

A further variation featured a linearly increasing factor between 25 and 55 mph to avoid large jumps in glance weighting as speed moved between thresholds, as shown in Figure 3. For example, if the driver was looking at a driving related (inner) location (factor of $\overline{0.33}$) and going 40 mph (factor of 1.5) for 1 second, then the InAttentionDuration would increase by $\overline{0.33}$*1.5.

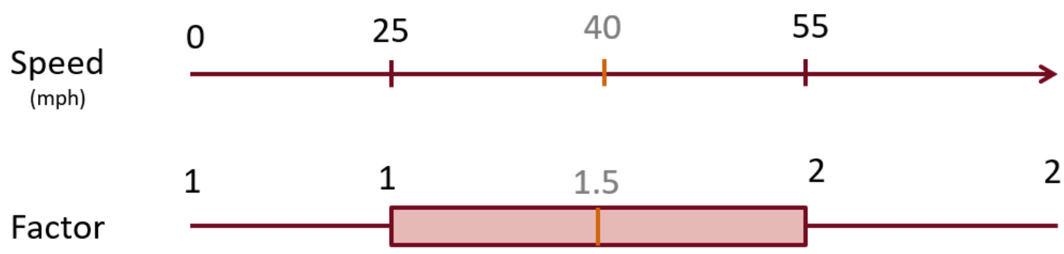

Figure 3. V2 Speed thresholds

Parameter Search: V3

The V1 and V2 algorithm variations were created by examining trends in the data and carefully crafting thresholds, which improved the resulting attention assessment, but this iterative process of determining the correct thresholds is an expensive process (in terms of computing power, computing time, and available data). Therefore, V3 was designed using a random parameter search, which allowed for efficient testing of thousands of parameter combinations, something that is simply not possible to manually perform. The optimization function used the average MSE of distraction levels (within each classification type) to ensure that larger differences between the expected and predicted classification would result in a lower scoring algorithm. This parameter optimization is a basic machine-learning approach but does not have the black box ambiguity of most machine learning methods since the factors applied to each variable are explicitly found.

Twelve individual glance locations (including glance unknown) parameters were included in the search, as well as thresholds for when to reset the In/AttentionDuration variables; thresholds above which a driver would be classified as slightly, moderately, or very distracted; and the speed threshold at which InAttentionDuration would increase by the highest factor. Reasonable selections of values within these parameters offered a potential 3,227,550,665,472,000 unique combinations, and a random search was performed with 20,000 randomly chosen combinations. The parameter value with the lowest median error was chosen for the final set of parameters as the representative buffer-based algorithm, V3.

Attention Filter: Threshold Values to Determine Driver Attention Through Logarithmic Regression.

Although the parameter search described in the previous section chose thresholds for classifying driver attentiveness levels, the research team also explored using aggregated outputs from the algorithm over a 10-second time span to mimic the approach used to determine the ground truth attention levels. To do so, the total maximum and the sum of the local maxima of each of the algorithm buffer values was calculated for 10-second sections of the data. Using an ordinal logistic regression model for the ground truth assessments, it was determined that using a combination of these values was best at determining the driver's attention state:

- AttentionDuration and InAttentionDuration at the end of the event
- MaxAttentionDuration and MaxInAttentionDuration during the 10-second event
- SumAttentionDuration and SumInAttentionDuration during the 10-second event

Therefore, instead of using only two values (AttentionDuration and InAttentionDuration) in determining the driver's attention level, the percentiles of these six aggregate values were used to determine the overall predicted attention state of the event (a continuous value between 0 and 3). This method was applied to all previous models (v0, v1, v2, and v3).

Results and Discussion

The average MSE of distraction levels between the algorithm output and the ground truth assessment was used to assess the accuracy of each algorithm. The attention levels of “Not Distracted,” “Slightly Distracted,” “Moderately Distracted,” and “Very Distracted” were converted into the numerical output of 0, 1, 2, and 3, respectively. Then the events were split into their respective ground truth assessments to determine the algorithm accuracy within each category.

Figure 4 shows the algorithm performance versus the ground truth assessment for each attention category. The x-axis is separated into each category used for the ground truth assessment, and the y-axis shows the average MSE of distraction values for each category and each algorithm. An ideal algorithm would have 0 error, so lower values indicate higher performance in this chart.

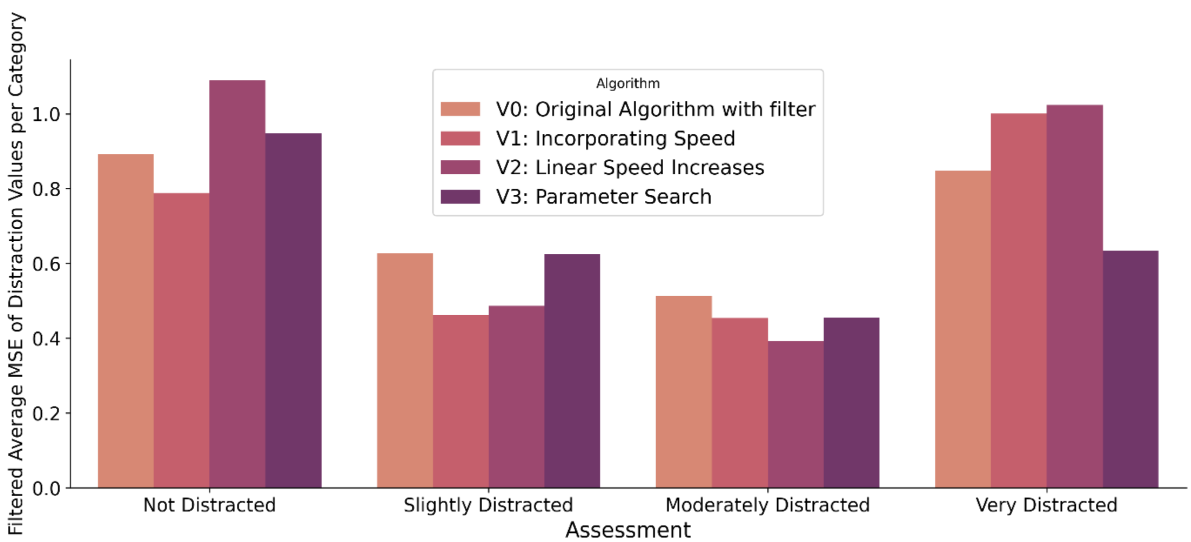

Figure 4. Average MSE of Distraction Values for each algorithm per category.

Because the average MSE changed within algorithms across distraction assessments (and standard deviation calculations did not illustrate the differences well), it was important to find out how many false alarms and misses occurred within each category. Therefore, a new scoring method was developed (Figure 5) to show the percentage of cases where the algorithm was correct, slightly incorrect, or very incorrect. If the algorithm returned the same AttentionStatus as the ground truth assessment, then it was marked as a “hit or correct rejection”; if it returned a different value, then it was marked as a false alarm (for cases where an algorithm incorrectly marked someone as more distracted than they were) or a miss (for cases where an algorithm indicated someone was less distracted than they were). If an algorithm was exactly perfect, then there would be green bars across each level of attentiveness. Grades of false alarms and misses were used to indicate whether the algorithm was incorrect by only one level. For example, a slight false alarm case would occur when the driver was actually “Not Distracted” but the algorithm returned “Slightly Distracted,” and a false alarm would occur when the assessment was off by two or more levels. Since the algorithm and ground truth do not have a binary output of “attentive” or “distracted”, the distraction level at which countermeasures are implemented may differ. Therefore, any slightly incorrect algorithm outputs may not necessarily correspond to an incorrect alert.

[INSERT FIGURE 5]

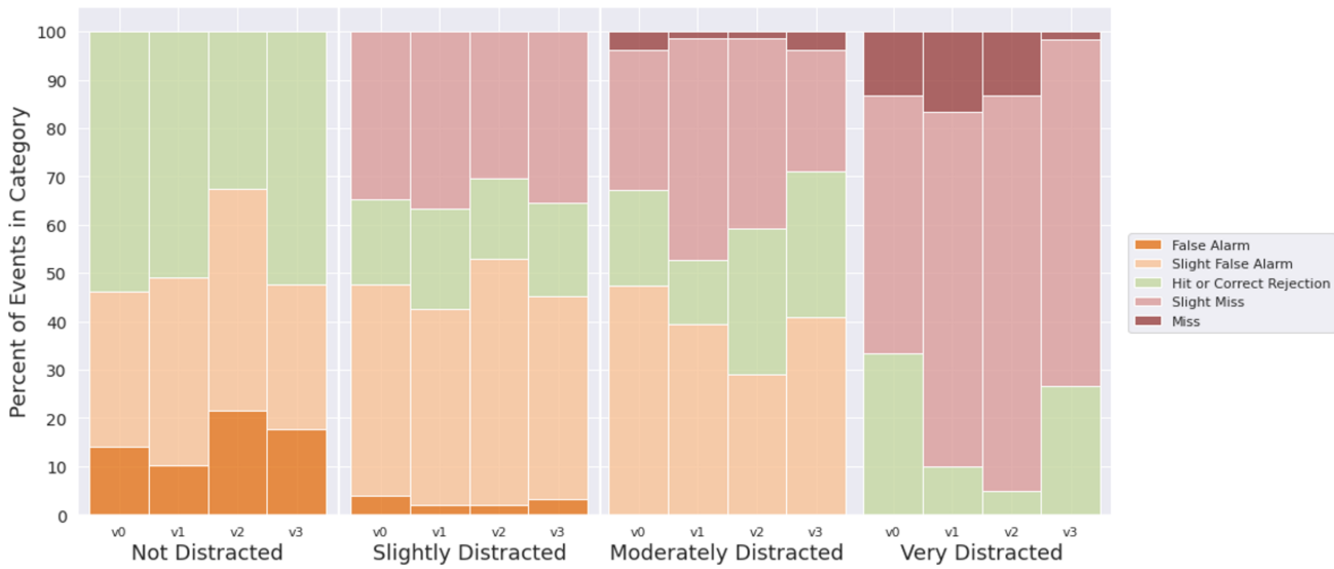

Figure 5. Percent of false alarms, misses, hits, and correct rejections for each algorithm separated by the ground truth assessment.

An average of 47.4% of “Not Distracted” epochs were correctly classified across all algorithms, while an average of 15.9% returned a false alarm (36.7% slight false alarm). If the algorithms had a binary output of distracted vs attentive, then the percentage of correct rejections don’t quite reach even chance performance. The misclassification of “Not Distracted” epochs as “Moderately Distracted” and “Very Distracted” was a recurring problem. Many of these epochs included drivers looking off road when they were about to perform a maneuver (e.g., a lane change or merge) at higher speeds. It could therefore be beneficial for future algorithms to include data elements like steering wheel angle or torque to better indicate maneuver-specific glance patterns, which could increase the percentage of correct rejections.

Correct classifications in the “Slightly” and “Moderately Distracted” epochs ranged from 13.1% to 30.3%, while slight misses and slight false alarms averaged at 34.6% and 41.7%, respectively. An average of 18.8% of “Very Distracted” epochs were correctly classified across all algorithms, while an average of 11.3% of these epochs were missed between all four algorithms (70% slight misses), with V3 having the lowest miss rate of 1.7%. These high error rates are important to note; slight misclassifications may be easy to tackle through differing levels of distraction alerts and countermeasures, but the false positives and misses that are off by more than one level must be addressed before being implemented in commercial vehicles.

Based on the average MSE of distraction levels across each attentiveness level (v0 = 0.720, v1 = 0.676, v2 = 0.748, v3 = .666), the greatest overall performance occurred with algorithm V3. This algorithm achieved the highest level of accuracy, especially in detecting “Moderately” and “Very Distracted” events. Properly assessing driver behavior at all speeds and in all traffic conditions proves very difficult (as seen by the high error rates in Figure 5), but the accuracy observed for true distraction events indicates promise for correctly identifying risky driving behaviors. Therefore, the algorithm must optimize accurately identifying highly distracted cases while also keeping the number of misclassifications low. This is discussed in more detail below.

Conclusion

Accurate assessments of driver attention will continue to be a difficult task even with the technological advancement of DMSs. Driver behaviors vary greatly under different conditions, and within various categories of attentiveness. This research highlights the fact that it is exceedingly difficult to use simple heuristics to assign a level of inattention to a driver based on the available data and a human understanding of what distraction looks like (based on the high error rates). For example, simply looking off-road for a set amount of time does not mean that the driver is inattentive. Additional information about a driver's environment and intended maneuvers is critical to define an attention level. One of the most important discoveries made during this project was regarding the differences in glance behavior at high and low speeds. From the initial video review, we found that drivers may have been looking at places other than the forward roadway at extremely low speeds to assess driving risks around the vehicle (such as pedestrians at crosswalks, or perpendicular traffic). Algorithm error was reduced in V1 and V2 when speed was incorporated. Adding additional variables to the attention algorithm increased algorithm performance only slightly, as shown in V3. Since some algorithm error could be attributed to maneuver-specific glance patterns at high speeds as well, it would be beneficial to add additional variables (i.e., steering wheel angle) that may indicate off-road glances that are still important to the current driving task.

Although this study did not directly tackle the question of driver acceptance of DMSs, it is logical to expect that driver response would be more favorable to systems that accurately assess their behavior. Given this expectation, it is critical that any feedback methods only notify truly inattentive drivers. If there are too many false alarm notifications, drivers may begin to ignore DMS alerts or turn the feature off, which defeats the purpose of the system entirely. Conversely, the purpose of the attentiveness assessment is to notify drivers when they are distracted to reduce any detrimental effects from distracted driving. To do so, it is imperative that the DMS detects all safety-critical distracting events and avoids any misses. The dichotomy between these two requirements creates a challenging problem to solve, in which an algorithm may have to prioritize either over- or under-notifying drivers. Additionally, even if the algorithm is 100% accurate in determining driver distraction, it is crucial that the alert system notifies distracted drivers effectively so that they can correct their behavior. An alert system that does not consider effective HMI design will not be successful at all.

Buffer-based algorithms were used in this study because they have the benefit of being interpretable and intuitive for humans to understand. However, even after developing multiple algorithms, there was little improvement in performance. The next step in this research would be to apply more sophisticated machine learning methods, which would allow for computer algorithms to determine the best output based on the input (for instance steering angle/driver torque, brake, and throttle pressure) and the ground truth. An appropriate next step in algorithm assessment would be increasing the amount of available training data to develop robust machine learning algorithms that can more easily incorporate more variables. An important piece of this project was using naturalistic data to develop a ground-truth assessment to compare algorithm output to real distracted drivers. Not only would performance likely improve by using more data containing distracted drivers, but integrating additional data sources into deep learning models is considerably less time-consuming than using the buffer-based models developed here.

Application

DMSs have the potential to be an important component in reducing inattention and crashes related to distraction. The results from these algorithms indicate that both gaze location and vehicle speed, at a minimum, should be used to assess driver distraction. Speed is particularly important for accurately establishing a driver's attentiveness level due to large deviations in acceptable glance patterns at low and high speeds. Designing a single algorithm across a wide range of speeds without incorporating changes based on speed itself is problematic based on our review of the described benchmark events.

An accurate distraction algorithm also has the potential to give insight into what causes these distracting behaviors. If new technology and HMIs in automated driving systems (ADS) are causing drivers to be more distracted, accurate distraction algorithms can also be used as a research tool to detect how in-vehicle technologies affect drivers' attention.

As mentioned in the introduction, components of a DMS can be broken down into the (1) data acquisition system (DAS), (2) distraction detection algorithm, and (3) distraction warnings and countermeasures. It is important to understand that no DAS is likely to be 100% accurate, and therefore algorithms making use of this data should be designed with an understanding of the limitations of this underlying data. Specifically, the data used in this project did not take drivers' race or ethnicity into account, but it is important to note that there are substantial disparities in the accuracy of facial analysis between races and genders (Buolamwini & Gebru, 2018). Also, any implemented countermeasures should take any potential distraction detection algorithm error into account. If a system is to gain a driver's trust, then it must correspond with the driver's own idea of distraction to an extent. Drivers engaging in distracting behaviors (e.g., texting, eating, or other non-driving related tasks) may argue that they are not distracted, and will thus not even agree with an accurate alert system. These cases require learning more about the underlying causes of distraction and addressing effective countermeasures to avoid them.

As driver assistance features continue to advance and L2 and L3 ADAS become more commonplace, the driving context and data available to underlying distraction detection algorithms also change. This should have an impact on the design of algorithms for detecting distraction since the level of automation and required manual input from the driver will change the available data. For example, an attention algorithm can no longer make use of changes in speed, lane position, and steering wheel torque when the vehicle is "in charge" of maintaining the vehicle's positioning and speed. Additionally, the way that distraction is mitigated may also change. For example, a forward-collision warning (FCW) can be issued sooner if the DMS detects that the driver is not looking at the forward roadway. Therefore, it is crucial that the vehicle environment and driver role is well understood and defined.

Finally, the models developed in this research did not account for cognitive distractions in which the driver is looking at the forward roadway, but their mind might be wandering elsewhere. Algorithms should also include a required scanning of the environment to maintain situational awareness to account for cognitive distractions. Counter to this, all glances away from the forward roadway were penalized in the algorithms developed, though likely driving-related glances were penalized to a lesser extent. Since scanning mirrors and surrounding traffic is a necessary component of safe driving, this is potentially insufficient for identifying truly attentive drivers. To further develop attention algorithms, future work could explore instances in which attentive drivers are ready to resume control of more highly automated vehicles. This was considered out of scope for the current research effort as the focus was on detecting driver distraction rather than levels of readiness or positive attention.

Biographies

Eileen Herbers, Graduate Researcher, Virginia Tech Transportation Institute, Virginia Tech

B.S. Engineering Science and Applied Mathematics, 2018, Northwestern University

Eileen Herbers is a PhD candidate in the Biomedical Engineering and Mechanics department at Virginia Tech. She has been conducting her research at the Virginia Tech Transportation Institute for the past 3 years and has helped with a variety of projects. Her dissertation focus is on developing measures to assess automated driving system safety through modeling and computational physics. Her previous research projects have included measuring the potential V2X impact on reducing crash and near-crash events and using traffic cameras to calculate kinematic measures of vehicles on the road in real time. She is also heavily involved in the Virginia Tech chapter of Women in Transportation (WTS), which is focused on advancing women in the field of transportation.

Marty Miller, Research Associate, Virginia Tech Transportation Institute, Virginia Tech

B.A. History (minor Computer Science), 2013, Berea College.

As a research associate in the Division of Vehicle, Driver & System Safety, Marty Miller splits his time between conducting human factors research for industrial proprietary projects for original equipment manufacturers and developing technologies to assist in the completion of such research. His areas of interest include: advanced driver assistance systems, automated vehicles, C-V2X, and convenience features. Marty's technological skills enable the development of novel research techniques and support fast, efficient data collection and analysis. He has a passion for automotive safety research and applies his skills and expertise to tackling the unique and challenging problems in this space. Before taking his current position, Miller spent a year living and working in Germany on a Fulbright English Teaching Assistantship.

Luke Neurauter, Division Director, Virginia Tech Transportation Institute, Virginia Tech

M.S. Industrial & Systems Engineering, Human Factors Focus, 2004, Virginia Tech

As director of VTTI's Division of Vehicle, Driver, & System Safety, Luke Neurauter leads a multifaceted team of transportation researchers encompassing a broad coverage of safety-focused topic areas. Luke's work consists primarily of gathering and analyzing human factors-related data to evaluate prototype concepts and advanced technologies. His research areas have traditionally focused on advanced vehicle technologies and active safety (collision avoidance) systems, evaluating how drivers comprehend and interact with these systems through both controlled and naturalistic exposure. Mental model development and driver response to both staged and naturally occurring events catered to the specific system being tested are routinely assessed and analyzed as part of these efforts.

Jacob Walters, Senior Research Specialist, Virginia Tech Transportation Institute, Virginia Tech

A.S. General Studies, New River Community College

Jacob Walters has been a researcher at the Virginia Tech Transportation Institute since 2014. He began with studies in infrastructure lighting, emergency vehicle lighting and retro reflectivity and has moved on to work on dozens of proprietary projects with a focus on connected vehicles, automation, and other advanced safety features. His role within those projects has ranged from primary data collection, data analysis, reporting, standards (FMVSS), and presentations. He is also responsible for practical safety driver exams for staff at VTTI.

Daniel Glaser, Senior HMI Researcher, General Motors

Ph.D. Human-Computer Interaction, 2012, Rice University

Daniel Glaser is the Driver Workload Technical Lead at General Motors. In this capacity he leads a validation body that ensures infotainment applications are consistent with internal and industry-wide driver workload guidelines. His research has largely focused on the impacts of emerging infotainment interactions and advanced driver assist systems on driver behaviors and how these systems may be leveraged to promote positive safety outcomes. Daniel conducts research leveraging an array of data sources ranging from simulator, closed course, and naturalistic studies. Dan holds a B.S. in Psychology from Texas Lutheran University (2000), as well as an M.A. (2008) & Ph.D. (2012) in Psychology from Rice University. Daniel joined General Motors in 2011.

Précis

Varying driver distraction algorithms were developed using information obtained from a camera-based driver monitoring system (DMS) and vehicle kinematics in real time. This information could be used to inform drivers of their attention to the driving task, as well as to improve performance and benefits associated with driver assistance systems.

Key Points

- Driver monitoring is an important component in detecting and potentially reducing distractions, and Driver Monitoring Systems (DMSs) make it possible to determine when this occurs, but DMSs alone require the correct usage of data elements to do so.
- Algorithms used to determine driver distraction should be designed with an understanding of their limitations: in terms of specific instances when the algorithm might fail to recognize a distracted driver, or areas in which the algorithm may produce high false positive distraction assessments.
- At a minimum, both glance location and speed should be used to assess driver distraction since glance patterns vary greatly between high and low speeds.

References

Ahlström, C., Georgoulas, G., & Kircher, K. (2022). Towards a Context-Dependent Multi-Buffer Driver Distraction Detection Algorithm. *IEEE Transactions on Intelligent Transportation Systems*, *23*(5), 4778-4790. https://doi.org/10.1109/TITS.2021.3060168

Ahlström, C., & Kircher, K. (2010). *Review of real-time visual driver distraction detection algorithms* Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research, Eindhoven, The Netherlands. https://doi.org/10.1145/1931344.1931346

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research.* Conference on Fairness, Accountability, and Transparency.

Dingus, T. A., Guo, F., Lee, S., Antin, J. F., Perez, M., Buchanan-King, M., & Hankey, J. (2016). Driver crash risk factors and prevalence evaluation using naturalistic driving data. *Proceedings of the National Academy of Sciences*, *113*(10), 2636-2641. https://doi.org/doi:10.1073/pnas.1513271113

Donmez, B., Boyle, L. N., & Lee, J. D. (2007). Safety implications of providing real-time feedback to distracted drivers. *Accid Anal Prev*, *39*(3), 581-590. https://doi.org/10.1016/j.aap.2006.10.003

Dunn, N., Dingus, T., & Soccolich, S. (2019). *Understanding the Impact of Technology: Do Advanced Driver Assistance and Semi-Automated Vehicle Systems Lead to Improper Driving Behavior?* (Report No. 19-0460). AAA: Foundation for Traffic Safety https://aaafoundation.org/wp-content/uploads/2019/12/19-0460_AAAFTS_VTTI-ADAS-Driver-Behavior-Report_Final-Report.pdf

Elvik, R. (2005). Speed and Road Safety: Synthesis of Evidence from Evaluation Studies. *Transportation Research Record*, *1908*(1), 59-69. https://doi.org/10.1177/0361198105190800108

Engström, J., & Mårdh, S. (2007). *SafeTE Final Report* (Report No. 2007:36). Vägverket. http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1363665&dswid=1300

Fernández, A., Usamentiaga, R., Carús, J. L., & Casado, R. (2016). Driver Distraction Using Visual-Based Sensors and Algorithms. *Sensors*, *16*(11), 1805. https://www.mdpi.com/1424-8220/16/11/1805

Kircher, K., & Ahlstrom, C. (2018). Evaluation of methods for the assessment of attention while driving. *Accident Analysis & Prevention*, *114*, 40-47. https://doi.org/https://doi.org/10.1016/j.aap.2017.03.013

Kircher, K., Ahlstrom, C., & Kircher, A. (2009). Comparison of Two Eye-Gaze Based Real-Time Driver Distraction Detection Algorithms in a Small-Scale Field Operational Test. *Proceedings of the Fifth International Symposium on Human Factors in Driver Assessment, Training and Vehicle Design.*

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). *The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data* (DOT HS 810 594). NHTSA. https://psycnet.apa.org/doi/10.1037/e729262011-001

Lee, J., Moeckli, J., Brown, T., Roberts, S., Victor, T., Marshall, D., Schwarz, C., & Nadler, E. (2013). *Detection of Driver Distraction Using Vision-Based Algorithms.* 23rd International Technical Conference on the Enhanced Safety of Vehicles, Seoul, South Korea. https://www-esv.nhtsa.dot.gov/Proceedings/23/files/23ESV-000348.PDF

Lee, J., Sawyer, B. D., Mehler, B., Angell, L., Seppelt, B. D., Seaman, S., Fridman, L., & Reimer, B. (2017). Linking the Detection Response Task and the AttenD Algorithm Through Assessment of Human–Machine Interface Workload. *Transportation Research Record*, *2663*(1), 82-89. https://doi.org/10.3141/2663-11

Lee, J. D., Moeckli, J., Brown, T. L., Roberts, S. C., Schwartz, C., Yekhshatyan, L., Nadler, E., Liang, Y., Victor, T., Marshall, D., & Davis, C. (2013). *Distraction Detection and Mitigation Through Driver Feedback* (DOT HS 811 547A). NHTSA. https://www.nhtsa.gov/sites/nhtsa.gov/files/811547a.pdf

Liang, Y., & Lee, J. D. (2008). *Comparing Support Vector Machines (SVMs) and Bayesian Networks (BNs) in detecting driver cognitive distraction using eye movements.* Passive eye monitoring: Algorithms, applications and experiments, Leipzig, Germany: Springer-Verlag Berlin Heidelberg.

Liang, Y., Lee, J. D., & Reyes, M. L. (2007). Nonintrusive Detection of Driver Cognitive Distraction in Real Time Using Bayesian Networks. *Transportation Research Record*, *2018*(1), 1-8. https://doi.org/10.3141/2018-01

Liang, Y., Lee, J. D., & Yekhshatyan, L. (2012). How Dangerous Is Looking Away From the Road? Algorithms Predict Crash Risk From Glance Patterns in Naturalistic Driving. *Human Factors*, *54*(6), 1104-1116. https://doi.org/10.1177/0018720812446965

Liu, Z., Ren, S., & Peng, M. (2021). Identification of Driver Distraction Based on SHRP2 Naturalistic Driving Study. *Mathematical Problems in Engineering*, *2021*. https://doi.org/10.1155/2021/6699327

Masala, G. L., & Grosso, E. (2014). Real time detection of driver attention: Emerging solutions based on robust iconic classifiers and dictionary of poses. *Transportation Research Part C: Emerging Technologies*, *49*, 32-42. https://doi.org/https://doi.org/10.1016/j.trc.2014.10.005

McCall, J. C., & Trivedi, M. M. (2004, 3-6 Oct. 2004). Visual context capture and analysis for driver attention monitoring. The 7th International IEEE Conference on Intelligent Transportation Systems (IEEE Cat. No.04TH8749),

Mehrotra, S., Wang, M., Wong, N., Parker, J. i., Roberts, S. C., Kim, W., Romo, A., & Horrey, W. J. (2022). *Human-Machine Interfaces and Vehicle Automation: A Review of the Literature and Recommendations for System Design, Feedback, and Alerts* (AAA Foundation for Traffic Safety) Safe SIM. https://aaafoundation.org/wp-content/uploads/2022/11/HMI-and-Automation-Design-Recommendations.pdf

Miller, M., Herbers, E., Walters, J., Neurauter, L., & through Disruption, S. (2022). *Improving Methods to Measure Attentiveness Through Driver Monitoring* (Report No. 05-019). Safety Through Disruption (Safe-D) University Transportation Centers (UTC), VTTI. https://safed.vtti.vt.edu/projects/improving-methods-to-measure-attentiveness-through-driver-monitoring/

NHTSA. (2004). *Reducing Distraction-Related Automotive Crashes with SAVE-IT*. Volpe Center Highlights. https://www.volpe.dot.gov/safety-management-and-human-factors/surface-transportation-human-factors/reducing-distraction

NHTSA. (2010). *Driver Distraction Program* (DOT HS 811 299). National Highway Traffic Safety Administration, Distraction.Gov. https://www.nhtsa.gov/sites/nhtsa.gov/files/811299.pdf

NHTSA. (2018). *Distracted Driving 2016* (DOT HS 812 517). National Highway Traffic Safety Administration. https://crashstats.nhtsa.dot.gov/Api/Public/Publication/812517#:~:text=There%20were%203%2C157%20fatal%20crashes,more%20than%20one%20distracted%20driver.

NHTSA. (2020). A Highway Safety Countermeasure Guide for State Highway Safety Offices. (10th Edition ed.). National Highway and Traffic Safety Administration. https://www.nhtsa.gov/book/countermeasures/countermeasures-work

NHTSA. (2022). Early Estimates of Motor Vehicle Traffic Fatalities and Fatality Rate by Sub-Categories in 2021. In *Traffic Safety Facts: Crash Stats*: US Department of Transportation.

Preston, B. (2021). Auto Industry and Lawmakers Call for Driver Monitoring Systems to Improve Safety. Retrieved Jan 15, 2023, from https://www.consumerreports.org/car-safety/call-for-driver-monitoring-to-improve-car-safety-a3949626347/

Seaman, S. R., Lee, J., Seppelt, B. D., Angell, L. S., Mehler, B., & Reimer, B. (2017). *It's All in the Timing: Using the Attend Algorithm to Assess Texting in the Nest Naturalistic Driving Database.* 9th

International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Manchester Village, Vermont.

Seppelt, B., Seaman, S., Angell, L., Mehler, B., & Reimer, B. (2017). *Differentiating Cognitive Load Using a Modified Version of AttenD.* Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Oldenburg, Germany. https://doi.org/10.1145/3122986.3123019

Seppelt, B. D., Seaman, S., Angell, L. S., Mehler, B., & Reimer, B. (2018). Assessing the effect of in-vehicle task interactions on attention management in safety-critical events. 6th Int. Conf. on Driver Distraction and Inattention,

Seppelt, B. D., Seaman, S., Lee, J., Angell, L. S., Mehler, B., & Reimer, B. (2017). Glass half-full: On-road glance metrics differentiate crashes from near-crashes in the 100-Car data. *Accident Analysis & Prevention*, *107*, 48-62. https://doi.org/https://doi.org/10.1016/j.aap.2017.07.021

SmartEye. (2023). *Smart Eye Technology*. Retrieved Jan, 21 2023 from https://smarteye.se/technology/#:~:text=By%20using%20sensors%20that%20detect,overall%20mood%20and%20awareness%20level.

Stewart, T. (2022). *Overview of Motor Vehicle Crashes in 2020* (DOT HS 813 266). National Highway Traffic Safety Administration. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813266

Victor, T., Harbluk, J., & Engström, J. (2005). Sensitivity of eye-movement measures to in-vehicle task difficulty. Transport Res F: Traffic Psychol Behav 8:167-190. *Transportation Research Part F: Traffic Psychology and Behaviour*, *8*, 167-190. https://doi.org/10.1016/j.trf.2005.04.014

Victor, T. W., Engström, J., & Harbluk, J. L. (2008). Distraction Assessmnet Methods Based on Visual Behavior and Event Detection. In *Driver Distraction: Theory, Effects, and Mitigation* (pp. 137-165). CRC Press. https://trid.trb.org/view/884613

Yekhshatyan, L. (2010). *Detecting distraction and degraded driver performance with visual behavior metrics* [Doctoral dissertation, University of Iowa].

Yilu, Z., Owechko, Y., & Jing, Z. (2004, 3-6 Oct. 2004). Driver cognitive workload estimation: a data-driven perspective. Proceedings. The 7th International IEEE Conference on Intelligent Transportation Systems (IEEE Cat. No.04TH8749)

Zhang, H., Smith, M. R., & Witt, G. J. (2006). Identification of real-time diagnostic measures of visual distraction with an automatic eye-tracking system. *Hum Factors*, *48*(4), 805-821. https://doi.org/10.1518/001872006779166307

Zhang, Y., Angell, L., Pala, S., Hara, T., & Vang, D. (2017). *Can You Still Look Up? Remote Rotary Controller vs. Touchscreen,* SAE International https://doi.org/10.4271/2017-01-1386