




# A Novel Approach to Identify Distracted Drivers: A Case Study in New Jersey

Authors:
Dr. Mohammad Jalayer, Rowan University
Mr. Ahmed Sajid Hasan, Rowan University
Mr. Deep Patel, Rowan University
Ms. Ruqaya Alfaris, Rowan University

December 2021

# Table of Contents

**DISCLAIMER** ..... **2**
**ABSTRACT** ..... **3**
**EXECUTIVE SUMMARY** ..... **4**
**CHAPTER 1: Introduction** ..... **6**
*Background* ..... *6*
*Research Objective* ..... *9*
**CHAPTER 2: Literature Review** ..... **10**
*Research Methodology* ..... *10*
*Data Collection* ..... *11*
*Data Analysis* ..... *15*
*Contributing Factors* ..... *17*
*Safety Countermeasures* ..... *18*
*Conclusion* ..... *22*
**CHAPTER 3: Analysis of Distracted Driving Crashes in New Jersey Using a Mixed Logit Model** ... **24**
*Methods and Data* ..... *25*
*Results* ..... *28*
*Discussion* ..... *31*
*Conclusion* ..... *34*
**CHAPTER 4: Identifying Distracted Driving Events in New Jersey Using Floating Car Method** ..... **36**
*Introduction* ..... *36*
*Methods* ..... *37*
*Results* ..... *51*
**CHAPTER 5: Detection of Distracted Driving Using Artificial Intelligence (AI)** ..... **69**
*Introduction* ..... *69*
*Methodology* ..... *69*
*Results and Discussions* ..... *77*
*Conclusion* ..... *79*
**CHAPTER 6: Conclusion and Recommendations** ..... **81**
*Introduction* ..... *81*
*Summary of Results* ..... *81*
*Recommendations and Future Work* ..... *83*
**REFERENCES** ..... **84**

## DISCLAIMER

## ABSTRACT

The main objective of this study is to identify the actual distracted driving events in the state of New Jersey. To this end, driver behavior was observed and recorded using video cameras along ten major highway corridors in New Jersey. This report provides a detailed analysis of the state's crash events involving distraction over the past five years. The event data were analyzed to identify the significance of various temporal features and geometric properties of roadways on the rate of distraction. The video data was analyzed to detect distracted driving using artificial intelligence. Together, this study reflects the status and pattern of distraction events across the selected corridors in New Jersey.

# EXECUTIVE SUMMARY

Every year, thousands of people die in the United States due to crashes involving distracted driving, with this cause contributing to almost 25% of all fatal traffic crashes in New Jersey in particular (Zutobi, 2022). Over the past several years, various techniques (e.g., surveys, crash reports, videos, and simulations) were developed and implemented by the transportation safety community to identify and evaluate distracted driving events. However, these methods collect cross-sectional data on individual subjects and do not provide the actual number of distractions on the road. To fill this gap, this study collected longitudinal data on distracted driving events in the state of New Jersey. The method involved a data collection crew continuously driving through the selected corridors to track driver distraction events by manual counting and video recording. The event data on distracted driving was analyzed to find the significance of various temporal features and geometric properties of roadways on the rate of distraction. The video data from the observational study was utilized to detect driving behaviors using a deep learning algorithm. This study also analyzed the historical crashes in the state of New Jersey to identify the factors contributing to the severity of crashes involving distracted driving. Some of the key findings of the crash analysis include:

- The proportion of young drivers in fatal crashes caused by distractions (13.7%) was less than for older drivers (18.2%).
- The likelihood of “Injury” crashes decreased by 21.9% and 43.1% for higher values of annual average daily traffic (AADT) and urban setting, respectively, i.e., more traffic volume or congestion leads to less propensity for serious injury.
- An increase in the “speed limit” decreased the probability of a “no injury” crash to a great extent (36.7%).
- The “Injury” severity level is more likely to increase for crashes during the early morning (midnight to 6:30 a.m.), for drivers under the influence of substances (e.g., alcohol or drugs), and for driving on wet surfaces.

The results from the analysis of distracted driving event data demonstrated that the number of distractions was significantly affected by both the time of day and by roadway type. To be specific, some of the key findings of the distraction driving event analysis include:

- Cellphone use is the most prominent type of distraction.
- Summer had a higher rate of distractions than spring.

- The “receiving calls” events significantly increased during the weekdays, on the unsignalized corridors, and non-toll roads compared to the weekends, signalized, and toll roads.
- The behavior of “eating/drinking” significantly increased on the signalized corridors compared to unsignalized corridors, and during the summer compared to the spring.
- The “radio/reaching object” distraction events significantly increased on the weekdays compared to the weekend.
- The “fidgeting/grooming” distraction events significantly increased during the summer season compared to the spring.
- An increase in speed limit significantly increased the distractions, while an increase in the number of lanes significantly decreased the distraction events.
- The “curbed” median encountered a significant reduction in distractions compared to the “unprotected” and “positive” medians.
- An increase in median width significantly decreased distractions, while an increase in shoulder width significantly increased distractions.

It is expected that the results obtained from this study will further assist state and local agencies in promoting awareness and reducing distracted driving in New Jersey. For instance, law enforcement and awareness campaigns should be focused more on the summer months. In addition, speeding should be adequately monitored since an increase in speed significantly increases the distraction events and increases the likelihood of being involved in a serious injury crash.

# CHAPTER 1: Introduction

## Background

Distracted driving is one of the major traffic concerns of the 21st century. According to the National Highway Traffic Safety Administration (NHTSA), a distracted driving event is "*Anything that takes the driver's attention away from the task of safe driving.*" (NHTSA, 2019). According to traffic safety experts, distraction events typically include the following three types (Regan, 2007):

- Visual—taking eyes off the road (e.g., while texting or talking to passengers).
- Manual—taking hands off the wheel (e.g., while texting, receiving calls, tuning the radio, or eating).
- Cognitive—taking their mind off what they are doing (e.g., when texting, receiving calls, or drowsy).

Distraction can include a wide variety of behaviors and activities, such as using a cellphone, eating and drinking, talking to passengers, grooming, reading, using an advanced traveler information system, watching a video, changing the radio station, switching music, monitoring children, getting lost in thought, and smoking.

Any non-driving activity is a potential distraction that can result in a high-profile crash. Nationwide in 2019, distracted driving was one of the top five factors contributing to fatal motor vehicle crashes (FARS, 2020; IIHS, 2020). Each year, thousands of people lose their lives in crashes resulting from distracted driving in the United States. The NHTSA reported 3,142 fatal crashes in the United States in 2019 due to this cause, which was 9.9% more than the 2,841 in 2018 (NHTSA, 2020). Figure 1 illustrates the trend of traffic crash fatalities involving distracted driving in the United States over a ten-year period (2010–2019). According to this figure, around 3,000 people die in motor vehicle crashes due to distraction each year.

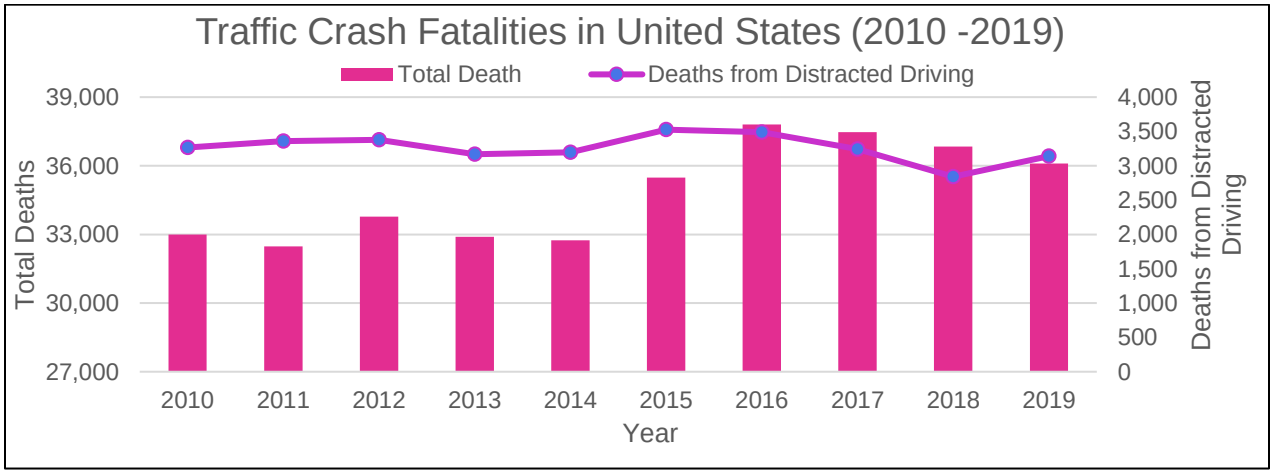



**Figure 1.** *Total Traffic Crash Fatalities in the United States (2010–2019) (Source: NHTSA, 2014; NHTSA, 2018; NHTSA, 2020)*

It should be noted that distracted driving is the leading cause of fatal crashes in New Jersey, accounting for almost 25% of the fatal motor vehicle crashes in the state, and ranking second-highest among all the states (FARS, 2020). In 2019, distracted driving caused 159 of the 524 fatal crashes in New Jersey that claimed 558 lives (New Jersey State Police, 2019). Figure 2 depicts the percentage of fatal motor vehicle crashes involving distracted driving across the country. According to this figure, New Jersey is among the top five states, each of which has experienced a rate of more than 15% fatal motor vehicle crashes being due to distracted driving.

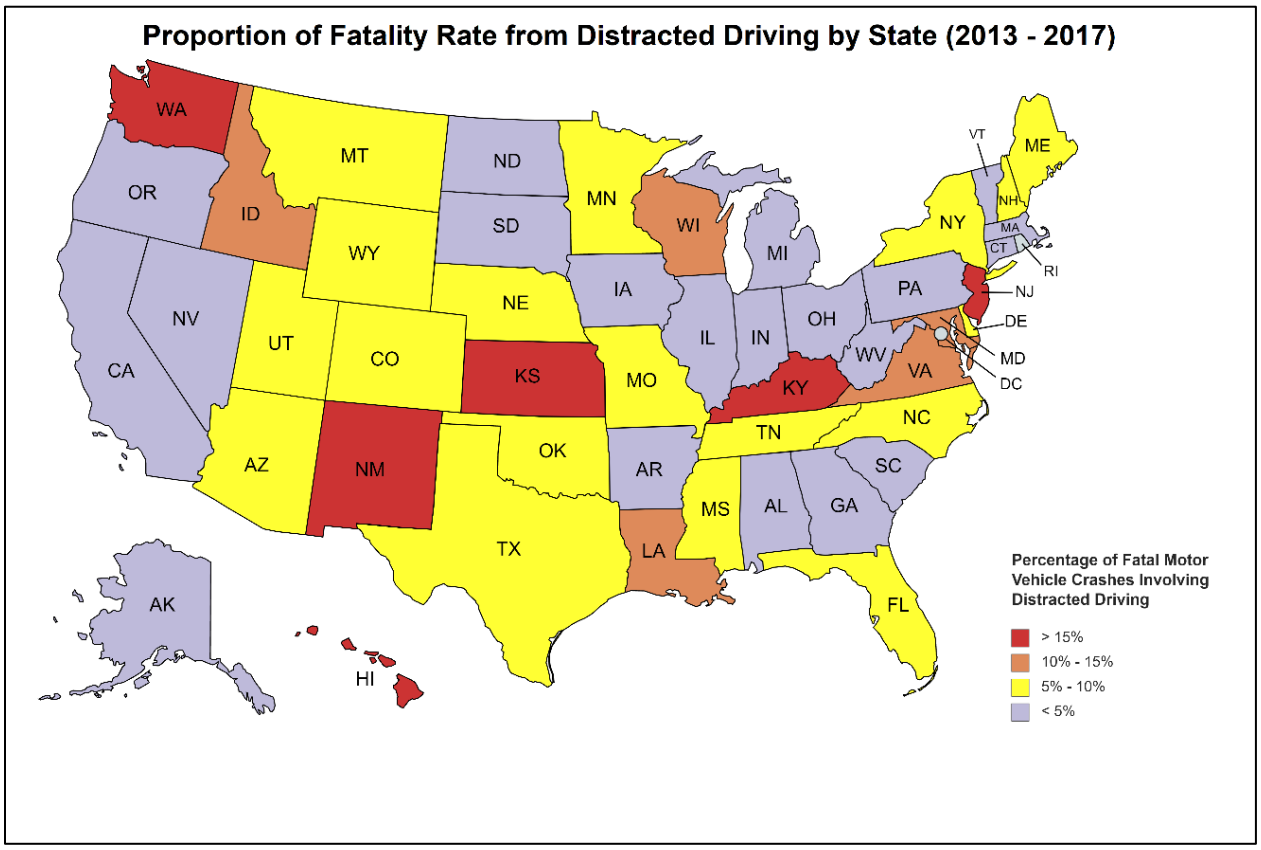



**Figure 2.** *The Proportion of Motor Vehicle Fatalities Involving Distracted Driving by State (Source: FARS, 2020)*

Figure 3 illustrates the total number of fatal crashes involving distracted driving in New Jersey over the last ten years (NJDHTS, 2021). We can observe that at least 500 fatal crashes occurred annually in New Jersey due to distractions (such as receiving calls, texting, and eating). According to this figure, this number went up to more than 550 fatal crashes between 2015 and 2019, necessitating further investigation to reduce the frequency and severity of such crashes.

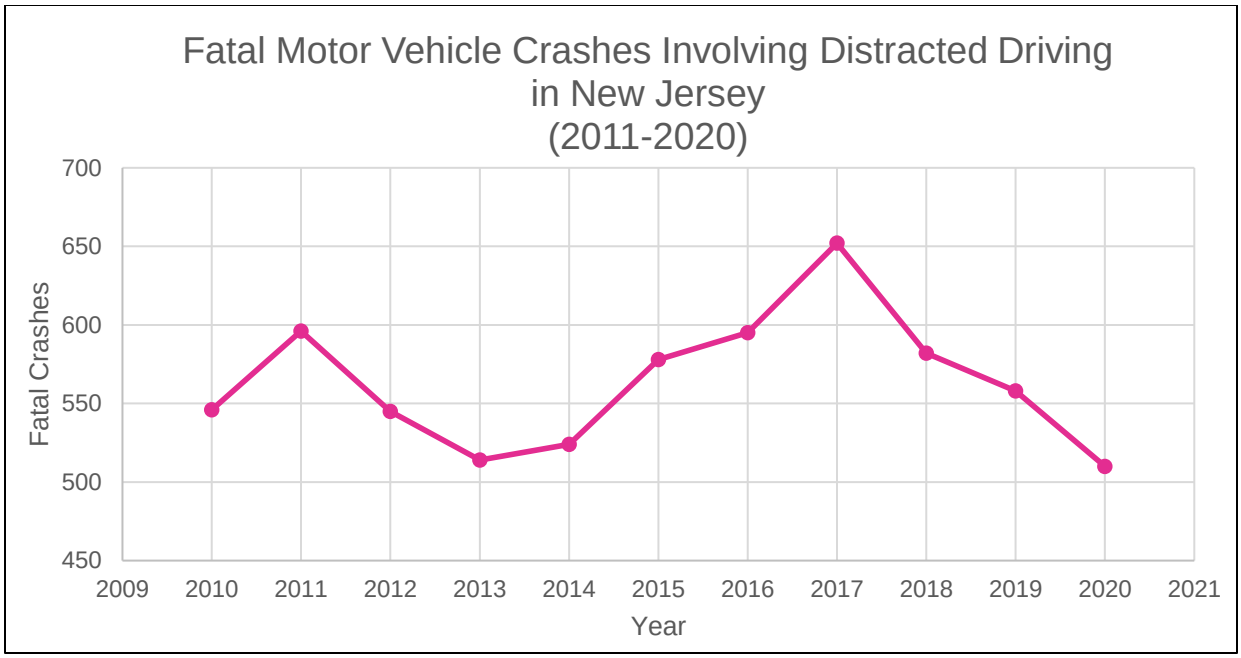



**Figure 3.** *Fatal Crashes Involving Distracted Driving in New Jersey (2010–2020) (Source: NJDHTS, 2021)*

It should be noted that, over the past few years, a number of safety countermeasures have been developed and implemented by state departments of transportation and local agencies across the nation to reduce crashes involving distracted driving. These countermeasures focus on reducing fatalities and injuries with the three E's of Engineering, Education, and Enforcement.

Engineers have developed countermeasures such as centerline and transverse rumble strips, edge lines, lighting, and wider and brighter striping—especially in crash-prone locations. Some in-vehicle technologies like Ford Sync have been developed to reduce visual distractions by enabling navigation, mobile communications, and audio device controls through voice commands (Shah, 2013).

Educational and awareness campaign materials presented on billboards and social media help to raise awareness about distracted driving in communities. For example, the month of April has been declared as a distracted driving awareness month by the NHTSA as a way of raising awareness regarding the dangers of distracted driving in order to eliminate preventable deaths and injuries on our roadways.

In addition to engineering and education strategies aiming to reduce the frequency and severity of distraction-related crashes, all the states have formulated laws to restrict risky behaviors such as texting or receiving calls while driving. Even with all these countermeasures put in practice, the number of crashes and fatalities associated with distracted driving is still high, thus necessitating further attention and investigation. This report will provide additional insights into the causes of distracted driving incidents and crashes, with the aim of reducing their frequency and severity.

## Research Objective

The primary objective of this study is to investigate distracted driving incidents and crashes in New Jersey using multiple data sources and analysis tools. To do so, longitudinal observational data were collected and statistically analyzed in ten important high-crash corridors in New Jersey. Statistical analysis was also performed on historical crash data to identify those factors contributing to distracted driving crashes. Following this, several safety countermeasures and strategies to reduce the frequency of distracted driving events were developed. It is expected that the outcomes of this study will assist transportation agencies and local officials in mitigating motor vehicle crashes involving distracted driving.

# CHAPTER 2: Literature Review

Distracted driving has been of interest to traffic safety scholars, practitioners, and policymakers since the early 2000s. However, researchers from diverse disciplines (e.g., psychology, medical science, and computer science) have put significant effort into investigating and addressing distracted driving. Depending on the focus of the study, the literature on distracted driving can be divided into several categories, including self-assessment surveys, crash analyses, driver simulations, observational studies, and methods for preventing distracted driving. Due to the versatility of the research on distracted driving, it is difficult to get a clear understanding of this problem and any solutions. The aim of this chapter is therefore to conduct a comprehensive literature review to better interpret distracted driving events and associated crashes. To be specific, this chapter will synthesize the findings of various data collection, data analysis, and crash prevention methods associated with distracted driving. This summary of the findings will help to suggest appropriate countermeasures of distracted driving to engineers, practitioners, and policymakers.

## Research Methodology

To provide a comprehensive view of the current state of practice regarding distracted driving, it was necessary to conduct a comprehensive search into what parts of the scientific literature focused on this topic. Database searches were therefore conducted on two well-known databases: the Transport Research International Documentation (TRID) service and Google Scholar. Keywords such as “distracted driving,” “distraction,” and “texting while driving” were searched for in the title, abstract, and keyword fields, and the scope was restricted to academic papers in English (including journal papers and conference proceedings). To capture any relevant “gray” literature (e.g., professional and agency reports), a Google search was also employed. Although a handful of studies on the topic were published before 2006, the primary focus was given to studies appearing during the last 15 years.

In the first step, 115 relevant papers and reports on distracted driving were gathered. In the second step, their titles, keywords, and abstracts were manually checked and refined in order to yield only papers specifically dealing with distracted driving. This filtration process brought down the number to 97. A final refinement based on the texts themselves was made, and all the papers with their full text available were sorted through this third step. After performing all of these steps, 74 focused articles were ultimately selected for this literature review. Figure 4 illustrates the stepwise track of this process.

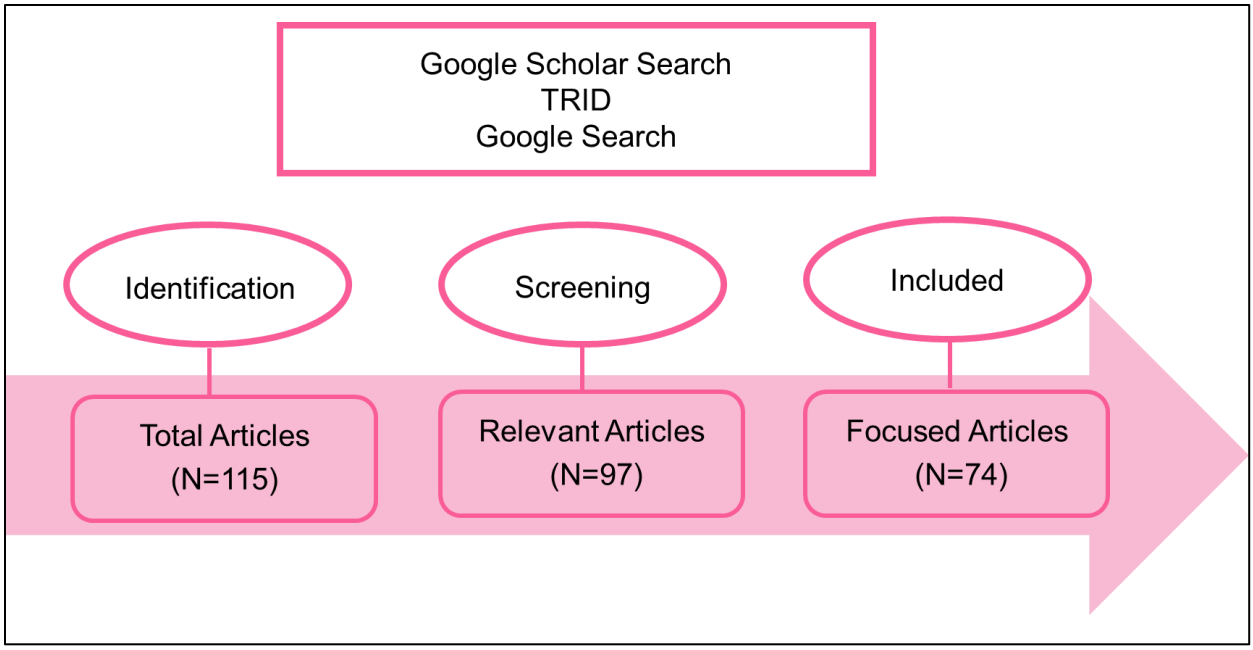



**Figure 4.** *Stepwise Literature Selection Process*

## Data Collection

Traditionally, data collected from surveys, crash reports, and observational studies have been used to investigate the dangers of distracted driving. However, more real-time approaches (e.g., from naturalistic driving studies, dashcam footage, or eye glance recorders) were also later introduced. This section will cover the various methods of data collection employed in distracted driving studies.

### *Survey-Based Studies*

Previous researchers have conducted numerous surveys to investigate driver concern and its involvement in distracted driving. Most of these survey studies ranked the level of driver involvement with secondary tasks and correlated this with their behaviors and their perception of distraction. For instance, Braitman et al. (Braitman & Braitman, 2017) conducted an ordinal/rank-based online survey of 266 young adult drivers. They demonstrated that those people engaging in distracting tasks like cellphone usage tend to rate these tasks as moderately risky. In contrast, those people undertaking less distracting tasks admitted that such visually demanding tasks do increase crash risk.

To investigate the driving behaviors and attitudes of Americans, the auto insurance company The Zebra operated a closed-end survey of 2000 people across the nation (Covington, 2021). The authors utilized the Google consumer survey platform and found that 24% of drivers text while driving. Although 36% of the respondents believe that texting is risky, half of them feel cognitive stress about texting back while driving. Through the Survey Sampling International (SSI) platform, the National Security Council (National Security Council, 2019) conducted a poll on 2,409 drivers across the nation, with 75% of respondents admitting that the main pressure on them is to receive calls while driving

comes from their family. Among these drivers, only 25% admitted that distraction creates dangers to other non-distracted drivers on the road.

Gliklich et al. (Gliklich, Guo & Bergmark, 2016) developed a nationwide web-based questionnaire to quantify the frequency of distraction events experienced by adult drivers in the United States. These authors demonstrated that there is a large correlation between cellphone-related distractions and younger driver age. Curry et al. (Curry, Hafetz, Kallan, Winston & Durbin, 2011) performed their research using the National Motor Vehicle Crash Causation Survey (NMVCC), finding that distracted driving is responsible for 20% of the critical errors made by teen drivers.

NHTSA has conducted the most comprehensive series of surveys on distracted driving, titled the National Survey on Distracted Driving (NSDDAB). Tison et al. (2011) summarized the findings of a nationwide telephone survey on distracted driving conducted in 2012 by the NHTSA, which utilized household landline phones and cellphones to interview 6,016 drivers across the nation. They found that talking with passengers, adjusting radios, eating, and drinking, and cellphone use are the leading sources of distraction. The younger drivers surveyed in this study were two to three times more prone to cellphone distractions than the other drivers. Overall, talking to passengers, tuning the radio, receiving calls, and eating were the leading sources of distraction while driving (NHTSA, 2014). In a later 2015 telephone survey study, it was found that 42% of respondents admitted that they received calls while driving. Interestingly, 8% of respondents used mobile apps while driving, with over half (56%) of those users believing that using apps while driving is not risky (Schroeder, Wilbur & Peña, 2018).

The findings from these nationwide surveys demonstrate that technology is a key factor behind the increased amount of distraction in recent years. Moreover, these studies also indicated that the involvement of younger drivers in distractions and subsequent crashes increases because of technology-oriented secondary tasks (i.e., texting, using apps, or receiving calls).

### *Observational Studies*

Many researchers have conducted observational studies on roadways to analyze distracted driving behavior. Most of these focused on investigating the prevalence of driver distractions involving handheld cellphones and receiving calls (Prat et al, 2015; Gras et al., 2012; Horberry, Anderson, Regan, Triggs & Brown, 2006; Narine, Walter & Charman, 2009). The NHTSA performs nationwide observational studies on intersections during the daytime, titled the National Occupant Protection Use Survey (NOPUS). Their 2019 study found that 3.2% of surveyed drivers use handheld phones while driving. Traffic at red light signals was recorded nationwide across 1,612 different intersections during daylight hours (7 a.m. to 7 p.m.) (Cambridge Mobile Telematics, 2020).

Bommer (2018) collected field data regarding distraction by cellphone use and found that it is more prevalent on local roads and in drive-alone cars compared to on highways and in the presence of occupants, respectively. Kidd & Chaudhary (2019)

conducted a roadside observational study across North Virginia and found 23% of the recorded drivers were distracted. Prat et al. (2015) performed a cross-sectional observational study in Spain to investigate driver distraction events and demonstrated that distracted driving behaviors vary due to temporal variations (e.g., during weekdays or at different times of the day). Sullman (2012) conducted a cross-sectional observational survey in six English cities and found that talking to passengers, smoking, and cellphone use were the leading types of distraction. Gras et al. (2012) found similar trends in their cross-sectional study across urban areas in Spain.

The findings from these studies give us a perception of the rate of distractions occurring on the observed roadways. However, these studies showed a wide range of variation in the recorded distraction rate (between 5% and 15%). Four reasons contributed to these variable distraction rates: different definitions of distraction, differences in data collection methods, regional differences, and lack of temporal diversity in the data. Regional variations in the observed studies may be due to different socioeconomic characteristics, roadway geometries, and the strictness of law enforcement (Adanu et al., 2017). More diverse observational studies on distracted driving should be conducted to address the regional and temporal biases of the data.

### *Cell Phone Tracking*

Cellphone apps can help in the collection of onboard data related to distracted driving. For example, Murray et al. (2019) utilized the TrueMotion app to encourage drivers to refrain from using their cellphones by offering incentives. The data it provided showed that handheld cellphone use increases during the holidays. Zendrive (2018), a smartphone-based platform that collects driver behavior data, is another platform that can be used to analyze cellphone usage while driving. By collecting driver behavior data, State Farm has produced a database of images that can be further used by researchers for training and detecting distracted driving behaviors (State Farm, 2016).

### *Crash Reports*

Researchers in this field have also made use of crash reports as a way of comprehending the patterns of driving behaviors that contribute to distracted driving crashes. One of the latest reports on distracted driving by NHTSA investigated distracted driving crash data from 2018 and demonstrated that cellphone use in drivers less than 40 years of age represents 69% of total drivers that are distracted. The Insurance Information Institute (2018) used the data provided by the NHTSA to publish an article on the facts regarding distracted driving occurring in 2017 and 2018. The authors demonstrated that cellphone use accounts for 13% of fatal crashes involving distracted driving.

Stimpson, Wilson & Muelleman (2013) investigated the Fatality Analysis Reporting System (FARS) records for 2005–2010 and found that pedestrians and bicyclists have a 1.6 times greater chance of getting hit by a distracted driver than a non-distracted one. Thus, the authors suggested implementing clear and lighted crosswalks and separate bicycle lanes, as well as measures to prevent distractions while driving. Marchese (2019) also used data from the FARS for crashes from 1991–2015 and found that young males

were represented as the group most involved in fatal crashes due to distracted driving. Once again, cellphone use was found to be the leading type of distraction in fatal crashes. These studies utilizing crash reports have been helpful in suggesting countermeasures to policymakers.

### *Eyeball Tracking and Gaze Detection*

Eyeball tracking has also been widely used by researchers to detect distraction by recording pupil movement and recognizing the direction of glance. Liang & Lee (2014) compared the performance of three detection algorithms—a dynamic Bayesian network (DBN), a layered Bayesian network, and a support vector machine (SVM)—by tracking the eye movements of participants with the faceLAB eye tracker.

Foss & Goodwin (2014) installed g-force cameras to monitor the behavior of teen drivers and found that they get engaged in more cellphone distractions when they are driving alone than when they drive with an occupant. Owens, McLaughlin & Sudweeks (2011) built an in-car system that integrated multiple dashcams, mobile systems, and eyeball tracking technologies to collect data on driver behavior. They found that texting degraded driver steering performance, accompanied by a more prolonged glance away from the road. Cabrall et al. (2016) utilized an eye tracker to measure gaze direction and found a deterioration in driving performance during visual distractions.

### *Simulations*

Yannis, Laiou, Papantoniou & Christoforou (2014) conducted a five-minute driving simulation with 34 young drivers by exposing them to different weather conditions and roadway features (e.g., rural roads, rainy weather, or jumping animals). The authors found that texting during driving results in slower speeds, an increase in driver reaction time, and an increase in the likelihood of getting involved in rear-end crashes.

Klauer, Dingus, Neale, Sudweeks & Ramsey (2006) utilized data from the 100-car naturalistic study to demonstrate that visual distractions decrease a driver's ability to keep in their lane during driving, resulting in safety-critical situations. Using the same data, Liang, Lee & Yekhshatyan (2012) found a positive correlation between crash risk and the degree of distraction estimated from driver eye-glance patterns.

Later on, Liang, Horrey & Hoffman (2015) also conducted a simulation of driver behavior on a straight highway and found that driving errors increase by 10% while texting. Gallahan et al. (2013) utilized the Microsoft Kinect motion-sensing device to detect distraction and to develop a distracted driving warning system in a driving simulator located at the Virginia Driving Safety Laboratory (VDSL). Using skeletal tracking, this model achieved a classification accuracy of 66%.

### *Use of Cameras*

To collect additional data on driver behavior, previous researchers have made use of cameras located on vehicle dashboards and on roadside poles. Victor et al. (2015) used onboard data from test cars and found that visually demanding tasks (like texting

while driving) are associated with high crash risks. Wang, Bao, Du, Ye & Sayer (2017) used images of drivers taken from various angles to assess their attention to driving under distracted and non-distracted conditions and calculated the entropy rate and a glance proportion matrix. They found that a higher scanning randomness level during visual-manual distractions shifts the attention of drivers from their primary task of driving.

Streiffer, Raghavendra, Benson & Srivatsa (2017) created a unified data analysis framework called DarNet to collect and analyze images of distracted test drivers, finding an increased classification accuracy compared to existing baseline models. De Castro et al. (2018) used cameras inside the car to track the eye gaze of drivers using the OpenFace model, and achieved a detection accuracy of 84% for distracted driving. Tran et al. (2020) used dual (front and side) cameras on a driving testbed to capture driver images, and they achieved a detection accuracy of 97%.

Johnson, Voas, Lacey, McKnight & Lange (2004) captured still images from the NJ Turnpike and found cellphone use was the major source of distraction. Elqattan, Moustafa & Shafey (2019) used innovative techniques to collect data from outside the car, either by mounting a camera to a police car or at locations on the roadside. The authors used the Xception model to detect distraction levels, and OpenALPR (with the help of GPS tracking) to detect and report the license plate numbers of distracted drivers.

## Data Analysis

### *Statistical and Discrete Choice Models*

In their work, researchers have used logistic regression models, ordered logit models, and probit models to analyze the results of distracted driving studies. These statistical and discrete choice models have helped researchers to discover the factors contributing to distracted driving events and crashes. For example, Qin et al. (2019) used Tukey's test, the chi-square test of independence, the Nemenyi posthoc test, and the Marascuilo procedure in their analysis and identifies vehicle devices (such as GPS, radio, music player) as major sources of distraction-related fatalities among young drivers.

Furthermore, D'Souza (2012) used multinomial logistic regression and found that driver age and fatigue levels are the most influential factors behind distractions. Jenkins, Codjoe, Alecsandru & Ishak (2017) likewise used the same method on the SHRP2 NDS data, showing that driver performance is affected by their involvement in secondary tasks. Neyens & Boyle (2007) studied four types of distractions in teen drivers using a multinomial logit model, while Claveria, Hernandez, Anderson & Jessup (2019) analyzed crash severity data on 515 distracted truck drivers using a random parameter logit model.

Many researchers have also employed mixed logit models to address the heterogeneity of the contributing factors. For instance, Hasan et al. (2021) investigated five years of New Jersey crash data involving cellphone use with a mixed logit model. They found that the urban setting and the age of the driver contribute significantly to the severity of crashes. Cao, Zhang, Song & Wang (2020) also investigated the SHRP2 naturalistic driving data using a mixed logit model and found that senior drivers are less

distracted by cellphones while driving.

### *Machine Learning and Deep Learning*

Machine learning algorithms are helpful in analyzing and detecting distractions, with Liang, Reyes & Lee (2007) doing so with an accuracy of 81% using an SVM. In comparison, Ahangari, Jeihani & Dehzangi (2019) used a Bayesian network to detect distracted driving, with an overall accuracy of 67.8%.

Quite a number of researchers have also been working on deep learning techniques to detect distracted driving. Liang & Lee (2014) did so with an 88% accuracy using a hybrid Bayesian model, while De Castro et al. (2018) achieved 89% using OpenFace—a feature extraction software. Eraqi, Moustafa, Abouelnaga & Saad (2019) proposed a genetically weighted ensemble of convolutional neural networks (CNNs), producing a reliable deep-learning-based system with a detection accuracy of 90%. Later, the authors proposed a thinned version of their ensemble with a classification accuracy of 84%. Abouelnaga, Eraqi & Moustafa (2017) also presented a robust vision-based system built with a genetically weighted ensemble of CNNs, achieving a 95.98% classification accuracy in driving posture estimation. Their simpler model AlexNet operates successfully in real-time with a classification accuracy of 94.29%.

In further examples, Elqattan et al. (2019) utilized pre-trained models of DarNet YOLO version 3 to detect the drivers inside the car and the Xception model to classify distraction, resulting in detection and classification accuracies of 89% and 95%, respectively. Leekha, Goswami, Shah, Yifang & Zimmermann (2019) have proposed another CNN-based system to perform real-time distracted driving detection. They achieved 98.48% and 95.64% test accuracies by training their model with State Farm and American University in Cairo (AUC) images.

Mase, Chapman, Figuredo, Torres & Chapman (2020) then presented a deep learning architecture that outclasses current CNN models (e.g., VGG-16, Resnet50, Inception V3-LSTM, and an ensemble of InceptionV3 with a GA-weighted algorithm) with an average accuracy of 92.7% when classifying distracted driving postures using static images. Later on, Huang, Wang, Wang, Zhang & Cao (2020) utilized a hybrid CNN framework (HCF) to detect distracted driving. The authors pretrained three different detection architectures (i.e., ResNet50, Inception V3, and Xception) using transfer learning, and merged them together to extract driver facial features. Their proposed HCF achieved a classification accuracy of 96% and an average processing time of 0.042 seconds.

In another approach, Alotaibi & Alotaibi (2020) proposed a model utilizing one block of ResNet, two layers of a hierarchical recurrent neural network (HRNN) built on top of the Inception architecture, and two dense layers with a softmax classifier. Their proposed method outperformed ResNet and HRNNs alone with an accuracy of over 92%. Baheti, Gajre & Talbar (2018) used the modified CNN architecture VGG-16, achieving a classification accuracy of 95.54%. Later, Baheti et al. (2020) introduced the Mobile VGG network and achieved 95.24% and 99.75% accuracy after training on the AUC and State

Farm datasets, respectively, and with less computational complexity and lower memory requirements. As a final example, Wang, Wu, Li & Zhang (2021) showed an enhancement of classification accuracy (to 96.97%) by detecting driving operation area (DOA) during the image preprocessing stage, and they did so using gradient-weighted class activation mapping (grad-CAM).

## Contributing Factors

One can categorize the factors contributing to distracted driving events as driver characteristics, roadway features, environmental features, and crash attributes. This section will summarize the previous research findings regarding these factors.

### *Driver Characteristics*

Driver characteristics are the most important factors behind distracted driving crashes. These include driver age, gender, and fatigue levels, each of which is described below.

**Driver Age.** This factor has been found to strongly impact the likelihood of getting distracted, especially by cellphones. Claveria et al. (2019), Neyens & Boyle (2007), and D'Souza (2012) all demonstrated that younger drivers are more prone to distraction than older drivers.

**Gender.** Distracted driving crashes are also significantly influenced by the gender of the driver. Behnood, Modiri & Roshandeh (2016) have found that young male drivers are more prone to distraction than female drivers. However, Qin et al. (2019) found that young female drivers are more prone to getting distracted by in-vehicle technology or devices.

**Driver Fatigue and Workload.** Driver fatigue is responsible for many crashes, especially for truck drivers. Claveria et al. (2019) and D'Souza (2012) both analyzed crash severity data and found that increased driver workload or fatigue is one of the most important parameters contributing to the severity of distracted driving crashes for truck drivers.

### *Roadway Features*

The geometric design of roadways and other road conditions can also play a vital role in the propensity of drivers to be distracted and thus to become involved in crashes.

**Surface Condition.** Poor surface conditions are harmful to traffic in all cases. However, their impact is even greater when drivers are impaired by other tasks. Neyens & Boyle (2007) found that poor surfaces are an important contributing factor to crashes caused by distracted driving.

**Urban Setting.** An urban driving setting contains more possible distractions than in rural environments because of factors such as congestion, speed limits, and intersections. D'Souza (2012), Neyens & Boyle (2007), and Chen & Lym (2021) all

investigated the influential factors behind the severity of distracted driving crashes and found that an urban setting was indeed one major influencing factor.

**Type of Roadway.** Crash types due to distraction can also depend on the type of roadway. Behnood et al. (2016) found that more severe crashes due to distracted driving occur on two-way roads. Different types of highways also have an impact on the severity of crashes due to distracted driving, with Chen & Lym (2021) finding that interstate highway has a higher severity level than other roadway types.

### *Environmental Features*

Because of vision impairment and susceptibility to glare, some people face problems in adverse weather conditions. However, clear weather can actually inspire drivers to get distracted, as Behnood et al. (2016) demonstrated.

### *Crash Attributes*

Crash type is associated with the severity of a distracted driving event. Neyens & Boyle (2007) demonstrated that teen drivers are more likely to get involved in fixed object and rear-end collisions, with cellphone distractions resulting in a greater likelihood of rear-end collisions.

## Safety Countermeasures

Safety countermeasures are the most important way to reduce traffic crashes. The implementation of such countermeasures against distracted driving can be described with the three Es: Engineering, Education, and Enforcement.

### *Engineering and Technology*

Various transportation agencies and companies provide external and internal engineering countermeasures to minimize crashes involving distracted driving. Previous studies suggested that wider and brighter striping and lighting (Behnood, 2016) or medians and shoulders (Chen & Lym, 2021) effectively reduce the severity of crashes due to distracted driving. Donmez, Boyle & Lee (2007) demonstrated that the real-time feedback provided to drivers could successfully bring their attention back to the road.

In another example, Ford SYNC helps reduce visual distractions by enabling navigation, mobile communications, and audio device controls through voice commands (Shah, 2013). By 2019, there were 29 cellphone applications that blocked features like texting, notifications, or receiving calls while driving. The most widely used app by Android users was Android Auto, while AT&T DriveMode was most widely used by iOS and Blackberry users (Oviedo-Trespalacios et al., 2019).

### *Education and Awareness*

The World Health Organization (WHO) has suggested that awareness campaigns could help reduce distracted driving (WHO, 2011). To train new drivers in adjusting to the

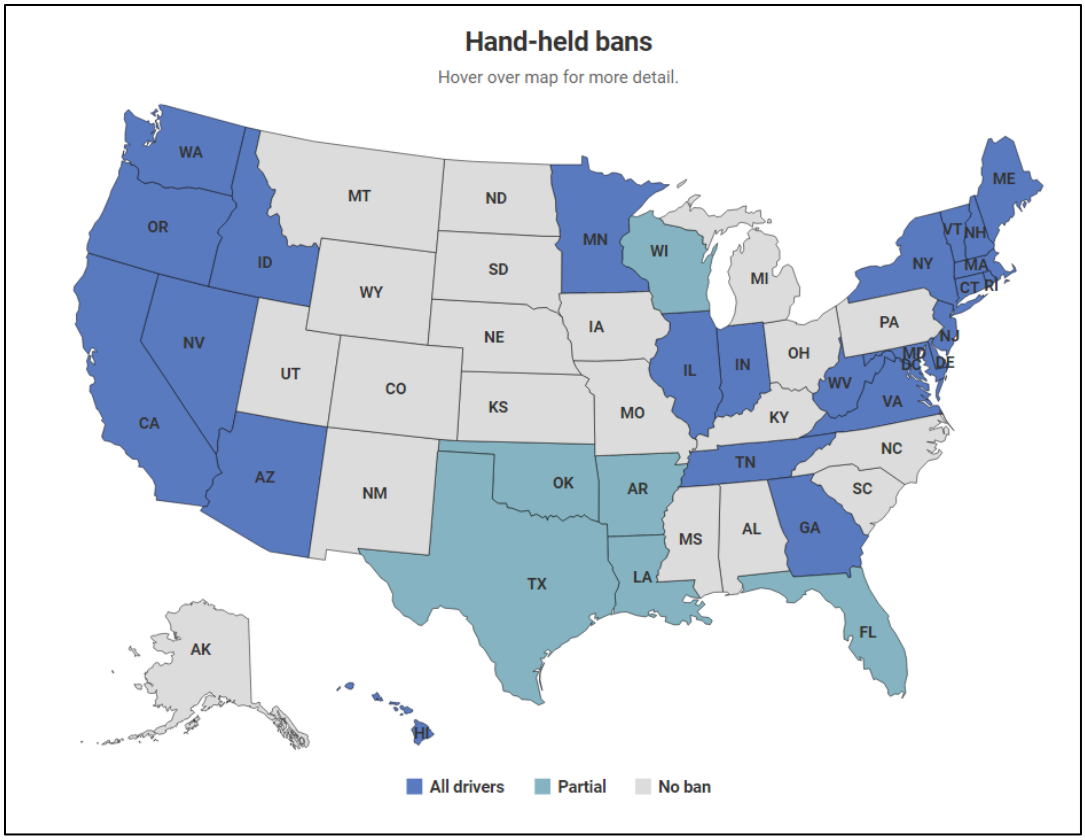
dangers of distracted driving, courses like “Impact Texas Teen Drivers” can prove useful (Texas Department of Public Safety, 2018). Other resources, like lesson plans from Toyota and safe driving training courses like Ford Driving Skills, can also be helpful in educating drivers about safe driving (TeenDrive365: In School, 2018; Ford Driving Skills for Life, 2018). Organizations like End Distracted Driving provides free educational materials, including safe driving agreements, quizzes, and surveys to help teen drivers deliver science-based presentations about the dangers of distracted driving to their parents (EndDD, 2021).

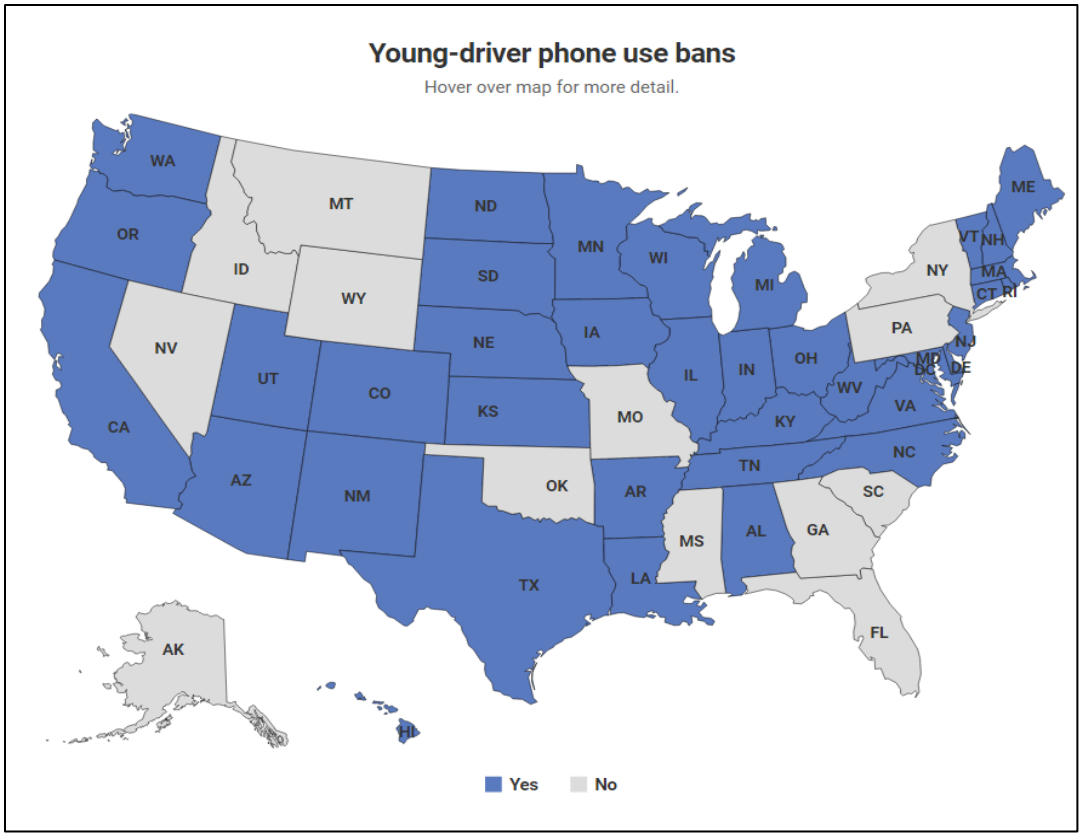
### *Enforcement*

The enforcement of current laws related to distracted driving helps to reinforce the educational and engineering countermeasures. Several laws have so far been passed in the states, including prohibitions on texting while driving and incremental fiscal punishments for the use of cellphones in general. A list of these laws from the various states is illustrated in Figure 5. To date, 21 states have bans on handheld cellphone use, 48 states have banned texting while driving, and 38 states have banned cellphone use by young drivers.

Aceable has developed a website that updates the fiscal punishments by the state for texting while driving (Aceable, 2014). The Governors Highway Safety Association (GHSA) has also listed the laws on distracted driving and cellphone use across the different states. The authors of one study by Stim et al. (2010) demonstrated that the general trend with regards to distracted driving laws is that they have become stricter over time. For instance, the fiscal punishment in Colorado for texting while driving was increased from $50 to $300 in 2017 (Kitch, 2018). In Connecticut, the penalty for first-time violators is $150, which increases to $300 for the second violation and $500 for any further offenses (Failla, 2019).

Some states also have rules whereby merit points are deducted for violations of traffic laws related to distracted driving. In Georgia, one point is deducted from the driver’s license for the first conviction, two points for the second conviction, three points for the third violation, and so on (Qi, Vennu, & Pokhrel, 2020). Figure 5 demonstrates that almost all states in the United States (except Montana) have laws restricting cellphone use while driving.


Hand-held bans
Hover over map for more detail.
WA
OR
CA
ID
NV
MT
WY
UT
AZ
CO
NM
ND
SD
NE
KS
OK
TX
MN
IA
MO
AR
LA
WI
IL
MI
IN
OH
KY
TN
MS
AL
GA
FL
SC
NC
VA
WV
PA
NY
VT
NH
ME
MA
CT
RI
NJ
MD
DC
DE
AK
HI
All drivers
Partial
No ban


Young-driver phone use bans
Hover over map for more detail.
WA
OR
CA
ID
NV
MT
WY
UT
AZ
CO
NM
ND
SD
NE
KS
OK
TX
MN
IA
MO
AR
LA
WI
IL
MI
IN
OH
KY
TN
MS
AL
GA
FL
SC
NC
VA
WV
PA
NY
VT
NH
ME
MA
CT
RI
NJ
MD
DC
DE
AK
HI
Yes
No

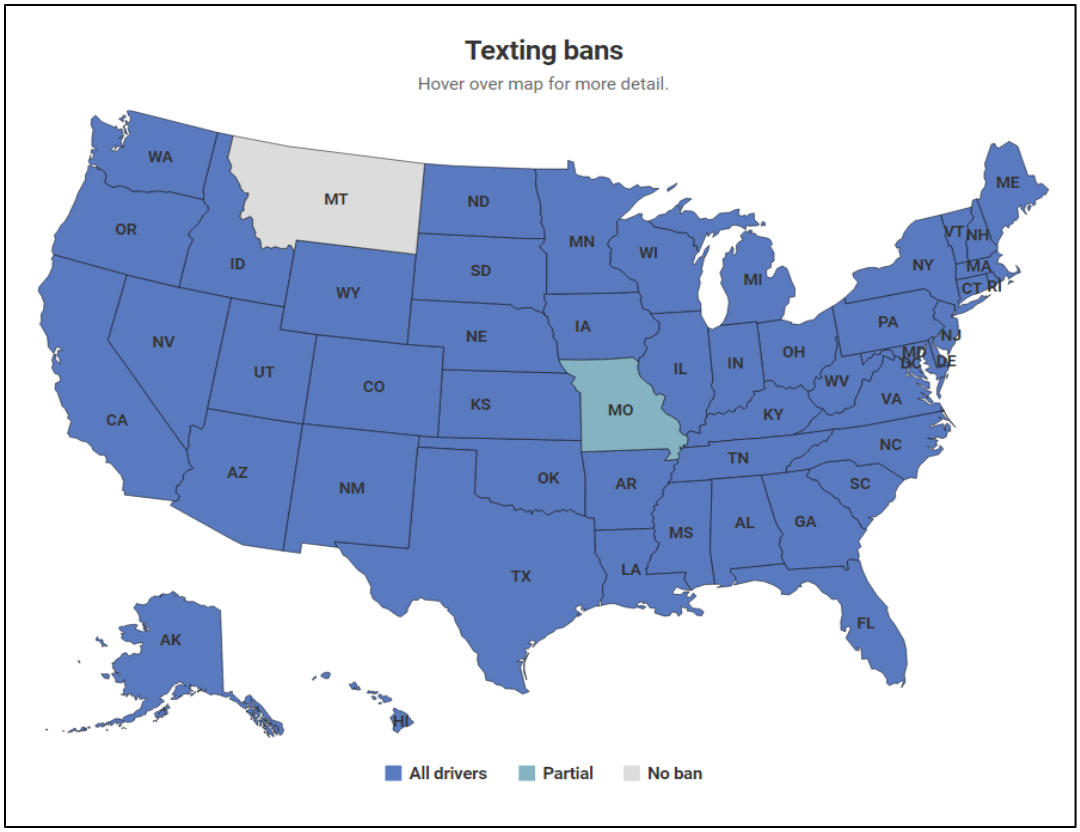



**Figure 5.** *Laws and Regulations on Cellphone Use While Driving in the Different States*

There have been various strategies employed across the nation in order to ensure that distracted driving laws are enforced. For instance, the Police Department of Austin, Texas, has taken the initiative of identifying texting while driving by monitoring drivers from the local mass transit lines running parallel to the roads. This action by officers is legal and is not a violation of rights (Skousen, Gulbrandsen & Patience, 2019).

Various government agencies have also participated in NHTSA's national high-visibility enforcement campaign "U Drive, U Text, U Pay." The goal of this campaign has been to increase distracted driving law enforcement efforts by catching distracted drivers. As of February 2020, 21 states, Washington D.C., Puerto Rico, Guam, and the US Virgin Islands have put into effect total handheld phone use bans, according to the Governors Highway Safety Association (NHTSA, 2019). This law prohibits all drivers from using handheld cellphones while driving and includes primary enforcement, thereby enabling an officer to cite a driver for using a handheld phone without there being any other traffic violation.

A study by NHTSA emphasized that the pre-deployment training of officers and the reallocation of resources is important to the enforcement of distracted driving laws. Distracted driving enforcement is different from traditional patrol strategies as it requires specialized skills to detect violators who conceal distracting devices. Law enforcement officers should be familiar with distracted driving laws in their jurisdictions. Police

departments should also provide training for officers to detect the observable cues of distracted driving and how to appropriately document violations (National Traffic Law Center, 2017).

A comprehensive way to mitigate distracted driving crashes or to change driver behavior can only be achieved by appropriate coordination of those entities that ensure the three Es: Engineering, Enforcement, and Education. Technology is an additional tool that can increase the effectiveness of the aforementioned three approaches. For example, updated Bluetooth technologies, drowsiness alert systems, call blocking, and eyeball tracking technologies have added new dimensions to the fight against distraction. Mobile carriers and insurance companies have also enticed drivers by providing discounts or monetary benefits in return for avoiding distractions while driving. Furthermore, the federal government and the NHTSA (2012) have prepared a blueprint for ending distracted driving through awareness and enforcement approaches, with the expectation of a reduced number of crashes due to distractions. Finally, a continuous or periodic evaluation of the three Es could help to quantify the safety benefits of particular countermeasures.

## Conclusion

Distracted driving is a concern in many fields, including academia, engineering, vehicle manufacturing, traffic safety and the healthcare industry. Researchers have found that heavy workloads, environmental factors, along with other factors such as driver and roadway characteristics, contribute to the severity of distraction-related crashes. The most common finding in the literature is the increased involvement of distractions in young drivers, especially with electronic devices. Variations in research approaches are reflected in the different findings, with survey-based methods mainly emphasizing the likelihood of getting distracted, while the more comprehensive methods like statistical modeling investigate the impacts of various contributing factors on crash severity resulting from distracted driving. Recent research on distracted driving has mostly focused on collecting driver behavior data during distractions by using newer types of equipment and, later, detecting them using artificial intelligence.

Most of the states have strict rules on distracted driving, including increasingly large fines for cellphone use. The lack of effectiveness of these legal measures and their enforcement brings into question their safety benefits. Therefore, education could prove a more effective role here. Every road user should be made aware of the dangers of distracted driving through the use of news media, TV or radio channels, social media, posters, and awareness campaigns like “U Text, U Drive, U Pay.” The USDOT has arranged for several awareness months in an attempt to reduce distracted driving, and has funded effective awareness campaigns through the state police in various states.

Future survey studies on distracted driving should cover more sociodemographic variation in order to investigate driving exposure and the factors behind distraction more comprehensively. The inclusion of surrogate safety measures and near misses under the various road, or environmental conditions could help investigate distraction-related crashes. These studies would help to determine strategic approaches to developing

campaigns or educational programs, specifically for those target groups prone to distracted driving. Also, these results can help us to evaluate the effectiveness of short- and long-term enforcement measures (like bans or fiscal punishments) enacted in various states.

When evaluating countermeasures, observational data can also play a significant role, combined with the appropriate use of technology. New techniques like dashcam recording, eyeball tracking, and glance recognition could be used to collect driver behavior data. At the same time, deep learning and other video processing algorithms could prove helpful in the identification of distracted driving. Future studies could also focus on the effectiveness of the countermeasure initiatives promoted by insurance companies and cellphone carriers, whereby cellphone usage is restricted while driving.

# CHAPTER 3: Analysis of Distracted Driving Crashes in New Jersey Using a Mixed Logit Model

Driving is a task that demands uninterrupted concentration. However, many drivers get engaged in other tasks that distract their attention away from driving. Technology-based distractions have been rising among drivers due to the rapid technological evolution of recent years. Hence, traffic safety engineers and researchers have emphasized the importance of understanding the impacts of technology on driver distraction, especially those of cellphones, which have received a lot of attention across the country due to the increased risks involved (Farmer & Braitman, 2010; Sullman & Baas, 2004; Klauer, Dingus, Neale, Sudweeks, & Ramsey, 2006; Oviedo-Trespalacios, Haque, King, & Washington, 2017; Stavrinos, et al., 2013; Cambridge Mobile Telematics, 2020). According to a report published by NHTSA, 385 people died in the United States in 2018 due to cellphone distractions (NHTSA, 2018). In addition, a total of 2,060 fatal crashes (412 per annum) occurred across the nation over a recent five-year period (2014–2018) due to cellphone distractions (NHTSA, 2018).

Comprehending which factors influence crash injury severity in crashes involving cellphone distractions is essential when suggesting countermeasures to reduce fatalities. Cellphone distractions have been at the center of attention for many scholars, who have used police crash data or naturalistic driving data in their analyses (Pickrell & Jianqiang, 2009; Dingus, et al., 2016; Beanland, Fitzharris, Young & Lenné, 2013; Caird, Willness, Steel & Scialfa, 2008; Dingus, et al., 2006; Jashami, Abadi & Hurwitz, 2017; McEvoy & Stevenson, 2007; Regan, Lee & Young, 2008; Cambridge Mobile Telematics, 2020; Gordon, 2009). Although crash reports miss a significant amount of information, naturalistic driving data is costly and time-consuming (Gordon, 2009; Asbridge, Brubacher, & Chan, 2013). The use of simple statistical methods in most of these studies provided insights into the factors associated with cellphone distractions in crash severity analysis (Asbridge, Brubacher, & Chan, 2013; Dingus, et al., 2016; Hanowski, Perez, & Dingus, 2005; Olson, Hanowski, Hickman, & Bocanegr, 2009). However, very few studies have emphasized the unobserved heterogeneity in these factors (Behnood & Modiri-Gharehveran, 2016; Gordon,2009).

This chapter overcomes the limitations of previous studies by implementing a mixed logit model to investigate this unobserved heterogeneity of cellphone-related crashes in the New Jersey crash data. Furthermore, the pseudo-elasticity of the important factors was determined in order to investigate the heterogeneity of the most crucial factors. To the best of the author's knowledge, only a few previous studies have focused on finding the factors contributing to cellphone distractions using a mixed logit model. Findings from this report will provide engineers, researchers, and policymakers with sufficient data to suggest appropriate countermeasures for

combating the devastating impacts of cellphone distractions in the state of New Jersey.

## Methods and Data

### *Data Collection*

Crashes due to cellphone distractions in New Jersey from 2015 to 2019 were considered for this research. The crash data was obtained from the NJDHTS Crash Analysis Tool and was filtered based on the involvement of cellphone use as a contributing factor. The source of the data identifies whether cellphone use was implicated as one of the major contributing factors to the crash. However, cellphone use is not the only factor described as contributing to crashes. For instance, a driver violating the speed limit could be involved in a crash without using a cellphone. In the second case, a driver using cellphone could get involved in a crash while abiding by the speed limit. Thirdly, a driver could be both speeding and texting while getting involved in a crash.

Based on previous studies and engineering judgments, 25 independent variables were selected for further analysis. The severity of the crash had five different categories: fatal, suspected severe injury, suspected minor injury, possible injury, and no injury. A small percentage of fatal, suspected minor and suspected major injuries were found among the 3,040 crash incidents. For the convenience of the study, fatal (7, 0.23%), suspected minor injury (19, 0.63%), and suspected major injury (241, 7.93%) were merged together and named “injury.” After merging, the final dataset contained 1,921 (63.19%) cases of no injury, 853 (28.06%) cases of possible injury, and 266 (8.75%) cases of injury.

The four continuous variables of AADT (Annual Average Daily Traffic), pavement width, shoulder width, and speed limit were used in the crash dataset. Based on previous studies on distracted driving, the categorical variables were divided into six categories: temporal features (i.e., season, day of the week, time of day), roadway features (i.e., number of lanes, highway type, functional class of road, temporary traffic control device, traffic control present), environmental conditions (i.e., weather, light, and surface conditions), driver characteristics (i.e., driver age, drunk or drugged driver involved, unsafe driving involved), vehicle characteristics (i.e., the total number of vehicles involved, vehicle type) and other crash attributes (i.e., curve related, crash type, severity, intersection related, pedestrian or bicyclist involved).

For the convenience of classification, the variable inputs were further categorized into different ranges, based on previous crash severity analysis studies (Claveria, Hernandez, Anderson, & Jessup, 2019; D’Souza, 2012; Neyens & Boyle, 2007; Behnood & Modiri-Gharehveran, 2016). For instance, the crash dates were grouped into four different seasons: spring, summer, fall, and winter. The surface

condition was also divided based on the status of being dry or wet. Similarly, the vehicles involved were classified depending on their weight.

The frequency of reported crashes provides some further crucial facts. For instance, 62.14% of all crashes due to cellphone distractions were rear-end collisions, which happens when a sudden speed reduction occurs as the result of the loss of attention due to cellphone use. Furthermore, a late reaction on the part of the following vehicle can also be the cause of a rear-end crash, which itself can again be the result of inattention due to the use of cellphones. Wet weather accounts only for 12.6% of crashes involving cellphone use. Young drivers accounted for 18.95% of total crashes, which is consistent with the concerns of traffic safety engineers regarding the behavior of teen drivers.

### *Model Selection*

Researchers have used various statistical methods in their crash severity analyses, which have been discussed comprehensively in some previous studies (Savolainen, Mannering, Lord, & Quddus, 2011; Mannering & Bhat, 2014). The multinomial logit (MNL) model is widely used among these statistical methods, and it can be used to establish a relationship between a categorical dependent variable and a set of continuous and categorical independent predictor variables (D'Souza, 2012). One of the shortcomings of the MNL model is its restrictive assumptions regarding the unobserved error terms: it considers them to be independently and identically distributed, which is not applicable for correlated unobserved errors (Manski & McFadden, 1981; Kim, Ulfarsson, Shankar, & Mannering, 2010). Sometimes, the unobserved terms for a single crash incident (e.g., weather, age of the driver, and speed limit violation) can be correlated, which could shape the outcome of the crash severity (Ulfarsson & Mannering, 2004; Islam & Mannering, 2006; Winston, Maheshri, & Mannering, 2006).

In order to address the unobserved heterogeneity of the model, a new flexible method—termed a mixed logit model (MLM)—has been developed that allows the error terms to vary across the random parameters (McFadden & Train, 2000; Train, 2009; Milton, Shankar, & Mannering, 2008). This model has already been implemented extensively in the domain of transportation research (Kim, Ulfarsson, & S. Kim, 2013; Roque, Jalayer, & Hasan, 2021; Haleem & Gan, 2013; Hao, Kamga, & Wan, 2016; Wu, et al., 2014), and it allows us to assume the effect of some parameters randomly while keeping the rest fixed. This flexibility is effective for studying crash severity, which contains significant heterogeneity (Behnood & Modiri-Gharehveran, 2016). This study used such an MLM for the interpretation of the impact of contributing variables to the injury severity in cellphone distracted driver crashes.

**Mixed Logit Model.** A mixed logit or random parameters logit model (MLM) is used for the analysis of discrete data (Washington, Karlaftis, & Mannering, 2020; McFadden & Train, 2000). It is different from MNL models as it allows for the

observation of heterogeneity effects (Train, 2009). Also, this model is not restricted only to normal distributions. If the injury severities are classified into $K$ levels (here $K$ = 3), the driver injury severity level $k$ ($k \in K$) for the $n$-th driver, $Y_{kn}$, is given by:

$$Y_{kn} = \boldsymbol{\beta}_k \boldsymbol{X}_{kn} + \varepsilon_{kn} \quad (1)$$

where $\boldsymbol{\beta}_k$ is a vector of parameters to be estimated for each driver injury severity level $k$, which may vary across observations; $\boldsymbol{X}_{kn}$ is a vector of explanatory variables (surface condition, driver age, etc.); and $\varepsilon_{kn}$ is the disturbance term, which is assumed to be described by a generalized extreme value distribution (Manski & McFadden, 1981). Consequently, the standard MNL model (neglecting the error components) can be expressed as:

$$P_n(k) = \frac{e^{\beta k X kn}}{\sum_{\forall k} e^{\beta k X kn}} \quad (2)$$

where $P_n(k)$ is the probability of the $k$-th severity level occurring for the $n$-th driver. The random parameters that capture unobserved heterogeneity on driver injury severity outcomes are given by ($\boldsymbol{\beta}_k|\boldsymbol{\varphi}$), where $\boldsymbol{\varphi}$ is a vector of the probability density function (PDF). According to previous studies using the MLM (McFadden & Train, 2000; Train, 2003), the resulting weighted outcome of probabilities, ($k|\boldsymbol{\varphi}$), is given by:

$$P_n(k|\boldsymbol{\varphi}) = \int_x \frac{e^{\beta k X kn}}{\sum_{\forall k} e^{\beta k X kn}} . (\boldsymbol{\beta}_k|\boldsymbol{\varphi})\, d\boldsymbol{\beta} \quad (3)$$

where $f(\beta_k|\phi)$ allows the parameters to vary across the distribution $\beta$ defined by the researcher, with $\beta$ being normally distributed most of the time. The injury severity outcome probabilities for observation-specific variations of explanatory variables ($X_{kn}$) can be described by $\beta$.

The final step in the modeling process is to determine the significance of the log-likelihood values, and this is done through a log-likelihood ratio test as follows (Washington, Karlaftis, & Mannering, 2020):

$$\chi^2 = -2[LL(\beta_{Fixed}) - LL(\beta_{Random})] \quad (4)$$

where $LL(\beta_{Fixed})$ represents the log-likelihood at convergence with fixed parameters (where $\beta$ does not vary its sign across observations), $LL(\beta_{Random})$ is the log-likelihood at convergence with random parameters (where the sign of the estimated random parameters can vary across observations), and $\chi^2$ is a chi-square statistic with degrees of freedom equal to the number of estimated random parameters in LL ($\beta_{Random}$).

**Elasticity**. Finding the value of coefficients from the MLM cannot adequately describe the changes of outcome probabilities due to changes in explanatory variables. Since the marginal effect of a variable depends on all the parameter estimates in the model, it is not possible to comprehend the net effect of one variable

from the value of this single parameter (Khorashadi, Niemerier, Shankar, & Mannering, 2005). Elasticity stands for the change in injury outcome probabilities due to the change of one variable input. The elasticities of the parameter estimates for each of the most severely injured drivers are expressed by Equation 5 below (Washington, Karlaftis, & Mannering, 2020):

$$E_{Xmn}^{Pkn} = [1-P_{kn}]\, \boldsymbol{\beta}_k X_{mk} \tag{5}$$

where $P_{kn}$ accounts for the probability of the outcome and $X_{mk}$ is the value for variable *m* at the injury severity level of *k*. Although elasticities are not applicable to dummy variables, the pseudo-elasticity of the *m*-th variable from vector $X_n$ or $X_{mn}$ for person *i* experiencing *j* outcome can be expressed by Equation 6 below (Ulfarsson and Mannering, 2004):

$$E_{Xmn}^{Pkn} = \left(e^{\beta mk} \frac{\sum_{k=1}^{K} e^{\beta' kXn}}{\sum_{k=1}^{K} e^{\Delta(\beta' kXn)}} - 1\right) x\ 100 \tag{6}$$

where *K* is the total number of possible outcomes; $\Delta(\beta'_k X_n)$ is the function's value for determining the outcome $Y_{kn}$ when $X_{mn}$ is changed from 0 to 1, with $\beta'_k X_n$ being the value when $X_{mn} = 0$; $X_n$ is a vector of *m* explanatory variables shared by all outcomes; $\beta_k$ is a vector of estimated coefficients on the *m* variables for outcome *k;* and $\beta_{mk}$ is the coefficient on $X_{mn}$ in outcome *k.*

The elasticity of a variable $X_{mn}$ is the effect on the change of outcome probability $P_{kn}$ due to a 1% change in $X_{mn}$. The pseudo-elasticity of a dummy variable with respect to injury severity represents the change in the probability of that injury severity category as a percentage when the variable is changed from zero to one. Hence, a pseudo-elasticity of 25% for a variable in a particular injury category means that, when the values of the variable in the subset of observations where $X_{mn} = 0$ are changed from 0 to 1, the average increase in the probability of an injury outcome for these observations is 25% (Savolainen, Mannering, Lord, & Quddus, 2011).

## Results

The MLM is developed for three severity classes (i.e., no injury, possible injury, and injury) by using the total dataset of 3,040 observations. This model is tested for temporal features (e.g., time of day), driver attributes (e.g., driver age, drugged driver, speed limit violations), roadway features (e.g., AADT, number of lanes, the speed limit, functional classification, divided or undivided road), vehicular information (e.g., total vehicles involved), environmental (e.g., surface conditions), and crash attributes (e.g., type of crash, pedestrian/bicycle involvement). A summary statistic of the variables retained in the final model is shown in Table 1.

**Table 1** Descriptive Summary of Independent Variables

| Variable Name | Definition | Mean | Std. Dev. | Minimum | Maximum |
|---|---|---|---|---|---|
| ***Temporal Characteristics*** | | | | | |
| Time of Day | Night =1, else = 0 | 0.323 | 0.467 | 0 | 1 |
| ***Driver Attributes*** | | | | | |
| Driver Age | Young =1, else = 0 | 0.189 | 0.392 | 0 | 1 |
| Drugged Driver | Drugged driver involved=1, else=0 | 0.078 | 0.269 | 0 | 1 |
| Speed Limit Violation | Speed limit violation =1, else=0 | 0.031 | 0.174 | 0 | 1 |
| ***Roadway Features*** | | | | | |
| Functional Classification | Urban=1, else=0 | 0.945 | 0.227 | 0 | 1 |
| Divided or Undivided | Undivided=1, else=0 | 0.488 | 0.499 | 0 | 1 |
| Speed Limit | Continuous | 45.391 | 10.597 | 0 | 65 |
| Number of Lanes | Continuous | 2.599 | 0.846 | 1 | 6 |
| AADT | Continuous | 39832 | 36935 | 0 | 247662 |
| ***Vehicular Attributes*** | | | | | |
| Total Vehicles Involved | Continuous | 2.009 | 0.633 | 1 | 8 |
| ***Environmental Condition*** | | | | | |
| Surface Condition | Wet=1, else=0 | 0.125 | 0.332 | 0 | 1 |
| ***Crash Attributes*** | | | | | |
| Type of Crash | Angular=1, else=0 | 0.079 | 0.271 | 0 | 1 |
| | Rear-end=1, else=0 | 0.621 | 0.485 | 0 | 1 |
| Pedestrian/ Bicycle Involved | Pedestrian/Bicycle involved=1, else=0 | 0.004 | 0.068 | 0 | 1 |

## *Mixed Logit Model*

A variety of variables were found to be statistically significant to the various injury severity classes of cellphone crashes. Most of the variables were significant for one of the significance tests values (i.e., the p-value or the t-stat), while some parameters showed statistical significance for both of the tests. The parameter estimation results are mentioned in Table 2. The results of this MLM are compared with the MNL model, with the goodness-of-fit measures suggesting it is a better fit—given the Akaike Information Criteria (AIC) value of 5,087.2 for the former is smaller than for the latter (5,095.4).

**Table 2** Results of the Mixed Logit Model

| Attributes | Mixed Logit Model | | | Multinomial Logit Model | | |
|---|---|---|---|---|---|---|
| | Coefficient | t-stat | Standard error | Coefficient | t-stat | Standard error |
| ***No injury*** | | | | | | |
| Speed Limit | -0.0364*** | -5.58 | 0.0065 | -0.0172*** | -3.71 | 0.0046 |
| Number of Lanes | 0.0804 | 1.22 | 0.0804 | 0.8867* | 1.90 | 0.4666 |
| Angular Crash | -0.3604*** | -8.94 | -1.3604 | -0.2664*** | -8.58 | 0.1476 |
| Young Driver Involved | 0.2139 | 1.45 | 0.2139 | 0.0475 | 0.48 | 0.0989 |
| Constant | 2.9988*** | 7.29 | 2.9988 | 2.0243*** | 7.10 | 0.2851 |
| ***Possible injury*** | | | | | | |
| #Rear-end crash | -4.0909 | -1.59 | 2.5729 | 0.6199*** | 6.59 | 0.0941 |
| Pedestrian or Bicyclist | 1.2731** | 2.30 | 0.5535 | 1.1776** | 2.15 | 0.5477 |
| Wet Surface | 0.3077 | 1.58 | 0.1947 | 0.1627 | 1.35 | 0.1205 |
| Undivided Highway | 0.4213** | 2.55 | 0.1652 | 0.2769*** | 2.73 | 0.1014 |
| ***Injury*** | | | | | | |
| #AADT/1000 | -0.0129** | -2.30 | 0.0056 | -0.0023 | -1.13 | 0.0020 |
| Night hours | 0.3068** | 2.04 | 0.1504 | 0.1870 | 1.31 | 0.1427 |
| Alcoholic or Drugged | 0.7632*** | 3.56 | 0.2144 | 0.7543*** | 3.73 | 0.2022 |
| Urban | -0.5209** | -2.21 | 0.2357 | -0.0711*** | -3.15 | 0.0226 |
| Speed Limit Violation | 1.3604*** | 4.82 | 0.2822 | 1.2439*** | 4.90 | 0.2539 |
| Total number of vehicles | 0.2204** | 2.06 | 0.1070 | 0.1229 | 1.22 | 0.1008 |
| Constant | -0.4027 | -1.22 | 0.3301 | -0.3393 | -1.13 | 0.3003 |
| ***Distance of random parameters' standard deviation (normally distributed)*** | | | | | | |
| Rear-end crash | 11.1600 | 2.18 | | - | - | |
| AADT/1000 | 0.0123 | 2.31 | | - | - | |
| **Model parameters** | | | | | | |
| AIC | 5087.2 | | | 5095.4 | | |
| Log-Likelihood | -2525.5937 | | | -2531.7242 | | |
| Number of Observations | 3,040 | | | 3,040 | | |
| ****, **, * represents the statistical significance at 99%, 95%, 90% confidence limits respectively.* | | | | | | |
| *# is the indication of Random parameter* | | | | | | |

## *Elasticity*

Elasticity accounts for the change in severity for one significant explanatory parameter on a severity class while keeping the other parameters unchanged. The magnitude of elasticity accounts for the change in probability of a particular severity class, while the sign provides the direction of the change. Table 3 shows the average pseudo-elasticity estimates of injury, possible injury, and no injury for cellphone crashes.

**Table 3** Elasticity Results of the Mixed Logit Model

| Attributes | Elasticity Effect |
|---|---|
| ***No injury*** | |
| Speed Limit | -0.3670 |
| Lane Count | 0.0448 |
| Angle | -0.0612 |
| Young | 0.0078 |
| ***Possible injury*** | |
| Rear-end | 0.0817 |
| Ped-bicycle | 0.0033 |
| Wet | 0.0138 |
| Undivided | 0.0803 |
| ***Injury*** | |
| AADT/1000 | -0.2192 |
| Night | 0.0836 |
| Drugged | 0.0465 |
| Urban | -0.4308 |
| Speed | 0.0281 |
| Total Vehicle involved | 0.3842 |

## Discussion

### *Continuous Variables*

The variable “speed limit” was found to be statistically significant for the “no injury” severity level, with an estimated value of −0.368. The negative value of the coefficient is interpretable since the “no injury” event is more likely to happen at a lower speed limit. The “number of lanes” variable has shown a positive coefficient value of 0.08035 for the same injury category. This suggests that there is a higher likelihood of such a “no injury” event when a higher number of lanes are present, perhaps because an increased number of lanes provides more separate ways for traffic traveling in different directions to maneuver.

The parameter AADT was found to be significant (at $p<0.05$) for the “injury” category, with a coefficient value of −0.00129. This implies that “injury” is less likely to occur for higher values of AADT (i.e., more traffic volume or congestion leads to less propensity for serious injury). The congested traffic condition of urban roadways might result in slower traffic movement. However, this parameter varies randomly, and the standard deviation for AADT was found to be significant.

The total number of vehicles involved in the collision was found to be positively associated with “injury,” with a higher number of vehicles involved increasing the complexity and severity of the crash. Therefore, an increase in the number of vehicles would indeed increase the probability of an “injury” class event. This finding is consistent with (Liu & Subramanian, 2009), who found that the speed limit had a significant impact on crash severity.

As the value for “speed limit” increases, the probability of a “no injury” crash decreases to a great extent (36.7%). Therefore, a decrease in speed for the vehicles involved reduces the probability of someone getting injured. On the other hand, the number of lanes slightly increases the probability of “no injury” (4.48%). When it comes to AADT, this variable could decrease the probability of “injury” by 21.92%, since more AADT is experienced in urban settings with lower speed limits. An increase in the number of vehicles would increase the probability of “injury” by 38.42%, which proves that vehicle involvement increases the complexity of the crash and thus increases the injury severity.

### *Temporal Variables*

Nighttime was found to have a significant impact on the severity class “injury,” with a coefficient estimate of 0.30678. This finding is consistent with previous studies. For example, injury severity level is more likely to increase for crashes during the early morning (midnight to 6:30 a.m.) (Hao, Kamga, & Wan, 2016). Injury severity can increase at nighttime due to visibility issues. These results are consistent with previous research (Kim, Ulfarsson, & S. Kim, 2013; Wu, et al., 2014; Xie, Zhang, & Liang, 2009), where it was found that daylight decreases the possibility of fatal crashes. The probability of a crash being of severity class “injury” increases by 8.36% due to cellphone crashes during the night. Because of the complex situation of lighting coming from different directions at nighttime, cellphone crashes are more prone to causing severe injuries during those times.

### *Driver Behavior*

The young driver parameter demonstrated a coefficient value of 0.21389 for the “no injury” severity level. Although young drivers frequently get involved in cellphone distractions (Ferguson, 2003), the proportion of young drivers in fatal crashes caused by distractions (13.65%) was less than for older drivers (18.17%). In this case, older drivers are less dynamic than drivers in other age categories (Adebisi, Ma, Masaki, & Sobanjo, 2019). The quick reaction times and dynamic behaviors of young drivers might play a role in this positive relationship with the “no injury” category.

Drivers under the influence of substances (e.g., alcohol or drugs) showed a coefficient value of 0.76318 for “injury,” which implies that the probability of serious injury crashes increases under these circumstances. This is consistent with the findings from the literature (Wang, Li, Wang, & Liu, 2020). In another study (Stimpson, Wilson, & Muelleman, 2013), it was also demonstrated that there was an increase in injury severity due to the involvement of alcohol or drugs.

Speed limit violations were found to be significant for the severity level of “injury.” A positive value of 1.36037 for the coefficient proves that speed limit violations are one of the primary reasons for cellphone-related severe crashes.

Higher speeds were also found to be associated with more severe injuries in previous studies (Jurewicz, Sobhani, Woolley, Dutschke, & Corben, 2016).

Although there is a major concern for young drivers involved in cellphone-related crashes, their likelihood of not getting injured only slightly increases by 0.78% compared to other age groups. This value demonstrates that care should still be taken for young drivers as well. The consumption of alcohol or drugs by drivers increases the crash severity level "injury" by 4.65%, which means cellphone crashes coupled with alcohol or drugs can be increasingly severe. Speed limit violations also increase the probability of "injury" slightly by 2.81%. Violating the speed limit is always harmful, especially when the road surface and lighting conditions are bad.

### *Roadway Features*

A negative value (−0.52097) of the coefficient was obtained for the severity class "injury" for crashes in urban settings. Previously, one study suggested that there is a higher probability of a no injury outcome for crashes in an urban setting (Al-Bdairi & Hernandez, 2020; Casado-Sanz, Guirao, & Attard, 2020). Furthermore, urban settings have higher traffic volumes, lower speed limits, and less sight distance (Kusano & Gabler, 2013), which discourage drivers from getting distracted and thus getting involved in a crash. The undivided highway scenario has a positive value of the coefficient (0.42129) for "possible injury." The probability of possible injury increases when a highway is not divided. This result is also consistent with the findings from the literature (Zhou & Chin, 2019), and the explanation here is that collisions from both sides of the road can increase the probability of injury on undivided roadways.

An urban road setting has a significant impact on injury severity, with the probability of "injury" found to decrease by 43.08%. The traffic count and congestion of urban settings make drivers more alert while driving, which mitigates crash severity. On the other hand, traveling on an undivided roadway increases the probability of "possible injury" by 8.03%. The lack of division between lanes is again likely the cause of this increase in possible injury.

### *Environmental Conditions*

A wet surface shows a positive value of the coefficient for the "possible injury" category. On a wet surface, a vehicle can lose control while speeding. Moreover, a slippery road makes it hard for a vehicle to decelerate to a safe speed, which in turn could expose drivers to a possible injury. Previous studies also demonstrated that driving on wet surfaces leads drivers to be more prone to injuries (Liu & Subramanian, 2009; Roy & Dissanayake, 2011). Improvements to the surface resistance and drainage facilities can help reduce crashes due to wet conditions. In these results, wet road conditions increased the probability of "possible injury" slightly by 1.38%,

again because wet surfaces make it harder to retain or regain control when driving in tough situations.

### *Crash Attributes*

Rear-end crashes demonstrated a negative value (−4.09097) for the coefficient for the severity class “possible injury.” Although rear-end crashes hold the major share of all crashes, these findings warrant further investigation. This random parameter showed a standard deviation of 2.18, which indicates that the impact of rear-end crashes can vary across the “possible injury” severity class. One report in the literature demonstrated that rear-end crashes are more likely to happen because of distractions (Yan, Radwan, & Mannila, 2009).

Angular crashes demonstrate a significant impact on the crash severity class “no injury,” with a confidence limit of 99% and a coefficient value of −1.36044. The negative value of the coefficient demonstrates that the probability of not getting injured is less when an angular crash occurs. These types of crashes happen due to improper decision taking and lane-keeping, which can result in a collision with a vehicle from the opposite side of the road. This can of course intensifies injury severity (Neyens & Boyle, 2007).

When it comes to the involvement of pedestrians or bicyclists, this factor significantly impacts the” possible injury” type of crash severity. The positive value of coefficient (1.27308) demonstrates that a possible injury is likely to happen when a pedestrian or bicyclist is involved in the crash. The presence of these two groups might discourage the use of a cellphone while driving since a pedestrian or a bicyclist themselves represent another type of distraction (Stutts and Hunter, 2003)

## Conclusion

This chapter investigated the New Jersey crash data over a five-year period (2015–2019) in order to find the factors contributing to cellphone distractions by utilizing a mixed logit model and pseudo-elasticity analysis. Several variables were found to be statistically significant for various injury severity levels. AADT and driving on an urban roadway significantly decreases the likelihood of “injury” crashes, while factors like driving at night, violations of the speed limit, the total number of vehicles involved, and the involvement of a drugged driver significantly increase the “injury” severity of crashes due to cellphone distractions. Rear-end crashes, pedestrian or bicyclist involvement, wet surfaces, and driving on an undivided road significantly increased the likelihood of “possible injury” category. In contrast, the age of the driver, and the number of lanes significantly increased the likelihood of “no injury” severity.

These results indicate a complex interaction of various classes of variables behind those crashes that are due to distracted driving. The elasticity analysis found the impact of various parameters in increasing or decreasing the probability of crash severity. For instance, AADT, urban setting, and the number of lanes decreased the

probability of crash severity. However, the total number of vehicles involved, speed limit violations, the involvement of a drugged driver, driving on an undivided highway, night hours, and wet surfaces all increased the severity of crashes. Some of the features had a neutral effect on the different crash types. For instance, the involvement of young drivers, pedestrians, or bicyclists, and the presence of a wet surface demonstrated a neutral relationship to crash severity.

The outcomes from both the calibrated models indicate that prioritization should be made for some issues over others when countermeasures against cellphone distractions are considered. Increasing the number of divided roads, speed controls, the separation of pedestrians and bicycle riders from vehicles, increases to the visibility of pedestrians and bicycle riders (Chen, et al., 2012), increases in road friction features with innovative materials, and awareness-raising campaigns like *U-drive, U text, U Pay* can all decrease the severity of crashes due to cellphone distractions.

Overall, cellphone distractions can be addressed comprehensively by adopting a “three Es” approach, with Engineering countermeasures, Enforcement of laws, and education of people (Qi, Vennu, & Pokhrel, 2020). Engineering countermeasures inside the vehicle can help drivers reduce cellphone use. For example, Donmez et al. (2007) provided real-time feedback to drivers who were getting involved in secondary tasks (distracted), and they found that the feedback provided was beneficial to helping the drivers to return their attention to the road. To decrease the secondary visual tasks that distract drivers, Ford and Microsoft launched Ford SYNC in 2007. This technology seamlessly combines the utility of navigation, mobile communication, and audio device controls with voice commands, allowing drivers to operate these features while driving (Shah, 2013). Cellphone companies have also come up with various applications to discourage distracted driving. Oviedo-Trespalacios et al. (2019) has reviewed 29 different cellphone applications, which can block certain features like texting, notifications, and receiving calls while driving. The most widely used application by Android users was Android Auto, whereas AT&T DriveMode was the application used most by both iOS and Blackberry users.

When it comes to the other three “Es,” the rigorous enforcement of laws cannot be successful unless drivers are provided effective education on the dangers of distracted driving. It is a positive sign that the campaign “U Drive. U Text. U Pay” (NHTSA, 2019) has done great work in encouraging people to sign up and pledge to drive safely without distraction. Also, the enforcement of laws on the roadways has been found to be the most effective measure of restricting people from using their cellphones while driving (Cambridge Mobile Telematics, 2020). Education materials specifically designed for target groups prone to distracted driving behaviors (e.g., heavy vehicle drivers or young drivers) can be distributed via insurance companies, DMVs, and trucking companies to provide these drivers with the most up-to-date information on distracted driving (Cambridge Mobile Telematics, 2020).

# CHAPTER 4: Identifying Distracted Driving Events in New Jersey Using the Floating Car Method

## Introduction

To investigate distracted driving events, various techniques (e.g., surveys, videos, and simulations) have been introduced in the transportation environment. The strengths, findings, and gaps of these methods are elaborately discussed in chapter 2 of this report. From the summary of chapter 2, it was found that the majority of the observational research collected cross-sectional event data on individual subjects. Researchers investigated the findings from these studies and found that four main reasons contributed to the variation in the rate of distractions between different observational studies: different definitions of the various types of distraction (Charlton, 2009); differences in methodology (e.g., still photography vs. roadside observations), the regional difference (Adanu et al., 2017), and temporal biases (Prat et al., 2015). The type of roadway observed, the density of the traffic, seasonal variations, weather conditions, legislation, and strictness of law enforcement are all potential reasons for the regional and temporal variations (Chen & Lym, 2021; Astrain, Bernaus, Claverol, Escobar & Godoy, 2003).

These findings from the literature emphasize that driver behavior is different in every region, state, or country because of variations in legislation, roadway features, levels of enforcement, and the surrounding environment. Moreover, some of these studies were limited to weekdays (Gras et al., 2012), collected data during a specific season (Prat et al., 2015), or in urban locations only (Gras et al., 2012). The influence of changes in the season, the day of the week, or in the roadways themselves could not be observed from these studies (Young & Lenne, 2010; Johnson et al., 2004; Astrain, Bernaus, Claverol, Escobar & Godoy, 2003; Walter, 2010). Apart from these limitations, cross-sectional studies can provide driver behavior patterns for some specific observational points on roadways. The variations in driver behavior throughout the whole section of a corridor, which could be affected by many roadside geometric factors (e.g., posted speed limit, median type, number of lanes), could not be observed from these cross-sectional studies.

To minimize this spatial bias in the observations, an innovative longitudinal observational study was designed, which collects driver behavior data continuously throughout the whole stretch of the corridor. The goals of this chapter are to investigate the rate of distracted drivers in New Jersey and evaluate the effects of factors such as day of the week (weekday/weekend), type of roadway (signalized/unsignalized, toll/non-toll), geometric features of the roadway (posted speed limit, number of lanes, median type) and season (spring/summer) on driver distraction rates. It is expected that the findings of this study will be of value to engineers, researchers, and policymakers immensely in suggesting appropriate countermeasures to mitigate distracted driving events in the state of New Jersey.

# Methods

## *Site Selection*

In this study, six corridors with historical distracted driving crashes were selected for further investigation. These corridors include US1, US9, US130, I-80, US22, and Garden State Parkway. Apart from these high crash corridors, important arterial and interstate roads in New Jersey (I-295, I-95, NJ18, and NJ55) were also selected for data collection. Figure 6 illustrates the frequency-based plot of the top six corridors with the highest number of distracted driving crashes.

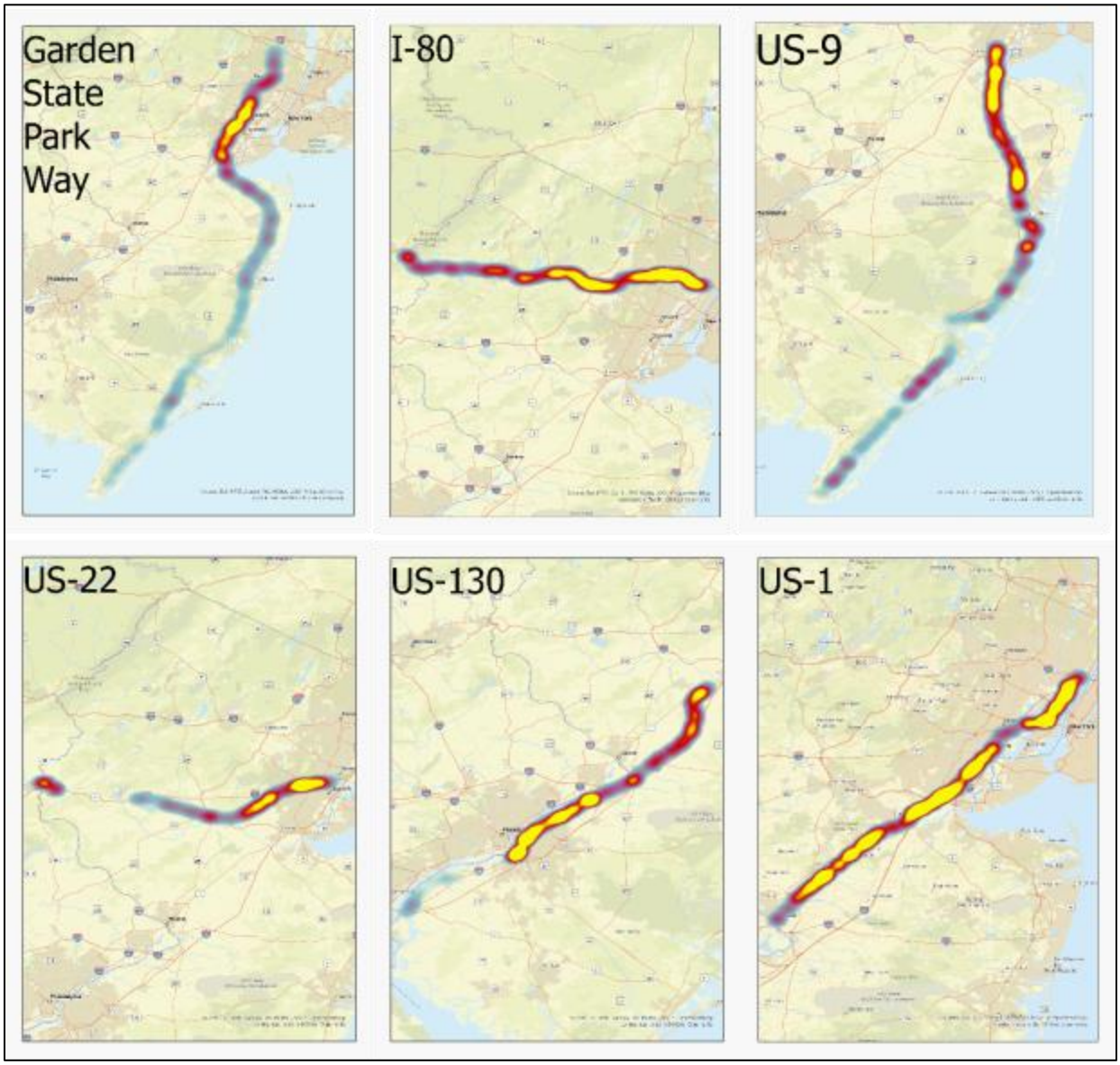



**Figure 6.** *A Map of Observational Study Locations*

Detailed information on all study corridors is listed in Table 4, including roadway type (toll/non-toll, signalized/unsignalized), total length, and total hours of observation. It is noteworthy that five rounds of data collection were performed during weekdays, and three rounds of data collection were completed during weekends. In addition, three rounds of data were collected during the spring, while five rounds were completed during the summer. Data were collected during both peak and off-peak hours for the same corridor.

**Table 4** *Detailed Information on the Selected Corridors*

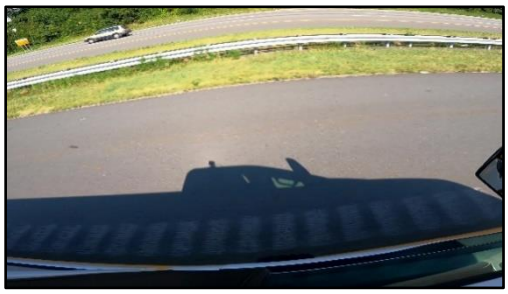
| Corridor | Signalized/ Unsignalized | Toll/ Non-Toll | Corridor Length in Miles (Round) | Total Miles of obs. | Total hours of obs. | AADT (2018) |
|---|---|---|---|---|---|---|
| **RT-18** | Signalized | Non-Toll | 85.5 | 855 | 25 | 27,424 |
| **US-1** | Signalized | Non-Toll | 76 | 760 | 25 | 31,395 |
| **US-130** | Signalized | Non-Toll | 156 | 1560 | 40 | 22,653 |
| **US-9** | Signalized | Non-Toll | 106 | 1060 | 50 | 25,836 |
| **US-22** | Signalized | Non-Toll | 80 | 800 | 20 | 29,933 |
| **RT-55** | Unsignalized | Non-Toll | 127 | 1270 | 25 | 27,819 |
| **I-295** | Unsignalized | Non-Toll | 142 | 1420 | 25 | 50,378 |
| **I-80** | Unsignalized | Non-Toll | 135 | 1350 | 25 | 61,355 |
| **I-95/ NJ Turnpike** | Unsignalized | Toll | 234 | 2340 | 40 | 60,213 |
| **Garden State Parkway** | Unsignalized | Toll | 342 | 3420 | 60 | 102,941 |

Variations in the geometric features of roadways (posted speed limit, number of lanes, median type, median width, and shoulder width) were recorded during data collection. The posted speed limit in selected corridors was divided into four groups: 25–35 mph, 36–45 mph, 46–55 mph, and 56–65 mph. The number of lanes was classified into three groups: two lanes, three lanes, and four or more lanes. The median width was divided into four groups: 0-10', 11'-20', 21'-30', and 31' or more. The shoulder width was categorized into four ranges: 0-3', 4'-6', 7'-9', and 10' or more. Finally, the median type was categorized into three groups: undivided, curbed, and positive. Different types of medians are illustrated in Figure 7. A comprehensive action map of the selected corridors is illustrated in Figure 8.

a) Curbed Median

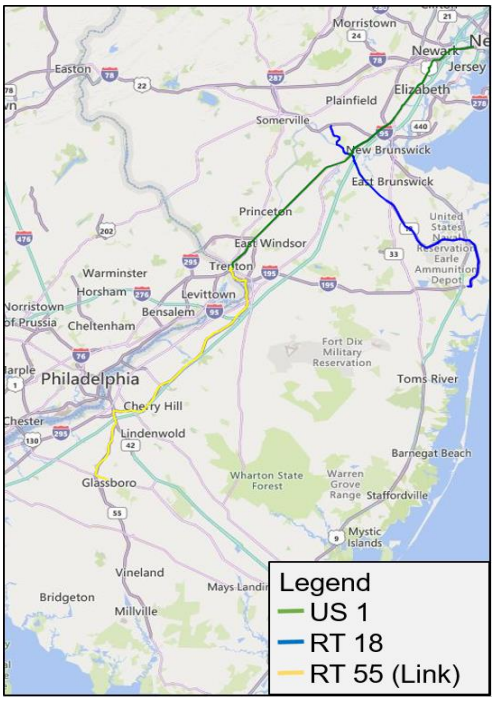
b) Unprotected Median

c) Positive Median (concrete)

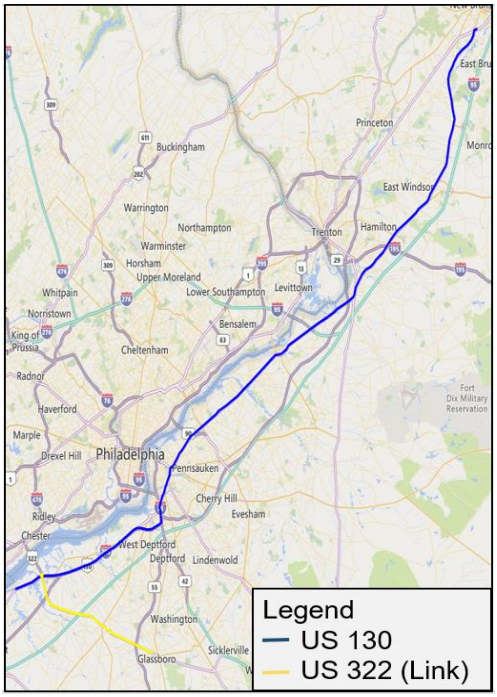


d) Positive Median (guardrail)

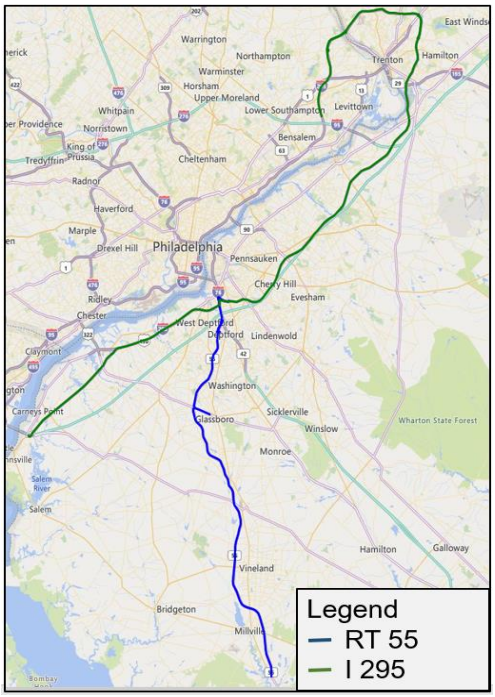
**Figure 7** *Different Types of Medians*



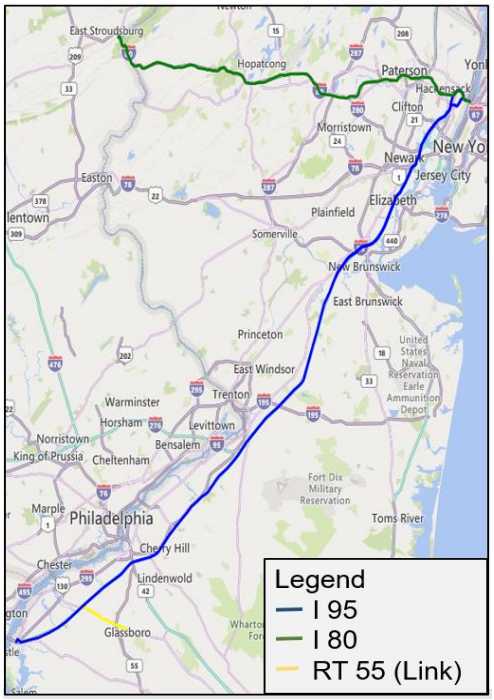

a) RT 18 and US1



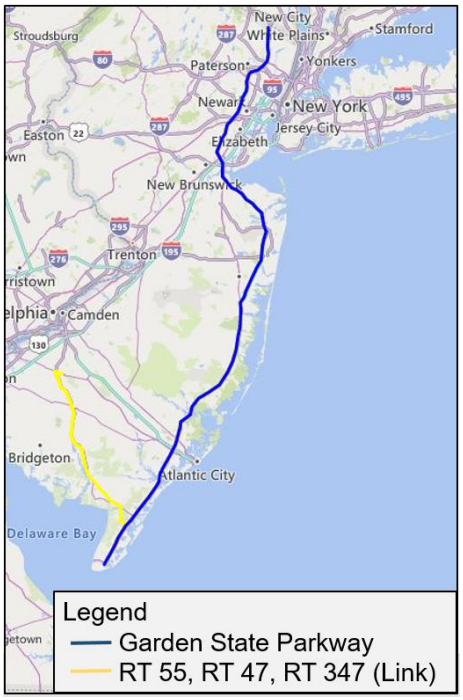

b) RT 130



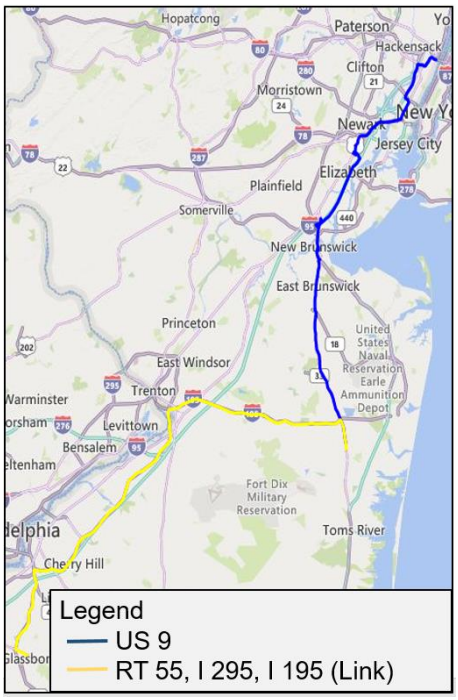

c) I-295 and RT55



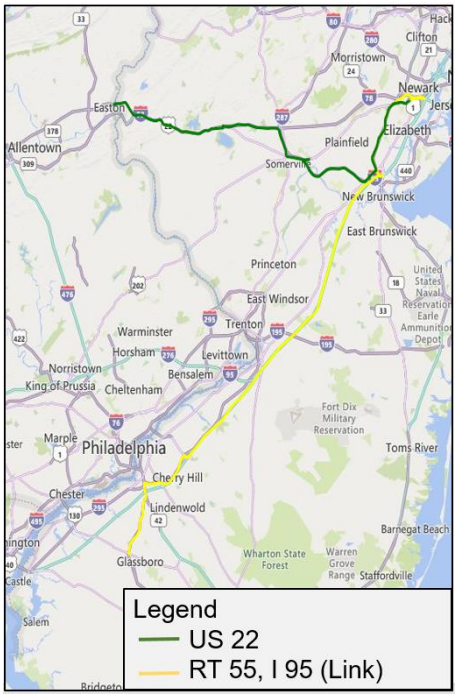

d) I-95 and I-80




e) Garden State Parkway




f) US9




g) US 22

**Figure 8** *Corridor Map for the Observational Study*

## *Data Collection*

For this observational study, a study participant drove through the selected corridors with a test vehicle. As the vehicle moved along, it would be passed or overtaken by vehicles from adjacent lanes. One of the study participants recorded driver behaviors in the adjacent lanes. The data collector used a mobile app named "Counter-Tally Count," accessible on iOS systems (iPhones and iPads). The 2021.6.3 version of this digital counter app was used, which was developed by Tevfik Yucek (Yucek, 2021). Eight categories (seven types of distraction and one non-distraction) were set in the counter app to record driver behaviors. Once the data collector places a passing vehicle into any one category, the counter app records the timestamp and the cumulative counts for that category. Following data collection, the data from the counter app can be exported into a CSV-formatted file, which gives an individual summary of all the types of distractions observed during the day. An illustration of the app is found in Figure 9.

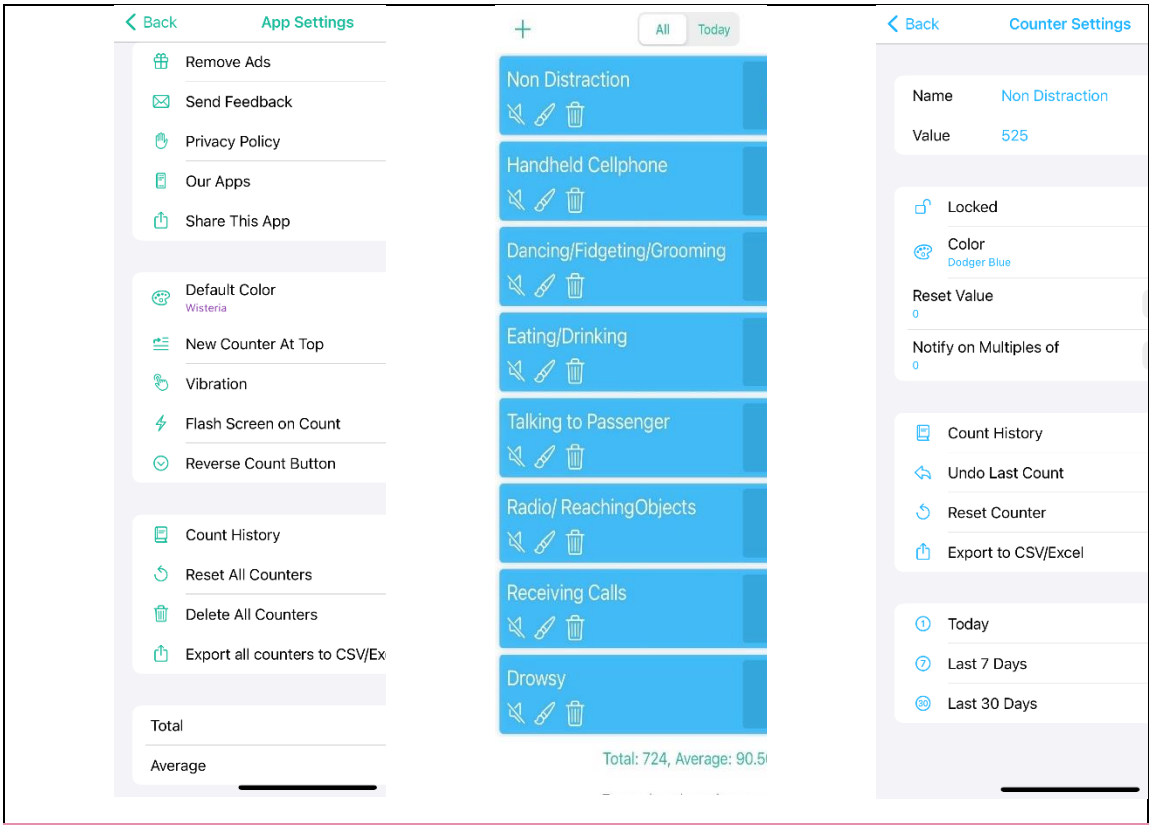



**Figure 9** *Illustration of "Counter- Tally Count" App*

The final dataset encompasses 335 hours and over 14,835 miles of data collected in New Jersey. To define the data collection time, the hourly variations in distracted driving crashes over a 14-year period (2006–2019) in New Jersey were assessed (Figure 10). As illustrated, the majority of distracted driving crashes happen

between 7 a.m. and 6 p.m. It was noted that, during daylight conditions, it is easier to reduce the effects of glare or reflection from mirrors.

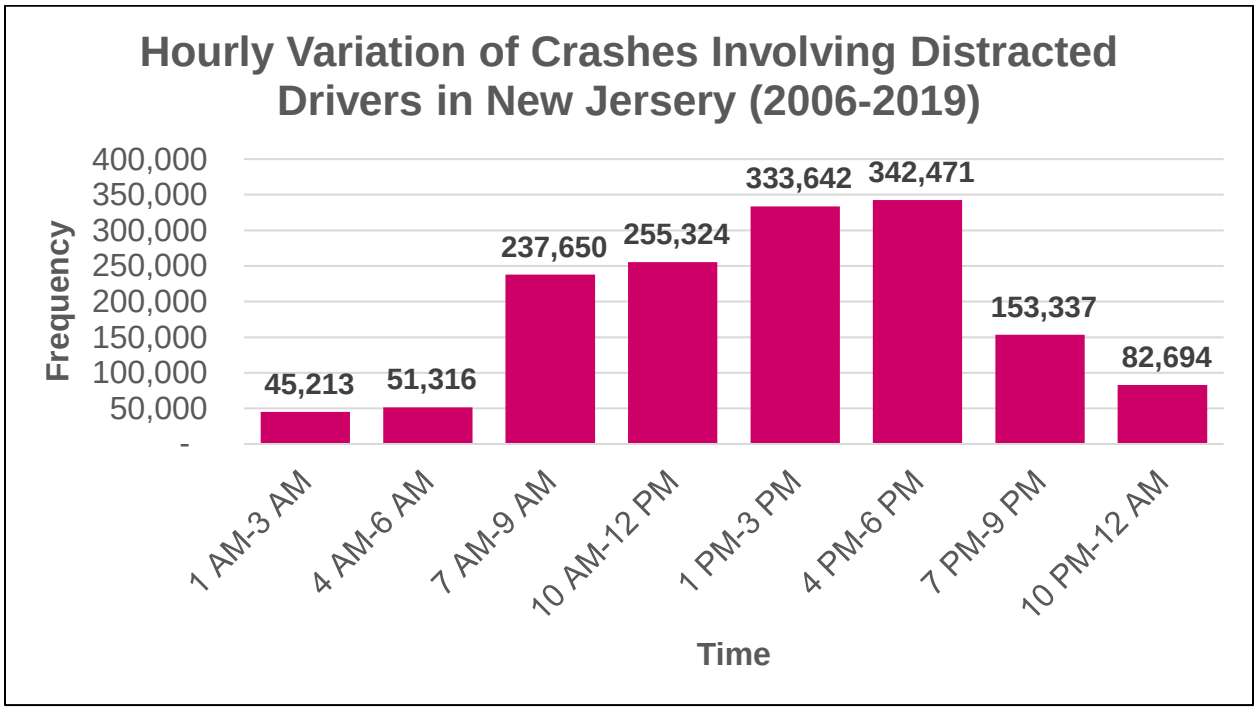



**Figure 10** *Hourly Variation of Crashes Involving Distracted Drivers in New Jersey (2006-2019) (Source: NJDHTS, 2021)*

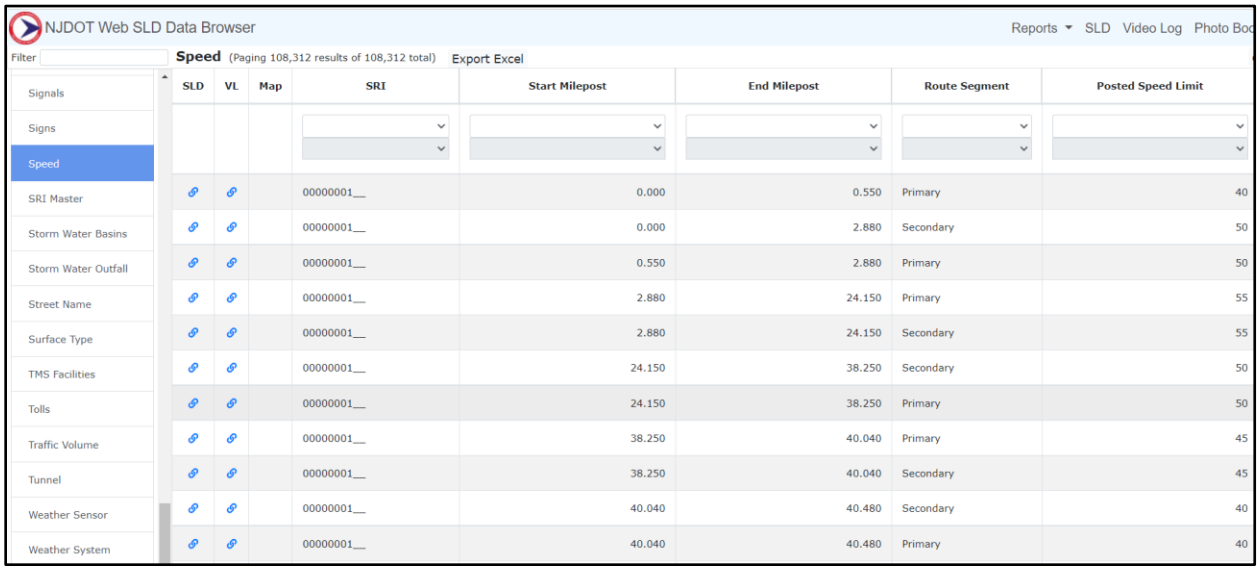
### *Data Integration*

Data integration is an essential step when determining the geolocation of distraction events and the geometric properties of the roadway (i.e., posted speed limit, number of lanes, median type, median width, shoulder width) for every distraction event. After data collection, the data was gathered, cleaned, and organized in accordance with the date and corridor name. This process is essential due to the variety of sources from which the data was obtained. A GPS tracking device (Figure 11) was used during the observational study to record the longitude and latitude at every second of the data collection, while the iOS application was recording a timestamp for every observation.


| | A | B | C |
|---|---|---|---|
| 1 | DateTime | Latitude | Longitude |
| 2 | 6/29/2021 9:20 | 39.08724 | -74.819 |
| 3 | 6/29/2021 9:20 | 39.08718 | -74.819 |
| 4 | 6/29/2021 9:20 | 39.08717 | -74.819 |
| 5 | 6/29/2021 9:20 | 39.08714 | -74.8191 |
| 6 | 6/29/2021 9:20 | 39.08709 | -74.8191 |
| 7 | 6/29/2021 9:20 | 39.08704 | -74.8192 |
| 8 | 6/29/2021 9:20 | 39.08698 | -74.8193 |
| 9 | 6/29/2021 9:20 | 39.08691 | -74.8193 |
| 10 | 6/29/2021 9:20 | 39.08683 | -74.8194 |

**Figure 11** *GPS Tracker with a Sample of Tracker Data*

Lastly, the web-based geographical database of the New Jersey Department of Transportation (NJDOT) was employed, which included the Standard Corridor Identifier (SRI) along with the geographical location of every 1/10$^{th}$ of a mile. Road features such as speed limits, lane counts, and median types were also provided by the NJDOT through their interactive online map, which was in accordance with the SRI and milepost location provided by the geographical database (Figure 12).



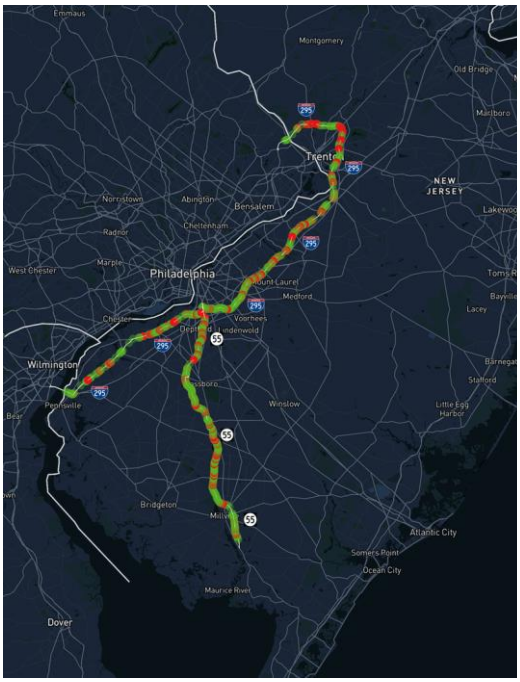

| SLD | VL | Map | SRI | Start Milepost | End Milepost | Route Segment | Posted Speed Limit |
|---|---|---|---|---|---|---|---|
| | | | 00000001__ | 0.000 | 0.550 | Primary | 40 |
| | | | 00000001__ | 0.000 | 2.880 | Secondary | 50 |
| | | | 00000001__ | 0.550 | 2.880 | Primary | 50 |
| | | | 00000001__ | 2.880 | 24.150 | Primary | 55 |
| | | | 00000001__ | 2.880 | 24.150 | Secondary | 55 |
| | | | 00000001__ | 24.150 | 38.250 | Secondary | 50 |
| | | | 00000001__ | 24.150 | 38.250 | Primary | 50 |
| | | | 00000001__ | 38.250 | 40.040 | Primary | 45 |
| | | | 00000001__ | 38.250 | 40.040 | Secondary | 45 |
| | | | 00000001__ | 40.040 | 40.480 | Secondary | 40 |
| | | | 00000001__ | 40.040 | 40.480 | Primary | 40 |

**Figure 12** *NJDOT SLD Web Browser (Source: NJDOT, 2021)*

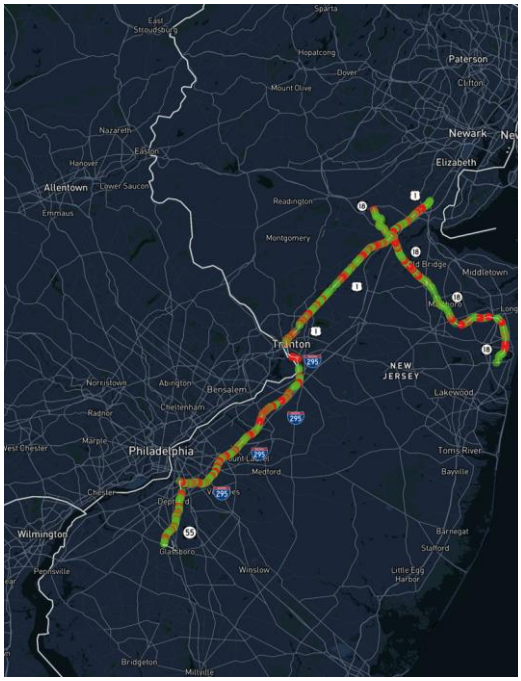
The clean-up and integration procedures were mainly aided by two pieces of software: Excel and ArcGIS Pro. The former was used to standardize the format of the data obtained from the iOS application and the tracking device. This specific data format was then required by ArcGIS Pro in order to join the two sources together in accordance with their time and date. In return, ArcGIS Pro estimates the location of each observation and allows a geographical data plot to be created on a map (Figure 13). All the distraction events were then marked as red dots on these maps, and they

will be helpful for locating those locations that are most prone to distracted driving events.

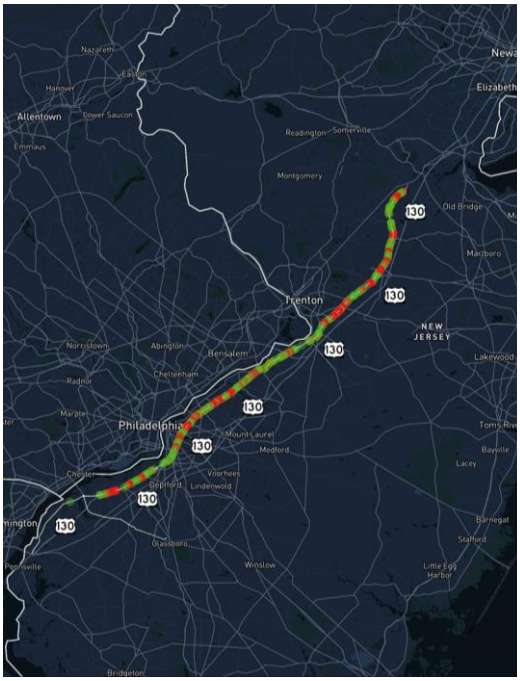

a) Distraction Events in US9



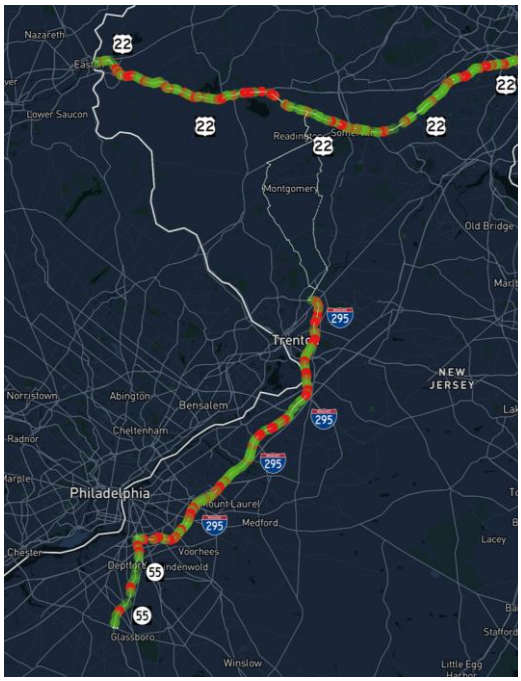

b) Distraction Events in I-295 & RT 55



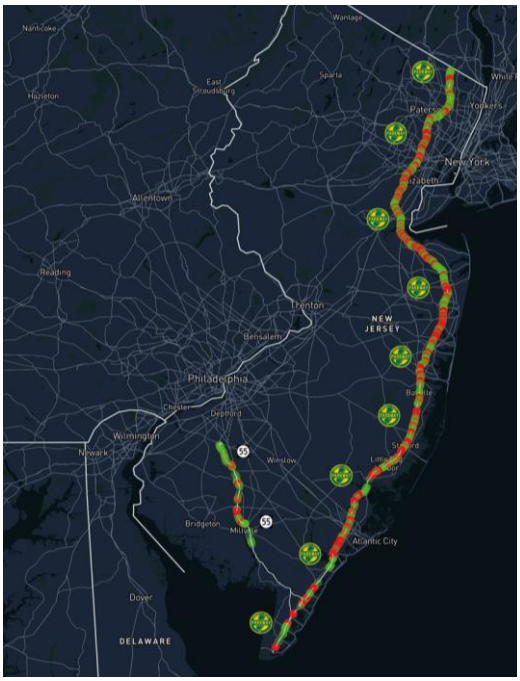

c) Distraction Events in US1 & RT 18



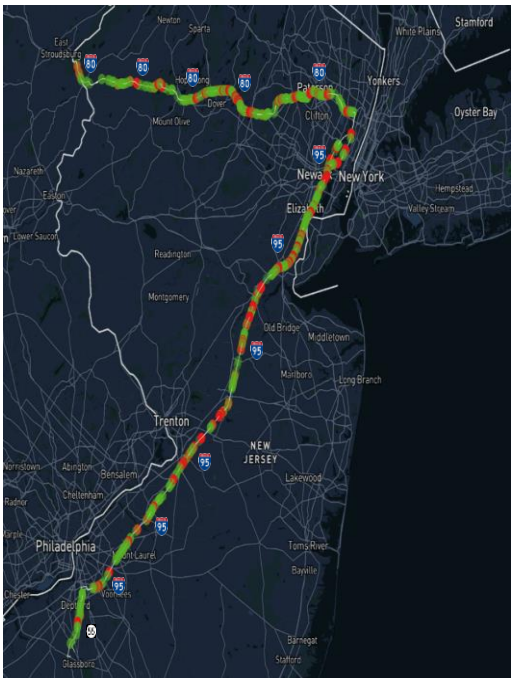

d) Distraction Events in US 130




e) Distraction Events in US 22




f) Distraction Events in Garden State Parkway




g) Distraction Events in I-95 and I-80

**Figure 13** *Distraction Events (Red Dots) in the Selected Corridors*

Finally, information about road features (i.e., posted speed limit, median type, and the number of lanes) was integrated into Excel by exporting the features of each SRI from the interactive event map and coordinating each geometric feature with the event data. The output file in Excel contains the event data of each type of distraction along with the geometric properties of the roadway at the time of the event (one example of the integrated file is illustrated in Figure 14). This integrated file is useful to correlate the temporal features and roadway geometric attributes of each data point. For example, data point 9 on the corridor US22 is located on a curbed median with a three-lane road and a posted speed limit of 25 mph.

| | E | G | H | I | P | Q | S | T | U | V |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Time | Distraction Index | Lat | Long | SRI | MP | SLD_NAMI | Posted Speed Limit | No. Lanes | Median Type |
| 2 | 7/1/2021 13:54 | 10 | 40.69574 | -75.1891 | 00000022_ | 0.8 | US 22 | 25 | 2 | Positive |
| 3 | 7/1/2021 13:54 | 1 | 40.69569 | -75.1895 | 00000022_ | 0.8 | US 22 | 25 | 2 | Positive |
| 4 | 7/1/2021 13:54 | 1 | 40.69618 | -75.1851 | 00000022_ | 1 | US 22 | 25 | 3 | Curbed |
| 5 | 7/1/2021 13:51 | 11 | 40.69654 | -75.1838 | 00000022_ | 1.1 | US 22 | 25 | 3 | Curbed |
| 6 | 7/1/2021 13:51 | 12 | 40.69655 | -75.1838 | 00000022_ | 1.1 | US 22 | 25 | 3 | Curbed |
| 7 | 7/1/2021 13:54 | 12 | 40.69635 | -75.1838 | 00000022_ | 1.1 | US 22 | 25 | 3 | Curbed |
| 8 | 7/1/2021 13:51 | 8 | 40.69653 | -75.1839 | 00000022_ | 1.1 | US 22 | 25 | 3 | Curbed |
| 9 | 7/1/2021 13:54 | 7 | 40.69624 | -75.1849 | 00000022_ | 1.1 | US 22 SEC | 25 | 3 | Curbed |
| 10 | 7/1/2021 13:51 | 3 | 40.69649 | -75.1843 | 00000022_ | 1.1 | US 22 | 25 | 3 | Curbed |
| 11 | 7/1/2021 13:51 | 1 | 40.69656 | -75.1838 | 00000022_ | 1.1 | US 22 | 25 | 3 | Curbed |
| 12 | 7/1/2021 13:51 | 1 | 40.69653 | -75.1839 | 00000022_ | 1.1 | US 22 | 25 | 3 | Curbed |
| 13 | 7/1/2021 13:51 | 1 | 40.69656 | -75.1838 | 00000022_ | 1.1 | US 22 | 25 | 3 | Curbed |
| 14 | 7/1/2021 13:54 | 1 | 40.69622 | -75.1848 | 00000022_ | 1.1 | US 22 SEC | 25 | 3 | Curbed |
| 15 | 7/1/2021 13:54 | 1 | 40.69622 | -75.1848 | 00000022_ | 1.1 | US 22 SEC | 25 | 3 | Curbed |
| 16 | 7/1/2021 13:54 | 1 | 40.69628 | -75.1842 | 00000022_ | 1.1 | US 22 | 25 | 3 | Curbed |
| 17 | 7/1/2021 13:55 | 8 | 40.69722 | -75.1776 | 00000022_ | 1.5 | US 22 SEC | 35 | 3 | Curbed |
| 18 | 7/1/2021 13:55 | 8 | 40.69723 | -75.1776 | 00000022_ | 1.5 | US 22 SEC | 35 | 3 | Curbed |
| 19 | 7/1/2021 13:56 | 1 | 40.69728 | -75.1774 | 00000022_ | 1.5 | US 22 SEC | 35 | 3 | Curbed |
| 20 | 7/1/2021 13:57 | 1 | 40.69753 | -75.1761 | 00000022_ | 1.5 | US 22 | 35 | 3 | Curbed |
| 21 | 7/1/2021 13:57 | 8 | 40.69771 | -75.1744 | 00000022_ | 1.6 | US 22 | 35 | 3 | Positive |
| 22 | 7/1/2021 13:57 | 7 | 40.69754 | -75.1759 | 00000022_ | 1.6 | US 22 SEC | 35 | 3 | Positive |
| 23 | 7/1/2021 13:50 | 3 | 40.69771 | -75.1756 | 00000022_ | 1.6 | US 22 SEC | 35 | 3 | Positive |
| 24 | 7/1/2021 13:50 | 1 | 40.69791 | -75.1745 | 00000022_ | 1.6 | US 22 | 35 | 3 | Positive |
| 25 | 7/1/2021 13:49 | 1 | 40.69847 | -75.1638 | 00000022_ | 2.3 | US 22 SEC | 35 | 3 | Positive |

**Figure 14** *Integrated Output File with Event Data and Geometric Features of Roadways*

A detailed flowchart of the data integration procedure is illustrated in Figure 15.

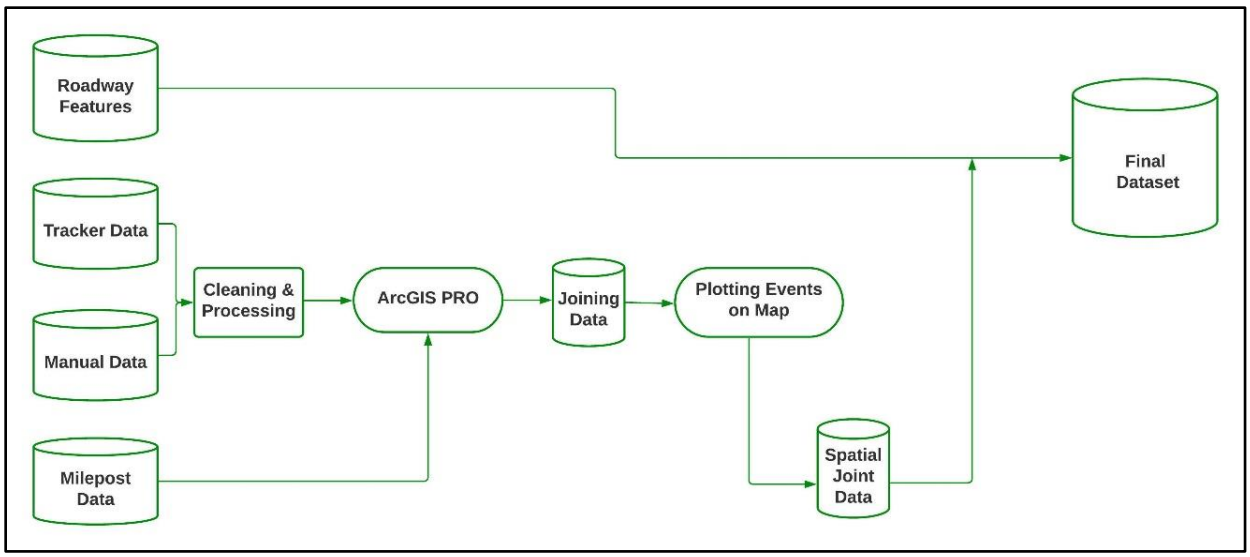



**Figure 15** *Steps of Data Integration*

### *Definitions of the Various Classes of Distractions*

The definitions of the different classes of distractions (secondary activities) were developed by refining categories previously used by researchers (Prat et al., 2015; Sullman, 2012; Gras et al., 2012). Moreover, multiple distractions could be coded for in the same driver (e.g., a person could both be smoking and tuning the radio). The categories and definitions used were:

1. Handheld Cellphone: The driver is holding or using a cellphone with his or her hands.
2. Receiving Calls: A cellphone is held near the ear of the driver.
3. Eating/Drinking: The driver is holding a cup, food, or a cigarette, or is seen eating or drinking.
4. Radio/Reaching Object: Their hands are reaching toward the radio or some other place on the dash.
5. Fidgeting/Grooming: Their hands are touching their face.
6. Drowsy: The driver is yawning.
7. Talking to Passenger: Their eye or face orientation is directed toward the passenger side.
8. Non-Distracted: Their hands are on the steering wheel.

### *Statistical Analysis*

**The Pearson Correlation Test.** The Pearson correlation test is widely used in statistical analysis to find the correlation or dependency among various parameters. The Pearson's correlation coefficient is defined by equation 7 (Thapngam et al., 2011).

$$\rho_{x,y} = \frac{E[(X-\mu_x)(Y-\mu_y)]}{\sigma_x \sigma_y} \quad (7)$$

Where, $\mu_x$ and $\mu_y$ are the expected values of the two variables X and Y; and $\sigma_x$ and $\sigma_y$ are the standard deviations of these values. The value of correlation varies from -1 to +1. According to the Pearson correlation test, any correlation value less than 0.3 is considered to have poor correlation (Ratner, 2009).

**The Mann–Whitney U Test.** The Mann–Whitney U test is a non-parametric method used to confirm if two independent sample means are equal. The test does not make any assumptions related to the distribution of scores. Initially, the test was proposed for equal sample sizes, but its application was later extended to unequal sample sizes. It should be noted that when the ranks of the two samples are collected from identical categories (e.g., signalized vs. unsignalized road segments, spring vs. summer, toll vs. non-toll roads, peak vs. off-peak hours, and weekday vs. weekends), the results for both samples can be expected to have an equal mean rank. This represents the case when the null hypothesis is true. However, if the sample result is affected by an independent variable, then it can be expected to impact their rank order and even cause the mean ranks of the two samples to be different. This represents the case when the null hypothesis is false. The calculation procedure for the Mann–Whitney test is as follows:

$$U_1 = R_1 - \frac{n_1(n_1+1)}{2} \quad (8)$$

$$U_2 = R_2 - \frac{n_2(n_2+1)}{2} \quad (9)$$

where $U_1$ and $U_2$ are the Mann–Whitney results for two different variations in the data (i.e., spring vs. summer, weekday vs. weekend, toll road vs. non-toll road, peak vs. off-peak hours, and signalized vs. unsignalized road); $n_1$ and $n_2$ are the numbers of events for the variations; $R_1$ and $R_2$ are rank sums for the variations. When the U value is less than or equal to the critical value, the two samples are statistically significant.

As a part of this analysis, a Python package was used to perform the Mann–Whitney U test. This package provided the mean ranks for each type of distraction using this test, as well as various other outcomes, such as mean ranks for each group, Mann–Whitney U statistics, Z-scores, and p-values. To be more specific, the Mann–Whitney U statistics consider the lowest sum of the rank for computing the p-value and identifying the statistical significance. Note that the package uses the approximation that the samples have a standard normal distribution, which is necessary in order to give the Z-score and the p-value (Salkind, 2010; Patel, 2020).

**The Kruskal–Wallis Test.** The Kruskal–Wallis test is a non-parametric, one-way analysis of variance. This test is a generalized form of the Mann–Whitney U test (Lambert, 2005) and is used when the number of groups to be compared is more

than two (Hashim et al., 2011; Sack et al., 2019). This test does not need the data to be normally distributed. However, the data points under consideration should be independent of each other (Smalheiser, 2017). The null hypothesis for this test states that there is no difference in mean rank between samples drawn from different groups (e.g., posted speed limit, number of lanes, median type). The alternative hypothesis is that the difference between the mean ranks of some groups is large enough to be considered statistically significant (Lambert, 2005).

In the Kruskal–Wallis test, the average rank for each sample or group is computed. If the samples are drawn from populations with different means, then the average rank will differ. The differences against the average ranks are assessed to test if the samples are drawn from the same population or not. Supposing the same number of samples are drawn from the same population, then the sampling distribution of the Kruskal–Wallis test statistic and the probability of observing the different values can be tabled.

It is noteworthy that, in the Kruskal–Wallis test, if the number of groups being compared exceeds the value of three, and if the number of observations in each group exceeds five, then the sampling distribution is well-approximated by the chi-square distribution (Hoffman, 2015). For this research, all the geometric features (posted speed limit, number of lanes, and median type) have at least three groups or variations for comparison. The approximation with the chi-square distribution performs better with increases in both the number of groups and the number of observations in each group.

The Kruskal–Wallis test is based on the following assumptions, “The observations in the data set are independent of each other; The distribution of the population is not necessarily normal, and the variances should not necessarily be equal; Observations are drawn from the population by random sampling” (Statistics Solutions, 2022).

The sample sizes in the Kruskal–Wallis test should be as equal as possible, but some differences are allowed. For this analysis, the critical value of the chi-square test was compared with the H-statistic from Kruskal–Wallis. Suppose the value of H is greater than the critical value of chi-square. In that case, the null hypothesis is rejected, which means there is a statistically significant difference of mean rank between at least two groups of the observations (Hoffman, 2015). The equation for the H-statistic of the Kruskal–Wallis test is as follows (Elliot and Hynan, 2011):

$$H = \left[\frac{12}{n(n+1)} \sum \frac{R_i^2}{n_i}\right] - 3(n+1) \qquad (10)$$

where,

$n_i$ = the number of items in sample *i*;

$R_i$ = the sum of the ranks of all items in sample *i*;

K = the total number of samples; and

$n = n_1 + n_2 + ..... + n_K$ is the total number of observations in all samples.

**P-value.** Determining the level of significance is necessary when analyzing the results of these observations. Usually, the value of the significance level varies between 0 and 1. It should be mentioned that researchers most often use significance values of 0.01, 0.05, or 0.10, corresponding to 99%, 95%, and 90% confidence levels, respectively. This study considered the significance level to be 95%. For a change to be statistically significant at the 95% level ($\alpha = 0.05$), the P-value must be less than 0.05 (Patel, 2020).

**Effect Size.** The effect size for the sample data is calculated by dividing the absolute standardized test statistic *z* by the square root of the total sample size *n*, as follows:

$$\text{Effect Size} = \frac{z}{\sqrt{n}} \tag{11}$$

Cohen's classification of effect size was used here to determine whether the changes are statistically significant. According to Cohen's classification, an effect size of between 0.1 and 0.3 is considered to have a small effect; between 0.3 and 0.5 to have a moderate effect; and 0.5 and above a large effect (Patel, 2020).

## Results

### *Sample Size Validation*

To analyze data on the subgroup level (e.g., with respect to the day of the week, the season, and the roadway type), the sample size should be representative of the total population. It was noted that most previous research studies used a 95% confidence interval with a margin of error of between 5% and 10% (Patel, 2020). For this study, the Annual Average Daily Traffic (AADT) is considered the population for the sample size for each selected corridor, indicating that the collected data is a representative sample of the total average daily traffic on the corridor. As shown in Figure 16, all the sample sizes of the collected corridors satisfy the minimum sample size requirement with a 99% confidence level and a 5% margin of error.

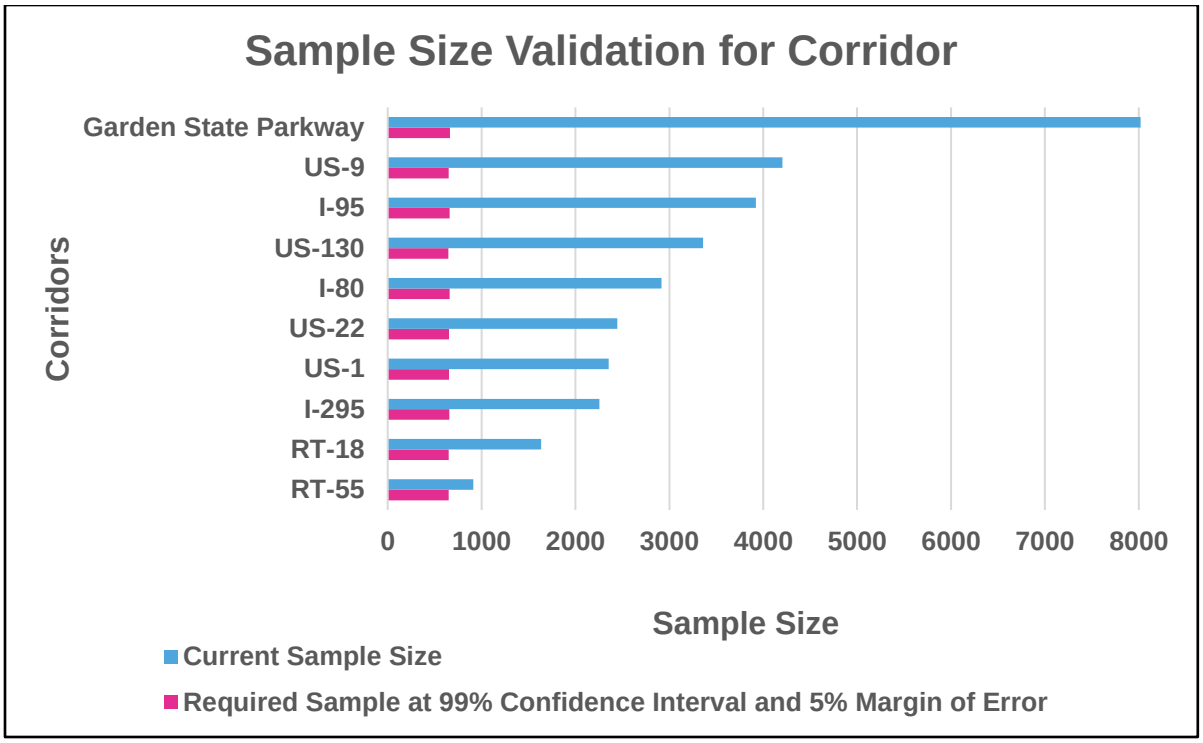



**Figure 16** *Sample size validation for the selected corridors*

### *Pearson Correlation Test*

A Pearson correlation test was conducted to test the correlation among the roadway's geometric properties (posted speed limit, median type, number of lanes, median width, and shoulder width). According to Table 5, there is poor and moderate correlation among the studied geometric properties. These findings imply that the roadways' geometric properties (posted speed limit, median type, number of lanes, median width, and shoulder width) are not dependent on one another.

**Table 5** *Correlation among Geometric Properties of Roadway*

| Geometric Property | Median Type | Posted Speed Limit | Number of Lanes | Median Width | Shoulder Width |
|---|---|---|---|---|---|
| **Median Type** | 1.00 | 0.23 | -0.21 | 0.28 | 0.47 |
| **Posted Speed Limit** | 0.23 | 1.00 | 0.08 | -0.22 | -0.07 |
| **Number of Lanes** | -0.21 | 0.08 | 1.00 | -0.22 | -0.11 |
| **Median Width** | 0.28 | -0.22 | -0.22 | 1.00 | 0.17 |
| **Shoulder Width** | 0.47 | -0.07 | -0.11 | 0.17 | 1.00 |

### *Data Description*

The details of the distraction events data is illustrated in table 6. Table 6 contains the number of various types of distractions for each of the ten routes. Eight

rounds of data were recorded of each route. In total, the behavior of 42,130 drivers was collected throughout the period of data collection, out of which 9,375 drivers were found distracted (22.3%). The rate of distraction for various routes ranges from 12.6% to 35.5%. However, for a better understanding of the data and the comparison of the distraction in various routes, an average value of distraction for each route is also taken. From the average rates of distractions, it was found that US 1 (25.8%) and Rt 18 (23.5%) had the highest rate of distractions while the minimum average rates of distraction are found in I-295 (22.1%), US-9 (21.2%) and Rt 55 (21.2%).

**Table 6** *Distraction Rates of various routes*

| Road Segment | Fidgeting/ Grooming | Radio/ Reaching Obj | Drowsy | Talking to Passenger | Calling | Eating/ Drinking | Handheld cell phone | Non- Distraction | Distraction Rate (%) |
|---|---|---|---|---|---|---|---|---|---|
| US-9 | 25 | 15 | 5 | 13 | 18 | 10 | 82 | 1165 | 12.6 |
| | 25 | 29 | 1 | 0 | 10 | 19 | 59 | 652 | 18 |
| | 3 | 2 | 0 | 0 | 0 | 1 | 6 | 59 | 16.9 |
| | 14 | 17 | 2 | 8 | 9 | 26 | 49 | 536 | 18.9 |
| | 49 | 18 | 2 | 8 | 7 | 27 | 37 | 390 | 27.5 |
| | 46 | 35 | 5 | 10 | 8 | 27 | 72 | 602 | 25.2 |
| | 30 | 8 | 3 | 3 | 2 | 15 | 25 | 280 | 23.5 |
| | 41 | 23 | 1 | 8 | 10 | 33 | 56 | 463 | 27.1 |
| US-22 | 27 | 13 | 3 | 11 | 1 | 15 | 19 | 329 | 21.3 |
| | 33 | 11 | 4 | 2 | 3 | 11 | 14 | 343 | 18.5 |
| | 19 | 19 | 1 | 5 | 3 | 10 | 40 | 282 | 25.6 |
| | 24 | 17 | 2 | 6 | 7 | 22 | 31 | 268 | 28.9 |
| | 33 | 13 | 1 | 7 | 3 | 16 | 26 | 283 | 25.9 |
| | 37 | 11 | 0 | 2 | 7 | 28 | 20 | 363 | 22.4 |
| | 23 | 9 | 1 | 7 | 2 | 15 | 19 | 275 | 21.7 |
| | 23 | 7 | 4 | 1 | 8 | 10 | 14 | 290 | 18.8 |
| US-130 | 19 | 22 | 2 | 4 | 7 | 22 | 47 | 435 | 22 |
| | 17 | 13 | 0 | 5 | 4 | 30 | 32 | 438 | 18.7 |
| | 41 | 16 | 1 | 18 | 2 | 23 | 31 | 356 | 27 |
| | 21 | 23 | 1 | 2 | 6 | 24 | 45 | 457 | 21.1 |
| | 34 | 14 | 3 | 6 | 6 | 30 | 58 | 479 | 24 |
| | 33 | 20 | 1 | 2 | 9 | 34 | 45 | 420 | 25.5 |
| | 49 | 12 | 5 | 13 | 9 | 38 | 49 | 549 | 24.2 |
| | 29 | 14 | 2 | 9 | 3 | 20 | 25 | 351 | 22.5 |
| US-1 | 14 | 27 | 4 | 13 | 8 | 17 | 41 | 414 | 23 |
| | 29 | 9 | 0 | 2 | 3 | 17 | 39 | 367 | 21.2 |
| | 26 | 16 | 0 | 12 | 1 | 18 | 27 | 275 | 26.7 |
| | 33 | 26 | 1 | 3 | 4 | 10 | 39 | 284 | 29 |
| | 26 | 11 | 4 | 1 | 2 | 20 | 16 | 237 | 25.2 |
| | 31 | 6 | 1 | 3 | 3 | 18 | 30 | 167 | 35.5 |
| | 32 | 13 | 0 | 10 | 4 | 22 | 20 | 309 | 24.6 |
| | 15 | 11 | 1 | 2 | 2 | 10 | 27 | 257 | 20.9 |
| RT-55 | 2 | 7 | 2 | 1 | 1 | 2 | 25 | 162 | 19.8 |

| Road Segment | Fidgeting/ Grooming | Radio/ Reaching Obj | Drowsy | Talking to Passenger | Calling | Eating/ Drinking | Handheld cell phone | Non- Distraction | Distraction Rate (%) |
|---|---|---|---|---|---|---|---|---|---|
| | 4 | 1 | 0 | 2 | 2 | 6 | 15 | 132 | 18.5 |
| | 11 | 2 | 0 | 2 | 4 | 3 | 9 | 88 | 26.1 |
| | 15 | 5 | 0 | 7 | 0 | 6 | 5 | 135 | 22 |
| | 2 | 5 | 0 | 2 | 0 | 2 | 6 | 105 | 13.9 |
| | 10 | 9 | 0 | 2 | 1 | 6 | 5 | 101 | 24.6 |
| | 10 | 3 | 1 | 2 | 5 | 7 | 16 | 120 | 26.8 |
| | 14 | 3 | 0 | 2 | 2 | 6 | 16 | 199 | 17.8 |
| **RT-18** | 6 | 23 | 3 | 4 | 6 | 8 | 24 | 315 | 19 |
| | 10 | 11 | 0 | 2 | 1 | 7 | 15 | 169 | 21.4 |
| | 24 | 8 | 2 | 3 | 2 | 5 | 7 | 189 | 21.3 |
| | 23 | 15 | 4 | 4 | 7 | 14 | 22 | 224 | 28.4 |
| | 20 | 7 | 0 | 2 | 5 | 9 | 14 | 169 | 25.2 |
| | 16 | 8 | 1 | 5 | 1 | 11 | 19 | 191 | 24.2 |
| | 21 | 5 | 2 | 7 | 2 | 17 | 37 | 299 | 23.3 |
| | 5 | 4 | 2 | 6 | 4 | 6 | 11 | 114 | 25 |
| **I-95** | 14 | 44 | 5 | 17 | 14 | 24 | 47 | 473 | 25.9 |
| | 23 | 19 | 1 | 9 | 11 | 19 | 42 | 579 | 17.6 |
| | 9 | 32 | 2 | 5 | 12 | 22 | 38 | 439 | 21.5 |
| | 33 | 24 | 2 | 6 | 9 | 26 | 44 | 458 | 23.9 |
| | 51 | 20 | 3 | 17 | 5 | 31 | 49 | 597 | 22.8 |
| | 35 | 25 | 1 | 3 | 12 | 34 | 65 | 472 | 27 |
| | 43 | 10 | 4 | 6 | 5 | 20 | 37 | 440 | 22.1 |
| | 41 | 9 | 3 | 10 | 1 | 17 | 16 | 401 | 19.5 |
| **I-80** | 22 | 23 | 1 | 11 | 7 | 13 | 18 | 344 | 21.6 |
| | 14 | 17 | 0 | 7 | 3 | 16 | 27 | 409 | 17 |
| | 21 | 11 | 0 | 6 | 7 | 14 | 34 | 300 | 23.7 |
| | 10 | 11 | 0 | 2 | 6 | 8 | 17 | 208 | 20.6 |
| | 29 | 32 | 3 | 7 | 9 | 26 | 28 | 414 | 24.5 |
| | 47 | 34 | 2 | 6 | 16 | 50 | 70 | 555 | 28.8 |
| | 27 | 9 | 4 | 2 | 6 | 25 | 23 | 357 | 21.2 |
| | 13 | 5 | 0 | 5 | 3 | 15 | 21 | 329 | 15.9 |
| **I-295** | 14 | 11 | 2 | 11 | 2 | 6 | 7 | 331 | 13.8 |
| | 14 | 17 | 1 | 6 | 4 | 10 | 35 | 364 | 19.3 |
| | 37 | 9 | 1 | 13 | 7 | 15 | 20 | 343 | 22.9 |
| | 24 | 27 | 3 | 15 | 7 | 16 | 38 | 358 | 26.6 |
| | 29 | 16 | 0 | 8 | 5 | 30 | 29 | 324 | 26.5 |
| | 22 | 12 | 1 | 4 | 3 | 10 | 16 | 277 | 19.7 |
| | 15 | 8 | 2 | 2 | 11 | 18 | 16 | 254 | 22.1 |
| | 39 | 17 | 2 | 6 | 5 | 25 | 28 | 558 | 17.9 |
| **Garden State Parkway** | 65 | 73 | 7 | 14 | 25 | 26 | 115 | 1002 | 24.5 |
| | 33 | 55 | 5 | 13 | 17 | 26 | 70 | 756 | 22.5 |
| | 74 | 65 | 6 | 27 | 18 | 64 | 110 | 1199 | 23.3 |
| | 68 | 28 | 6 | 9 | 31 | 41 | 93 | 1112 | 19.9 |

| Road Segment | Fidgeting/ Grooming | Radio/ Reaching Obj | Drowsy | Talking to Passenger | Calling | Eating/ Drinking | Handheld cell phone | Non- Distraction | Distraction Rate (%) |
|---|---|---|---|---|---|---|---|---|---|
| | 63 | 31 | 3 | 11 | 23 | 42 | 99 | 915 | 22.9 |
| | 94 | 34 | 5 | 30 | 11 | 39 | 89 | 1277 | 19.1 |
| | 75 | 22 | 5 | 15 | 17 | 31 | 80 | 875 | 21.9 |
| | 92 | 15 | 5 | 20 | 18 | 35 | 77 | 948 | 21.7 |

## *Summary of Distractions*

The details of the distraction events data are illustrated in table 7. From the table, we can observe that handheld cellphone had the highest rate of distractions, irrespective of the temporal variety or variation in roadway geometry. The distraction rate on weekdays (22.9%) is higher than on the weekends (22.2%). Distraction rate in peak hours (24.6) is higher than the off-peak hours, while the signalized roads had a higher distraction rate than the unsignalized roads. Toll roads and non-toll roads have almost similar rates of distraction. The distraction rate of summer was 3.6% more than the distraction rate of the spring season. In terms of median type, the curbed median had more rate of distractions over the positive or undivided medians. 3 lane roads encountered more rate of distraction compared to the 2 lane and 4 or more lane roads. The rate of distraction was more in the speed range of 25-35 mph compared to the other speed limits. A median width less than 10' had a higher rate of distraction compared to the other categories of shoulder width.

**Table 7** *Distraction Rates of various classes for temporal and geometric variation*

| Temporal/ Roadway Feature | Handheld cell phone | Fidgeting/ Grooming | Eating/ Drinking | Radio/ Reaching Object | Talking To Passenger | Receiving Calls | Drowsy | Non-Distraction | Total Distraction Rate (%) |
|---|---|---|---|---|---|---|---|---|---|
| **Distraction Rates (%)** | | | | | | | | | |
| ***Day of Week*** | | | | | | | | | |
| **Weekday** | 6.8 | 5.5 | 3.7 | 3.8 | 1.4 | 1.3 | 0.4 | 77.1 | 22.9 |
| **Weekend** | 7.1 | 5.3 | 3.6 | 3.6 | 1.4 | 1.0 | 0.2 | 77.8 | 22.2 |
| ***Hour of Day*** | | | | | | | | | |
| **Peak Hour** | 7.1 | 6.6 | 4.0 | 3.6 | 1.4 | 1.4 | 0.5 | 75.4 | 24.6 |
| **Off-Peak Hour** | 6.6 | 5.3 | 3.7 | 3.7 | 1.2 | 1.1 | 0.4 | 78.0 | 22.0 |
| ***Signalized/ Unsignalized*** | | | | | | | | | |
| **Signalized Road** | 7.2 | 6.0 | 3.9 | 3.6 | 1.2 | 1.0 | 0.4 | 76.7 | 23.3 |
| **Unsignalized Road** | 6.7 | 4.8 | 3.4 | 3.8 | 1.6 | 1.4 | 0.3 | 78.0 | 22.0 |
| ***Toll/Non-toll road*** | | | | | | | | | |
| **Toll Road** | 7.3 | 4.5 | 3.5 | 4.0 | 1.3 | 1.6 | 0.4 | 77.4 | 22.6 |
| **Non-Toll Road** | 6.8 | 5.7 | 3.7 | 3.6 | 1.4 | 1.1 | 0.4 | 77.3 | 22.7 |
| ***Season*** | | | | | | | | | |
| **Spring Season** | 6.7 | 4.5 | 3.0 | 3.6 | 1.5 | 1.2 | 0.3 | 79.2 | 20.8 |
| **Summer Season** | 7.1 | 6.3 | 4.3 | 3.8 | 1.3 | 1.3 | 0.3 | 75.6 | 24.4 |

| Temporal/ Roadway Feature | Handheld cell phone | Fidgeting/ Grooming | Eating/ Drinking | Radio/ Reaching Object | Talking To Passenger | Receiving Calls | Drowsy | Non-Distraction | Total Distraction Rate (%) |
|---|---|---|---|---|---|---|---|---|---|
| ***Type of Median*** | | | | | | | | | |
| **Unprotected Median** | 7.3 | 5.5 | 3.0 | 3.3 | 1.5 | 0.9 | 0.2 | 78.3 | 21.7 |
| **Curbed Median** | 9.2 | 3.8 | 5.1 | 3.9 | 0.9 | 2.3 | 0.2 | 74.7 | 25.3 |
| **Positive Median** | 6.6 | 5.2 | 3.9 | 3.7 | 1.4 | 1.3 | 0.3 | 77.5 | 22.5 |
| ***No. of Lanes*** | | | | | | | | | |
| **2** | 8.2 | 5.6 | 3.1 | 3.3 | 1.2 | 0.9 | 0.2 | 77.4 | 22.6 |
| **3** | 6.8 | 5.5 | 4.0 | 4.1 | 1.5 | 1.0 | 0.4 | 76.7 | 23.3 |
| **4 or more** | 7.1 | 6.2 | 2.6 | 2.3 | 2.0 | 0.9 | 0.2 | 78.7 | 21.3 |
| ***Posted Speed Limit (mph)*** | | | | | | | | | |
| **25-35** | 3.7 | 10.1 | 3.1 | 5.4 | 3.1 | 0.5 | 0.3 | 73.9 | 26.1 |
| **36-45** | 7.2 | 5.5 | 3.4 | 3.4 | 1.4 | 1.2 | 0.7 | 77.1 | 22.9 |
| **46-55** | 6.2 | 5.0 | 3.3 | 2.9 | 1.1 | 1.6 | 0.4 | 79.3 | 20.7 |
| **56-65** | 6.6 | 4.7 | 3.9 | 4.1 | 1.4 | 1.3 | 0.3 | 77.8 | 22.2 |
| ***Median Width (ft.)*** | | | | | | | | | |
| **0-10** | 7.0 | 5.6 | 4.0 | 3.2 | 1.4 | 1.2 | 0.4 | 77.1 | 22.9 |
| **11-20** | 7.1 | 5.4 | 3.2 | 3.2 | 1.1 | 1.4 | 0.3 | 78.3 | 21.7 |
| **21-30** | 6.0 | 4.6 | 3.4 | 3.3 | 1.5 | 1.1 | 0.3 | 79.8 | 20.2 |
| **30 or more** | 6.3 | 5.4 | 3.6 | 3.2 | 1.6 | 1.2 | 0.4 | 78.4 | 21.6 |
| ***Shoulder Width (ft.)*** | | | | | | | | | |
| **0-3** | 7.9 | 6.7 | 4.0 | 3.5 | 1.1 | 1.1 | 0.5 | 81.3 | 18.7 |
| **4-6** | 6.5 | 4.3 | 3.9 | 2.0 | 1.1 | 0.6 | 0.3 | 81.3 | 18.7 |
| **7-9** | 7.0 | 4.3 | 3.1 | 3.0 | 0.9 | 1.5 | 0.5 | 79.8 | 20.2 |
| **9 or more** | 6.6 | 5.4 | 3.6 | 3.2 | 1.5 | 1.2 | 0.3 | 78.1 | 21.9 |

### *Interpretation of the Mann–Whitney U Test*

A Mann–Whitney U test gives the ordered rank of the variation of two data samples. These pairs were the seasonal variation (spring vs. summer), the variation of the day of the week (weekday vs. weekend), the roadway variation (toll road vs. non-toll road), and the roadway classification (signalized vs. unsignalized). The results for weekdays and weekends demonstrate that the former had a slightly higher distraction rate (22.9%) compared to the latter. However, regarding the overall percentage of distractions, the category “handheld cellphone” was the leading type of distraction during both weekdays and weekends (6.8; 7.1, respectively). The percentage attributed to “fidgeting/grooming” increased during the weekdays (5.5%) compared to the weekends (5.3%).

It was also noted that the distraction categories of “radio/reaching objects,” and “receiving calls” were found to be statistically significant, indicating that the rate of distraction due to these driver behaviors were significantly different during weekdays and weekends (Table 8).

**Table 8** *Mann-Whitney U Test for the Comparison of Weekday and Weekend*

| Driver Behavior | Mean Rank Weekdays | Mean Rank Weekends | Delta Mean Rank | Mann-Whitney U | Z-score | P-Value | Effect Size |
|---|---|---|---|---|---|---|---|
| **Fidgeting/Grooming** | 40.28 | 40.87 | 0.59 | 739.0 | -0.10 | 0.92 | 0.01 |
| **Radio/Reaching Objects** | 44.46 | 33.9 | -10.56 | 552.0 | 1.96 | 0.05* | 0.22 |
| **Drowsy** | 42.87 | 36.55 | -6.32 | 631.5 | 1.17 | 0.24 | 0.13 |
| **Talking to Passenger** | 40.17 | 41.05 | 0.88 | 733.5 | -0.16 | 0.87 | 0.02 |
| **Receiving Calls** | 45.44 | 32.27 | -13.17 | 503.0 | 2.45 | 0.01* | 0.27 |
| **Eating/Drinking** | 41.78 | 38.37 | -3.41 | 686.0 | 0.63 | 0.53 | 0.07 |
| **Handheld Cell Phone** | 41.22 | 39.3 | -1.92 | 714.0 | 0.35 | 0.73 | 0.04 |
| **Non-Distracted** | 37.22 | 45.97 | 8.75 | 586.0 | -1.62 | 0.11 | 0.18 |
| *** statistically significant ($p < 0.05$)** | | | | | | | |

The delta mean rank also illustrates that weekdays exhibit more “receiving calls” and “reaching object” events than the weekends. Prat et al. (2015) found similar trends for “receiving calls” and explained that one possible reason was that work-related calls or calls to keep in touch with the family increased on weekdays. As per Cohen’s classification, the effect size of “receiving calls” also had a small significant impact. In addition to “receiving calls” category, “handheld cellphone,” “drowsy driving,” “radio/reading object,” “fidgeting/grooming,” and “eating/drinking” events also increased on weekdays. This finding is consistent with the results of previous studies by Prat et al. (2015) and Johnson et al. (2004), where the researchers found that cellphone use increases on weekdays. However, “talking to passenger” increases during weekends compared to weekdays. Perhaps the presence of more passengers on weekend trips plays a role here. The mean rank of “non-distraction” significantly increased during the weekends, which is obvious since the sources of distraction like cellphone use and receiving calls are less during the weekends.

**Table 9** *Mann-Whitney U Test for the Comparison of Signalized vs. Unsignalized Road*

| Driver Behavior | Mean Rank Signalized Road | Mean Rank Unsignalized Road | Delta Mean Rank | Mann-Whitney U | Z-score | P-Value | Effect Size |
|---|---|---|---|---|---|---|---|
| **Fidgeting/Grooming** | 45.18 | 35.83 | -9.35 | 613.0 | 1.79 | 0.07 | 0.20 |
| **Radio/ Reaching Objects** | 41.28 | 39.72 | -1.56 | 769.0 | 0.29 | 0.77 | 0.03 |
| **Drowsy** | 43.19 | 37.81 | -5.38 | 692.5 | 1.03 | 0.30 | 0.12 |
| **Talking to Passenger** | 37.08 | 43.92 | 6.84 | 663.0 | -1.31 | 0.19 | 0.15 |
| **Receiving Calls** | 35.45 | 45.55 | 10.1 | 598.0 | -1.93 | 0.05* | 0.22 |
| **Eating/Drinking** | 45.75 | 35.25 | -10.5 | 590.0 | 2.02 | 0.04* | 0.23 |
| **Handheld Cell Phone** | 44.12 | 36.88 | -7.24 | 655.0 | 1.39 | 0.16 | 0.16 |
| **Non-Distracted** | 35.98 | 45.02 | 9.04 | 619.0 | -1.74 | 0.08 | 0.19 |
| *** statistically significant ($p < 0.05$)** | | | | | | | |

Unsignalized roads recorded a lower distraction rate (22%) than signalized roads (23.3%), with “handheld cellphone” being the major distraction for both (6.7%; 7.2%, respectively). Signalized roads experienced more distractions from grooming and “eating/drinking” than the unsignalized roads, while the latter had a greater proportion of “talking to passengers” events. Both the “eating/drinking” and “receiving calls” events were statistically significant for both types of roads (Table 9). One explanation is that signalized intersections promote behaviors like tuning radios, reaching for objects, texting, fidgeting, and eating/drinking when the driver has to reduce their speed or stop. However, both “talking to passenger”, and “receiving calls” increase in rank on unsignalized roads compared to the signalized roads. Interestingly, both of these distractions require verbal actions. Hence, the drivers are more comfortable performing verbal actions on the road without stopping or reducing speed.

In terms of the seasonal factor, the overall distraction rate during the summer season increased by 3.6% compared to the spring, particularly the distraction rates associated with “grooming” and “eating/drinking.” All the classes, including “grooming,” “eating/drinking,” and “non-distracted,” were also found to have moderate significance in terms of the effect size (Table 10). However, “non-distracted,” “eating/drinking,” and “grooming” were all more statistically significant during spring compared to summer. These findings are further explained by the fact that non-distraction had a significant drop in its delta mean rank value for the summer compared to the spring. Interestingly, most of the distractions increased their mean rank during the summer season. Similar findings were highlighted by Qin et al. (2019), demonstrating that summer has a greater distraction rate than the other seasons.

**Table 10** *Mann-Whitney U Test for the Comparison of Spring vs. Summer Seasons*

| Driver Behavior | Mean Rank Spring | Mean Rank Summer | Delta Mean Rank | Mann-Whitney U | Z-score | P-Value | Effect Size |
|---|---|---|---|---|---|---|---|
| **Fidgeting/Grooming** | 28.83 | 47.14 | 18.31 | 401.0 | 3.38 | 0.00* | 0.38 |
| **Radio/Reaching Objects** | 45.55 | 37.63 | -7.92 | 593.0 | -1.46 | 0.14 | 0.16 |
| **Drowsy** | 36.81 | 42.60 | 5.79 | 632.5 | 1.07 | 0.28 | 0.12 |
| **Talking to Passenger** | 41.07 | 40.18 | -0.89 | 723.0 | -0.16 | 0.87 | 0.02 |
| **Receiving Calls** | 38.17 | 41.82 | 3.65 | 672.0 | 0.67 | 0.50 | 0.07 |
| **Eating/Drinking** | 27.38 | 47.96 | 20.58 | 359.0 | 3.80 | 0.00* | 0.42 |
| **Handheld Cell Phone** | 41 | 40.22 | -0.78 | 725.0 | -0.14 | 0.89 | 0.02 |
| **Non-Distracted** | 50.86 | 34.61 | -16.25 | 439.0 | -3.00 | 0.00* | 0.34 |
| *** statistically significant (p < 0.05)** | | | | | | | |

In respect of the toll and non-toll roads, almost similar overall rates of distraction events (22.6%, 22.7%, respectively) were recorded for both categories. It should be noted that “handheld cellphone” was the primary type of distraction on toll (7.3%) and non-toll roads (6.8%). “Receiving calls” was found to be statistically significant in both types of roads (Table 11). Based on the delta mean rank, receiving calls increased significantly on toll roads compared to non-toll roads.

**Table 11** *Mann-Whitney U Test for the Comparison of Toll Road vs. Non-Toll Road*

| Driver Behavior | Mean Rank toll road | Mean Rank non-toll road | Delta Mean Rank | Mann-Whitney U | Z-score | P-Value | Effect Size |
|---|---|---|---|---|---|---|---|
| **Fidgeting/Grooming** | 36.56 | 41.48 | 4.92 | 449.0 | 0.75 | 0.45 | 0.08 |
| **Radio/Reaching Objects** | 39.78 | 40.68 | 0.9 | 500.5 | 0.13 | 0.90 | 0.01 |
| **Drowsy** | 47.34 | 38.79 | -8.55 | 402.5 | -1.31 | 0.19 | 0.15 |
| **Talking to Passenger** | 43.5 | 39.75 | -3.75 | 464.0 | -0.57 | 0.57 | 0.06 |
| **Receiving Calls** | 50.62 | 37.97 | -12.65 | 350.0 | -1.94 | 0.05* | 0.22 |
| **Eating/Drinking** | 33.12 | 42.34 | 9.22 | 394.0 | 1.41 | 0.16 | 0.16 |
| **Handheld Cell Phone** | 42.06 | 40.11 | -1.95 | 487.0 | -0.29 | 0.77 | 0.03 |
| **Non-Distracted** | 42.06 | 40.11 | -1.95 | 487.0 | -0.29 | 0.77 | 0.03 |
| *** statistically significant (p < 0.05)** | | | | | | | |

The results of the peak hour (9 a.m. to 12 p.m. and 3 p.m. to 6 p.m.) and off-peak hours (12 p.m. to 3 p.m.) comparison demonstrate that the overall rate of distraction was greater during peak hours (24.6%) than during off-peak hours (22 %). Both categories saw “handheld cellphone” (7.7% during peak hours, and 6.6% during off-peak hours) as the major type of distraction. However, peak hours experienced considerably greater rates of “fidgeting/grooming” and “eating/drinking” distractions compared to the off-peak hours. “Non-distraction” was found with statistically significant p-values and a small significance value of the effect size (Table 12). From the delta mean rank, “non-distraction” significantly increased during off-peak hours compared to peak hours.

**Table 12** *Mann-Whitney U Test for the Comparison of Peak Hour vs. Off-Peak Hour*

| Driver Behavior | Mean Rank peak hour | Mean Rank off-peak hour | Delta Mean Rank | Mann-Whitney U | Z-score | P-Value | Effect Size |
|---|---|---|---|---|---|---|---|
| **Fidgeting/Grooming** | 54.34 | 47.97 | -6.37 | 1111.5 | -1.09 | 0.28 | 0.04 |
| **Radio/Reaching Objects** | 49.86 | 52.03 | 2.17 | 1217.5 | 0.37 | 0.71 | 0.09 |
| **Drowsy** | 49.35 | 52.49 | 3.14 | 1193.0 | 0.53 | 0.60 | 0.08 |
| **Talking to Passenger** | 54.44 | 47.89 | -6.55 | 1107.0 | -1.12 | 0.26 | 0.03 |
| **Receiving Calls** | 55.73 | 46.72 | -9.01 | 1045.0 | -1.54 | 0.12 | 0.02 |
| **Eating/Drinking** | 52.68 | 49.48 | -3.20 | 1191.5 | -0.54 | 0.59 | 0.08 |
| **Handheld Cell Phone** | 54.43 | 47.90 | -6.53 | 1107.5 | -1.12 | 0.26 | 0.03 |
| **Non-Distracted** | 44.86 | 56.56 | 11.70 | 977.5 | 2.00 | 0.05* | 0.01 |
| *** statistically significant (p < 0.05)** | | | | | | | |

### *Interpretation of the Kruskal–Wallis Test*

The Kruskal–Wallis test was performed on the various ranges of the speed limit to determine whether there is a significant change in the mean rank of the distraction rates for the various classes due to changes in speed limit. As shown in Table 13, all types of distraction are statistically significant (p-value < 0.05, H value > 7.81) due to speed variations. All types of distractions showed significance with a large effect size except the category of drowsy driving which has a medium effect size.

**Table 13** *Effect of the Variation of Speed on the Drivers' Behaviors*

| Driver Behavior | H score | P-Value | Effect Size |
|---|---|---|---|
| **Handheld Cell Phone** | 68.73* | < 0.0001* | large, 0.21 |
| **Fidgeting/Grooming** | 47.68* | < 0.0001* | large, 0.14 |
| **Radio/Reaching Objects** | 51.44* | < 0.0001* | large, 0.15 |
| **Eating/Drinking** | 66.27* | 0.0001* | large, 0.20 |
| **Talking to Passenger** | 58.48* | < 0.0001* | large, 0.18 |
| **Receiving Calls** | 71.96* | < 0.0001* | large, 0.22 |
| **Drowsy** | 34.20* | < 0.0001* | medium, 0.10 |
| **Non-Distracted** | 69.79* | < 0.0001* | large, 0.21 |
| *** statistically significant** | | | |
| **(Degree of freedom=3, α= 0.05, Critical $\chi^2$ value = 7.8147)** | | | |

Table 14 demonstrates the direction of the significance (increase/decrease) of the rates of distractions associated with changes in the speed limit. For instance, the "handheld cellphone" class shows a significant increase ($p < 0.05$, $H > 7.8147$) due to changes in the speed limit from 25-35 mph to higher speed limits. Previous studies also found that the drivers select a higher speed while using a mobile phone (Oviedo-Traspalacios et al., 2017; Fitch et al., 2014). It should be noted that reducing the speed limit from (56–65 mph) to (25–35 mph) is associated with a significant reduction for all types of distraction. These findings demonstrate that the speed limit has a significant impact on driver distraction behaviors.

**Table *14*** *Kruskal Wallis Test for the Pairwise Comparison of Speed Limits*

| Type of Distraction | Mean Rank Values | | | | Direction of Significance (↑ for increase, ↓ for decrease) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Posted Speed Limit (mph) | | | | | | | | | |
| | 25-35 | 36-45 | 46-55 | 56-65 | 25-35 vs. 36-45 | 25-35 vs. 46-55 | 25-35 vs. 56-65 | 36-45 vs. 46-55 | 36-45 vs. 56-65 | 46-55 vs. 56-65 |
| **Handheld Cell Phone** | 100 | 157.2 | 214.2 | 170.7 | ↑ | ↑ | ↑ | - | ↑ | ↓ |
| **Fidgeting/Grooming** | 117.5 | 148.3 | 211.3 | 164.9 | ↑ | ↑ | ↑ | ↑ | - | ↓ |
| **Radio/Reaching Objects** | 123.9 | 145.0 | 196.3 | 176.9 | ↑ | ↑ | ↑ | ↑ | ↑ | - |
| **Eating/Drinking** | 104.1 | 146.3 | 210.4 | 181.2 | ↑ | ↑ | ↑ | ↑ | ↑ | - |
| **Talking to Passenger** | 109.1 | 148.3 | 202.8 | 181.8 | ↑ | ↑ | ↑ | ↑ | ↑ | - |
| **Receiving Calls** | 104.7 | 146.1 | 204.8 | 186.5 | ↑ | ↑ | ↑ | ↑ | ↑ | - |
| **Drowsy** | 128.7 | 150.9 | 188.1 | 174.3 | ↑ | ↑ | ↑ | ↑ | - | - |
| **Non-Distracted** | 114.1 | 136.8 | 226.2 | 165.0 | ↑ | ↑ | ↑ | ↑ | ↑ | ↓ |

The test results from the Kruskal–Wallis test demonstrate that the variation of the number of lanes is significant for all types of distraction. However, only the distraction classes “radio/reaching object” and “talking to passenger” show a demonstrated significance with a large effect size (Table 15). The H-score for all classes of distraction is greater than the critical chi-square value, which indicates that the difference between the mean ranks of some groups is big enough to be statistically significant.

Table 16 demonstrates the pairwise comparison of different roadway numbers of lanes on the drivers’ behaviors. According to this table, several distraction types are statistically significant ($p < 0.05$, $H > 5.9915$) concerning the number of lanes. For instance, the “drowsy driving” class shows a significant increase with the changes in the number of lanes from “2 lanes” to “3 lanes,” while it shows a significant decrease with a change in the number of lanes from “2 lanes” to “4 or more lanes” and from “3 lanes” to “4 or more lanes.” As the likelihood of fatal distracted driving crashes is increased by increasing the roadway number of lanes, the drivers are more cautious and less distracted on roadways with more lanes (Chen & Lym, 2021; Stavrinos et al., 2013). In addition, changes in the number of lanes from “2 lanes” to “4 or more lanes” and from “3 lanes” to “4 or more lanes” decreased all types of distraction except drowsy. These findings demonstrate that driver distraction behavior is significantly influenced by the number of lanes.

**Table 15** *Effect of the Variation of Number of Lanes on the Drivers’ Behaviors*

| Driver Behavior | H score | P-Value | Effect Size |
|---|---|---|---|
| **Fidgeting/Grooming** | 19.03* | < 0.00001* | medium, 0.07 |
| **Radio/Reaching Objects** | 32.20* | < 0.00001* | large, 0.13 |
| **Drowsy** | 20.48* | < 0.00001* | medium, 0.08 |
| **Talking to Passenger** | 26.35* | < 0.00001* | large, 0.10 |
| **Receiving Calls** | 18.89* | < 0.00001* | medium, 0.07 |
| **Eating/Drinking** | 24.56* | < 0.00001* | medium, 0.09 |
| **Handheld Cell Phone** | 24.68* | < 0.00001* | medium, 0.10 |
| **Non-Distracted** | 9.43* | 0.00895* | small, 0.03 |
| *** statistically significant** | | | |
| **(Degree of freedom=2, α= 0.05, Critical $\chi^2$ value = 5.9915** | | | |

**Table 16** *Kruskal Wallis Test for the Pairwise Comparison of the Number of Lanes*

| Type of Distraction | Mean Rank Values | | | Direction of Significance (↑ for increase, ↓ for decrease) | | |
|---|---|---|---|---|---|---|
| | Number of Lanes | | | | | |
| | 2 | 3 | 4 or more | 2 vs. 3 | 2 vs. 4 or more | 3 vs. 4 or more |
| **Handheld Cell Phone** | 142.8 | 128.3 | 90.4 | - | ↓ | ↓ |
| **Fidgeting/Grooming** | 133.7 | 134.7 | 93.1 | - | ↓ | ↓ |
| **Radio/Reaching Objects** | 131.0 | 144.5 | 86.0 | - | ↓ | ↓ |
| **Eating/Drinking** | 131.5 | 140.2 | 89.9 | - | ↓ | ↓ |
| **Talking to Passenger** | 126.7 | 143.9 | 90.9 | - | ↓ | ↓ |
| **Receiving Calls** | 125.5 | 140.0 | 96.0 | - | ↓ | ↓ |
| **Drowsy** | 119.3 | 141.8 | 100.5 | ↑ | ↓ | ↓ |
| **Non-Distracted** | 139.8 | 109.0 | 112.6 | ↓ | ↓ | - |

From Table 17, we can see that the effect of variations in the median type is significant for all types of distraction ($p < 0.05$, $H > 5.9915$). Except for the non-distraction class, all other classes demonstrated significance with a large effect size.

**Table 17** *Effect of the Variation of Median Type on the Drivers' Behaviors*

| Driver Behavior | H score | P-Value | Effect Size |
|---|---|---|---|
| **Fidgeting/Grooming** | 38.05* | < 0.00001* | large, 0.15 |
| **Radio/Reaching Objects** | 34.90* | < 0.00001* | large, 0.14 |
| **Drowsy** | 44.16* | < 0.00001* | large, 0.18 |
| **Talking to Passenger** | 41.37* | < 0.00001* | large, 0.17 |
| **Receiving Calls** | 57.67* | < 0.00001* | large, 0.23 |
| **Eating/Drinking** | 41.04* | < 0.00001* | large, 0.16 |
| **Handheld Cell Phone** | 57.67* | < 0.00001* | large, 0.23 |
| **Non-Distracted** | 24.12* | < 0.00001* | medium, 0.09 |
| *** statistically significant** | | | |
| **(Degree of freedom=2, α= 0.05, Critical $\chi^2$ value = 5.9915** | | | |

According to Table 18, the "handheld cellphone" distractions were significantly decreased ($p < 0.05$, $H > 5.9915$) due to changes in the median type from "positive" to "curbed" and from "unprotected" to "curbed." The crash analysis presented in Chapter 3 also found that the type of median has a significant effect on crash injury level involving cellphone use. Similarly, "fidgeting/grooming," "talking to passenger," and "non-distraction" were also significantly affected by median type. It should be noted that a change in median type from "positive" to "curbed" was found to reduce all types of distractions significantly. These findings indicate that variations in median type have significant impacts on driver distraction behaviors. Chen & Lym (2021)

showed that the presence of medians reduces the frequency of crashes involving distractions.

**Table 18** *Kruskal Wallis Test for the Pairwise Comparison of Median Type*

| Type of Distraction | Mean Rank Values | | | Direction of Significance (↑ for increase, ↓ for decrease) | | |
|---|---|---|---|---|---|---|
| | Median Type | | | | | |
| | Unprotected | Positive | Curbed | Unprotected vs Positive | Unprotected vs Curbed | Positive vs Curbed |
| **Handheld Cell Phone** | 115.9 | 161.1 | 84.4 | ↑ | ↓ | ↓ |
| **Fidgeting/Grooming** | 127.3 | 149.8 | 84.4 | ↑ | ↓ | ↓ |
| **Radio/Reaching Objects** | 117.8 | 153.4 | 90.3 | ↑ | ↓ | ↓ |
| **Eating/Drinking** | 117.1 | 156.6 | 87.8 | ↑ | ↓ | ↓ |
| **Talking to Passenger** | 123.5 | 152.2 | 85.8 | ↑ | ↓ | ↓ |
| **Receiving Calls** | 115.9 | 161.1 | 84.4 | ↑ | ↓ | ↓ |
| **Drowsy** | 110.6 | 153.6 | 97.3 | ↑ | - | ↓ |
| **Non-Distracted** | 127.1 | 143.4 | 91.0 | ↑ | ↓ | ↓ |

The Kruskal–Wallis test results are shown in Table 19 illustrates that all types of distraction are statistically significant (p-value < 0.05, H value > 7.81) due to the variation in median width. Only “talking to passenger” distraction category showed significant change with a large effect size, all other types of distractions showed significant changes with medium or small effect size. Non-distracted driving behavior did not record a significant change by the means of median width variation.

**Table 19** *Effect of the Variation of Median Width on the Drivers’ Behaviors*

| Driver Behavior | H score | P-Value | Effect Size |
|---|---|---|---|
| **Handheld Cell Phone** | 13.79* | 0.0032* | small, 0.04 |
| **Fidgeting/Grooming** | 21.20* | < 0.0001* | small, 0.05 |
| **Radio/Reaching Objects** | 20.77* | < 0.0001* | medium, 0.06 |
| **Eating/Drinking** | 16.04* | 0.0011* | medium, 0.04 |
| **Talking to Passenger** | 9.31* | 0.0254* | large, 0.32 |
| **Receiving Calls** | 13.86* | 0.0031* | small, 0.04 |
| **Drowsy** | 34.21* | < 0.0001* | medium, 0.1 |
| **Non-Distracted** | 3.76 | 0.2886 | small, 0.002 |
| *** statistically significant** | | | |
| **(Degree of freedom=3, α= 0.05, Critical $\chi^2$ value = 7.8147)** | | | |

Table 20 demonstrates the direction of the significance (increase/decrease) of the rates of distractions associated with changes in the median width. For instance, the “handheld cellphone” class shows a significant reduction ($p < 0.05$, $H > 7.8147$) due to changes in the median width to 21'-30' from lower median widths. At the same time, it increases significantly when the median width increases further (>30'). Similar trends are observed for distractions like “Fidgeting/Grooming” and “Eating/Drinking.” A recent study by Kong et al. (2021) demonstrated that the median width is associated with distracted driving crashes. It should be noted that changing the median width from 21'-30' to higher or lower ranges of median width significantly increases the ranks for distraction. These findings demonstrate that the median width has a significant impact on driver distraction behaviors.

**Table 20** *Kruskal Wallis Test for the Pairwise Comparison of Median Width*

| Type of Distraction | Mean Rank Values | | | | Direction of Significance (↑ for increase, ↓ for decrease) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Median Width (ft) | | | | | | | | | |
| | 0-10 | 11-20 | 21-30 | >30 | 0-10 vs. 11-20 | 0-10 vs. 21-30 | 0-10 vs. >30 | 11-20 vs. 21-30 | 11-20 vs. >30 | 21-30 vs. >30 |
| **Handheld Cell Phone** | 178.0 | 161.0 | 128.1 | 166.9 | - | ↓ | - | ↓ | - | ↑ |
| **Fidgeting/Grooming** | 178.8 | 172.2 | 119.8 | 163.2 | - | ↓ | - | ↓ | - | ↑ |
| **Radio/Reaching Objects** | 178.0 | 161.1 | 133.8 | 161.1 | - | ↓ | - | ↓ | - | - |
| **Eating/Drinking** | 183.8 | 159.8 | 127.9 | 162.5 | - | ↓ | - | ↓ | - | ↑ |
| **Talking to Passenger** | 181.6 | 153.2 | 142.2 | 157.0 | ↓ | ↓ | - | - | - | - |
| **Receiving Calls** | 183.6 | 155.4 | 135.3 | 159.6 | - | - | ↓ | - | - | - |
| **Drowsy** | 186.5 | 160.2 | 138.4 | 149.0 | ↓ | ↓ | ↓ | ↓ | - | - |
| **Non-Distracted** | 173.2 | 154.4 | 146.2 | 160.2 | - | - | - | - | - | - |

The Kruskal–Wallis test was performed on the various ranges of the shoulder width to determine whether there is a significant change in the mean rank of the distraction rates for the various classes due to changes in shoulder width. As shown in Table 21, all types of distraction are found to be statistically significant (p-value < 0.05, H value > 7.81) due to shoulder width changes.

**Table 21** *Effect of the Variation of Shoulder Width on the Drivers' Behaviors*

| Driver Behavior | H score | P-Value | Effect Size |
|---|---|---|---|
| **Fidgeting/Grooming** | 38.43* | < 0.00001* | large, 0.11 |
| **Radio/Reaching Objects** | 46.44* | < 0.00001* | large, 0.14 |
| **Drowsy** | 57.67* | < 0.00001* | large, 0.18 |
| **Talking to Passenger** | 102.74* | < 0.00001* | large, 0.32 |
| **Receiving Calls** | 71.54* | < 0.00001* | large, 0.22 |
| **Eating/Drinking** | 47.36* | < 0.00001* | large, 0.14 |
| **Handheld Cell Phone** | 28.59* | 0.000002* | medium, 0.082 |
| **Non-Distracted** | 17.14* | 0.0007* | medium, 0.045 |
| *** statistically significant** | | | |
| **(Degree of freedom=3, α= 0.05, Critical $\chi^2$ value = 7.8147** | | | |

Table 22 shows the direction of the significance (increase/decrease) of the rates of distractions associated with changes in the shoulder width. For instance, the "handheld cellphone" class shows a significant increase ($p < 0.05$, $H > 7.8147$) due to changes in the shoulder width to >9' from lower shoulder width. Moreover, all the distraction classes significantly increase the distraction rate when the shoulder width is increased to >9' from lower shoulder widths. Previous studies also found that an increase in shoulder width encourages distraction, especially using a mobile phone (Kong et al., 2021). It should be noted that increasing the shoulder width from 0-3' to 4'-6,' and 6'-9' decreases the distraction events. These findings demonstrate that the shoulder width has a significant impact on driver distraction behaviors.

**Table 22** *Kruskal Wallis Test for the Pairwise Comparison of Shoulder Width*

| Type of Distraction | Mean Rank Values | | | | Direction of Significance (↑ for increase, ↓ for decrease) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Shoulder Width (ft) | | | | | | | | | |
| | 0-3 | 4-6 | 7-9 | >9 | 0-3 vs. 4-6 | 0-3 vs. 7-9 | 0-3 vs. >9 | 4-6 vs. 7-9 | 4-6 vs. >9 | 7-9 vs. >9 |
| **Handheld Cell Phone** | 154.3 | 139.0 | 138.2 | 202.6 | - | - | ↑ | - | ↑ | ↑ |
| **Fidgeting/Grooming** | 155.7 | 138.5 | 131.7 | 208.1 | - | - | ↑ | - | ↑ | ↑ |
| **Radio/Reaching Objects** | 157.1 | 128.8 | 136.3 | 211.8 | ↓ | - | ↑ | - | ↑ | ↑ |
| **Eating/Drinking** | 156.1 | 129.9 | 135.2 | 212.8 | ↓ | - | ↑ | - | ↑ | ↑ |
| **Talking to Passenger** | 149.0 | 127.9 | 123.1 | 234.0 | ↓ | ↓ | ↑ | - | ↑ | ↑ |
| **Receiving Calls** | 155.5 | 124.5 | 135.2 | 218.8 | ↓ | - | ↑ | - | ↑ | ↑ |
| **Drowsy** | 155.9 | 134.8 | 138.7 | 204.5 | ↓ | - | ↑ | - | ↑ | ↑ |
| **Non-Distracted** | 155.3 | 147.1 | 138.5 | 193.1 | - | - | ↑ | - | ↑ | ↑ |

## Conclusion

This chapter investigated distracted driving events in New Jersey using the floating car method and identified the effects of temporal characteristics and roadway types on the rate of distraction events. The Mann–Whitney U test was used to confirm the statistically significant changes in the rate of distraction events as a function of the day of the week, the season, and the roadway type. The results showed that "receiving calls" events statistically significantly decreased in unsignalized roads compared to signalized roads, during weekdays compared to the weekend, and on toll roads compared to non-toll roads. Similarly, the "food/drink" events were significantly increased during summer compared to the spring and on signalized roads compared to the unsignalized corridors. It was also found that summer had more distraction events than spring. The Kruskal–Wallis test was then employed to investigate the significance of the variation of geometric features in roadways (speed limit, median width, number of lanes, shoulder width, and median type) on the types of distractions. An increase in speed limit significantly increased most of the distractions. "Curbed" median encountered significant reduction in distractions compared to "unprotected" and "positive" median. An increase in median width significantly decreased distraction events while an increase in shoulder width significantly increased distraction events. It also should be noted that an increase in the number of lanes significantly decreased the distraction events.

These results indicate a complex interaction between various classes of distractions with temporal and spatial variations. The outcomes from the statistical analysis indicate that safety countermeasures such as law enforcement and awareness campaigns should be focused more on the summer months. Moreover, such countermeasures should also prioritize corridors with signalized intersections and weekday driving. Strict enforcement should also be ensured on corridors with

lower speed limits, wider shoulder width, higher median width, and positive-type medians.

In general, distracted driving events can be addressed by adopting the “three Es” strategies: Engineering, Enforcement, and Education (Qi et al., 2020). According to the results obtained here, the distracted driving awareness month in April has shown a positive response toward mitigating distracting driving rates. To reduce “cellphones” and “receiving call” events, various insurance companies and mobile carriers have been offering incentives, and these could also play a vital role in reducing distractions.

# CHAPTER 5: Detection of Distracted Driving Using Artificial Intelligence (AI)

## Introduction

Distracted driving is a major traffic safety concern in the United States. The existing high visibility enforcement campaigns against distracted driving focus on monitoring driver behaviors (Elqattan et al., 2019). Hence, various techniques to monitor or detect distracted behaviors have been introduced into the transportation safety community. As we saw in Chapter 2, previous research in detecting distracted driving includes methods that rely on computer vision and deep learning algorithms (Elqattan et al., 2019; De Castro et al., 2018).

Although an overwhelming majority of the research into the detection of distracted driving was found to have used dashcam videos, images of driver behaviors taken from these devices have some limitations. A dash camera (dashcam) can capture the behavior of only one driver at a single time, and the behavior of drivers is biased when they are aware of being recorded. Due to these limitations, images captured with dashcams cannot identify or detect actual distracted drivers on a road segment. In order to perform real-time detection of distracted drivers, the videos of the drivers should be taken from cameras located outside the car.

Observing driver behaviors in this way can be conducted using either a cross-sectional or a longitudinal setting. To date, most of the related research performed cross-sectional data collection (Elqattan et al., 2019), where driver behavior is captured from a single point. However, to fully capture the real picture of distracted driving, temporal variations should be employed in order to capture driver behaviors from different roadway geometries (e.g., different speed limits, on signalized or unsignalized roads, different median types, and different numbers of lanes).

With these ideas in mind, this chapter presents a novel longitudinal observational data collection approach to capturing videos of drivers. In this study, a test vehicle equipped with cameras was driven along the selected corridors, where driver motions were captured through their side windows. All the video recordings were done during the daytime with the permission of the state authorities. The driver behaviors captured through this method are unbiased and versatile and can thereby contribute to the research regarding the real-time detection of distracted drivers.

## Methodology

### *Data Collection*

The proposed system collects driver data through the side window of a moving test vehicle equipped with two cameras. The configuration of the cameras from which

the video data were collected is displayed in Figure 17. A set of two GoPro HERO 9 cameras was mounted on top of the vehicle facing left and right. Portable chargers were also attached inside the vehicle to provide the cameras with an uninterrupted power supply. The orientation and alignment of the cameras were optimized to maintain an appropriate angle and distance for capturing driver behaviors from the cars in adjacent lanes while moving along the road.


**Figure 17** *Test Vehicle Equipped with Go Pro Hero 9 Cameras*

### *Model Selection*

Object detection methods for computer vision are divided into two main categories: one-stage and two-stage algorithms. The two-stage algorithms (e.g., R-CNN and Fast R-CNN) first locate the target object before sending the images to the next stage for bounding box regression and object classification. On the other hand, one-stage classifiers (e.g., YOLO and SSD) use convolutional neural networks to simultaneously predict the target category and position (Gao et al., 2021). Compared to one-stage algorithms, two-stage algorithms were initially considered slower but more accurate at detection (Jiang et al., 2020). However, following on from YOLOv3, the more recent algorithms of the YOLO series (i.e., YOLOv4 and YOLOv5) have achieved a much better trade-off in object recognition accuracy and speed (Bochkovskiy et al., 2020). Due to the requirements of faster detection and accuracy, a one-stage YOLOv5 algorithm was used here to detect distracted driving behaviors.

Version 5 of YOLO (You Only Look Once) was developed by Ultralytics in 2020. YOLO is a real-time object detector that has gone through continuous developments over recent years (Xu et al., 2021). YOLOv5 is a fast learner primarily due to three reasons. Firstly, it uses CSPDarknet (a cross partial network incorporated into Darknet) as its backbone, ensuring detection speed and accuracy with reduced model size (Wang et al., 2020). Secondly, YOLOv5 utilizes a path aggregation network (PANet) to boost its information flow and efficiency in finding the target object location (Wang et al., 2019). Thirdly, the YOLO layer or the network head can generate three different feature map sizes (18×18, 36×36, or 72×72), thus

enabling the model to detect small, medium, or large objects (Redmon and Farhadi, 2018). This multi-scale detection method enhances the performance of YOLOv5 when objects of diverse sizes need to be detected. For the detection of distracted driving, this multi-scale detection function is particularly helpful. The architectural network of YOLOv5 is shown in Figure 18.

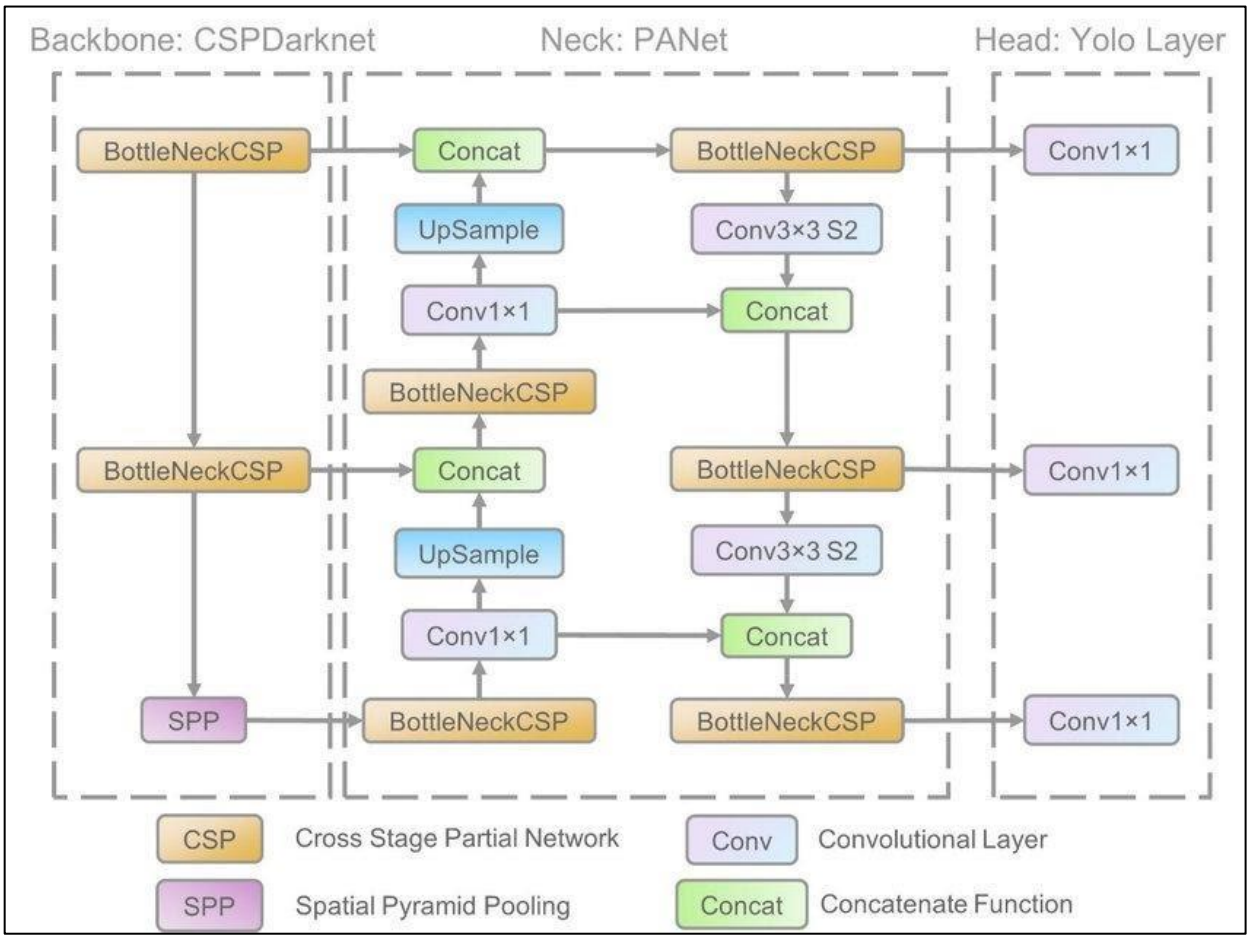



**Figure 18** *YOLOv5 Architecture (Source: Xu et al., 2021)*

For the detection of distracted drivers, images were extracted from the video data. These contained representative images for all of the distraction classes defined in Chapter 4. Following collection, each image was classified into one of these various distraction types. Then, they were split up with a ratio of 9:1 between the training and testing datasets. The training images were first annotated using the LabelImg software, and then the YOLOv5 model was trained on them. This model evaluated its learning performance by predicting the distraction class of 10% of the training images—known as validation. Once validation was completed, the model was ready to test its performance by running some of the testing images through it. Figure 19 describes the stepwise procedure of the overall detection process, with a brief description of all these steps listed as follows:

1. Extraction of Images: The video recordings were reviewed frame-by-frame to extract representative images of the various distraction classes.
2. Labeling Images: Eight folders with the name of each of the distraction classes were created in the working directory of the computer. Then, each of the extracted images was reviewed one by one and moved to the folder with the name of its respective class. For instance, all images categorized as “non-distraction” were moved to the folder named “non-distraction,” all images where the driver was holding a cellphone were moved to “handheld cellphone,” and so on.
3. Splitting Training and Testing Data: A train-to-test ratio of 9:1 was maintained while distributing the images into the two training and testing datasets.
4. Annotation of Images: The training images were annotated according to the YOLOv5 annotation standard using a tool named LabelImg. Images are annotated in order to define representative areas in the defined classes, such as “radio/reaching object” or “talking to passengers.” The activity or body part of the driver was annotated to help the model when learning that driver's behavior during its training.
5. Training the Model: The YOLOv5 algorithm was used to train the model, using different pre-trained model weights combined with customized data.
6. Validation of the Model: Once the model was trained, it was validated using 10% of the training images.
7. Testing the Model: Once the model was validated, it was tested using the testing images.
8. Evaluating Model Performance: After testing, the overall accuracy of the model was evaluated by dividing the total number of correctly classified images by the total number of images tested.

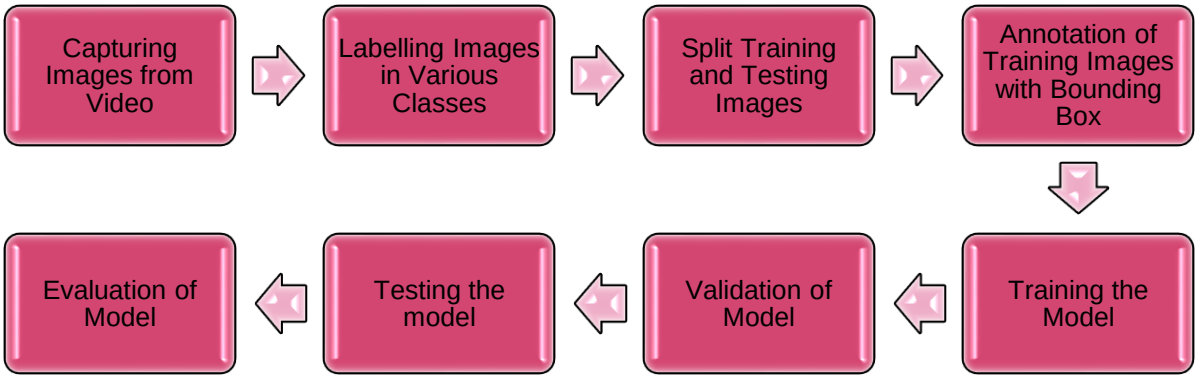



**Figure 19** *The Stepwise Procedure Involved in the Detection of Distracted Driving Behaviors*

### *Definition of Various Classes of Distractions*

The definitions of the different classes of distractions (secondary activities) were discussed in Chapter 4. Representative images of each class were collected and used for detection, as shown in Figure 20.


**Figure 20** *Examples of Various Types of Distraction from the Extracted Frames*

### *Preparation of Datasets*

Detection of driver behaviors needs a dataset containing training images of various distraction types. Most of the publicly available training datasets (e.g., the State Farm dataset or the American University of Cairo dataset) captured their images from inside the car using a dashcam (State Farm, 2016). Hence, they would not be suitable for training this model to detect driver behaviors from cameras located outside the car. A customized dataset was therefore created in this study for model training purposes.

As mentioned in the previous section, a set of training and testing images was organized for the study after extracting them from the frames of the video recordings. It is worth noting that the classes with the fewest images (e.g., drowsy, talking to a

passenger, and radio/reaching object) were oversampled in order to avoid an imbalanced dataset. The LabelImg annotation tool was further utilized to annotate the training images (Tzutalin, 2017). Figure 21 illustrates the annotation process performed using this software tool.

During the annotation process, a rectangular bounding box is added around the image of the driver, and then each of the images is attributed to one single class. Once an image is annotated, a text file is created containing information like the driver attribute label, the width and height of the bounding box, and the centroid of the bounding box. Once all training images are annotated, the model is ready to be trained by them in order to learn the different classes of driver behaviors. It is worth mentioning that YOLOv5 runs the training process on a parallel batch of images (usually 16). Figure 22 and Figure 23 illustrate one such example of training and validation done by YOLOv5 with a batch of 16 images.

### *Camera Selection and System Specifications*

Clear images are essential if the model is to detect the different driver behaviors. However, it is challenging to obtain clear images on the road, given the changing weather conditions, the different lighting conditions during the day, and the high speed of the vehicles. Hence, a preliminary investigation was performed to find a suitable camera that both works well at high speeds and produces images of stable quality. GoPro brand cameras were then found to be the most suitable, with one of their Hero 9 cameras managing to catch the clearest driver images from adjacent lanes. The videos were taken at 60 frames per second and at a resolution of 2704-by-1520 pixels.

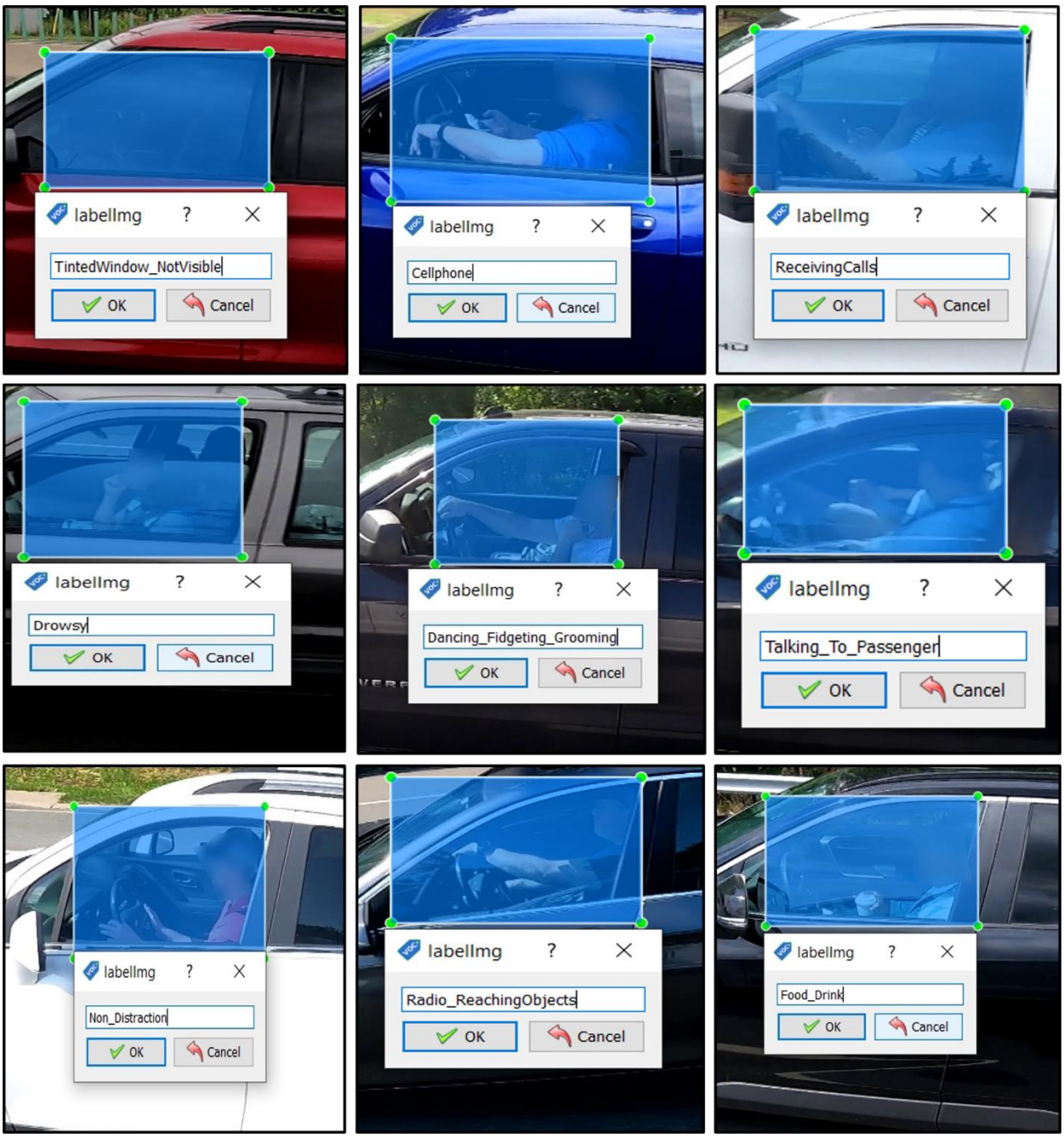



**Figure 21** *Annotation of the Images Using Labeling Tool*


**Figure 22** *Training Images for a Batch of 16 Images*


**Figure 23** *Validation Images for a Batch of 16 Images*

### *Model Performance Evaluation*

The performance of the model was evaluated by using its accuracy when detecting the testing images. The equation to find the accuracy is as follows (Elqattan et al., 2019):

$$\text{Accuracy} = \frac{Total\ number\ of\ correctly\ predicted\ images}{Total\ number\ of\ images} \quad (12)$$

## Results and Discussions

### *Distribution of the Training Dataset*

The model was trained and tested using 5,670 images. Training to the testing ratio of 9:1 was maintained throughout the entire model evaluation process. The distribution of the testing dataset is shown in Figure 24. It was observed that the classes “non-distraction,” “tinted window/not visible,” and “fidgeting/grooming” each had a larger number of training images associated with them than for the other classes. On the other hand, the classes “drowsy,” “reaching object,” and “talking to passenger” had fewer training images than the others.

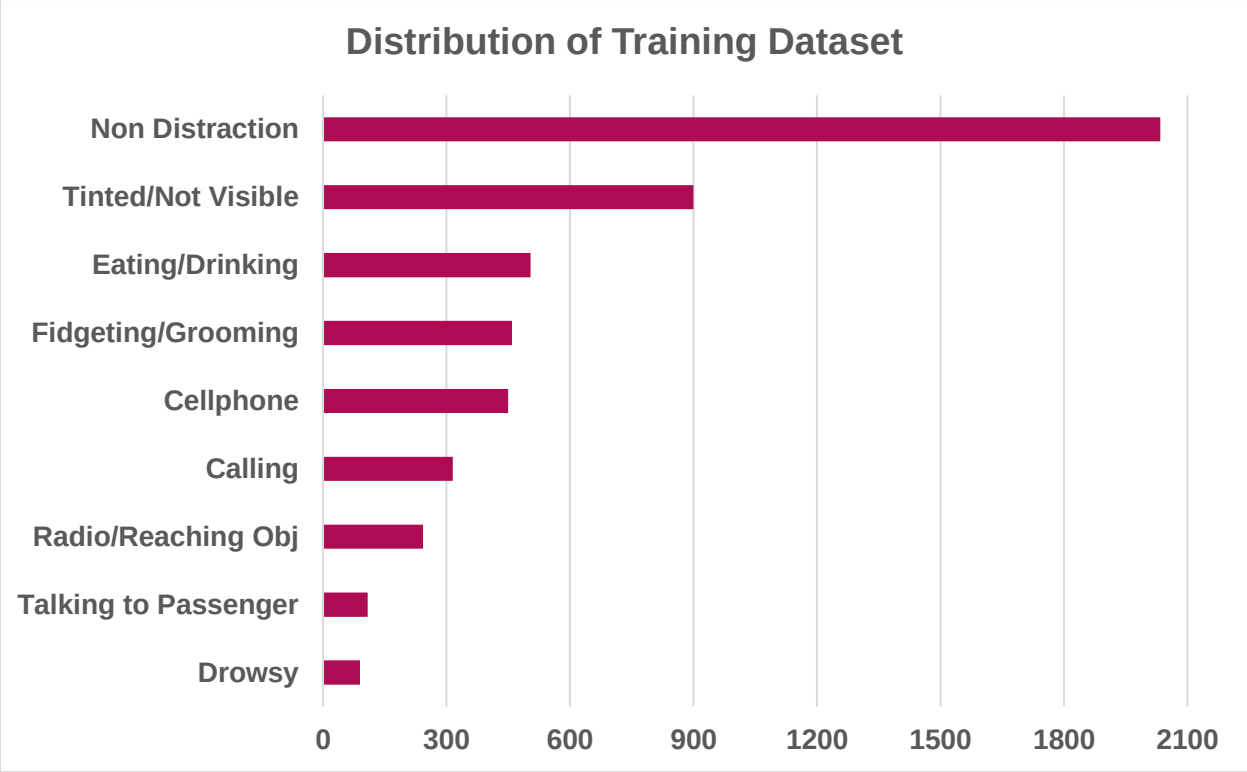



**Figure 24** *Distribution of the Training Dataset*

### *Testing Results*

From the 5,670 images in total, the model was trained on 5,103 images and tested on 567 images. Out of these 567 images, 482 were correctly predicted by the model, which gave an overall testing accuracy of 85.01%. Both “tinted window/not visible” and “non-distraction” classes showed accuracy rates above 90%, and “handheld cellphone” and “receiving calls” had rates of 86% and 78%, respectively. The model predicted 77% of the “eating/drinking” distraction class images correctly, while it had an accuracy of 76% for the “fidgeting/grooming” class. Lastly, the classes “radio/reaching objects,” “talking to passengers” and “drowsy,” were the least accurate of all. A confusion matrix with these results is illustrated in Figure 25.

| Class | Non-Distraction | Cellphone | Tinted/ Not Visible | Fidgeting/ Grooming | Talking to Passenger | Radio/ Reaching Obj | Eating/ Drinking | Drowsy | Receiving Calls |
|---|---|---|---|---|---|---|---|---|---|
| Non-Distraction | **0.91** | 0.10 | 0.10 | 0.16 | 0.33 | 0.30 | 0.23 | 0.20 | 0.11 |
| Cellphone | 0.00 | **0.86** | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Tinted/ Not Visible | 0.05 | 0.00 | **0.90** | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Fidgeting/ Grooming | 0.04 | 0.04 | 0.00 | **0.76** | 0.00 | 0.00 | 0.00 | 0.20 | 0.00 |
| Talking to Passenger | 0.00 | 0.00 | 0.00 | 0.00 | **0.67** | 0.00 | 0.00 | 0.00 | 0.11 |
| Radio/ Reaching Obj | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | **0.70** | 0.00 | 0.00 | 0.00 |
| Eating/ Drinking | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | **0.77** | 0.00 | 0.00 |
| Drowsy | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | **0.60** | 0.00 |
| Receiving Calls | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | **0.78** |

**Figure 25** *Confusion Matrix for the Detection of Distracted Driving*

Figure 26 contains one example of a correct prediction from each distraction class. As shown in that figure, the name of the class is added over the top of the bounding box containing the image of the driver. The model uses what it has learned from the training images to sort the target area, and it predicts the distraction class based on the movement of the drivers. Along with the name of the predicted class, the model also gives the level of confidence. For instance, the first image in Figure 26 shows that the model has predicted that the driver is holding a “cellphone” with 90% confidence. As it can be seen, the model has predicted most images with a confidence of 90% or higher.

In some cases, the model predicted distractions as “non-distraction,” which could be termed “false negatives.” On the contrary, very few images with “non-distraction” were predicted by the model as one kind of distraction or another. The

dominance of false-negative over false-positive values indicates that the model rarely predicts non-distracted drivers as law violators. On the other hand, it would also be unable to detect some law violators due to the prevalence of false negatives.


**Figure 26** *Correctly Detected Distractions*

## Conclusion

This chapter demonstrates the stepwise procedure used to detect driver behaviors associated with various distractions using video data. The study presented

here proposed a novel approach to longitudinal video data collection on the high crash corridors of New Jersey. After data preprocessing, an artificial intelligence model titled YOLOv5 was used to train and test driver behaviors. The system was trained with 90% of the representative images of each distraction class and tested with the rest. The model achieved a fair accuracy of 85.01% when predicting the type of distraction visible in the testing images. From the subsequent model evaluation, it was found to have delivered more false negatives than false positives.

Further improvements in system accuracy could be made by training the model with more representative images of the various types of distraction. Collecting more observational data under diverse weather, seasonal, and lighting conditions could also add diversity to the training data and improve the accuracy of the model. Further enhancing the detection of distracted driving could also be achieved through data augmentation. Future research on these kinds of driver images could be done by annotating multiple objects in the training images (i.e., with micro annotation) or by fine-tuning the detector. Driver behaviors could also be trained and tested using other detection algorithms like Reset, HRNN, and Xception. The findings of this study could help researchers and policymakers to understand both the challenges and the possible scope of detecting driver behaviors from cameras located outside the car.

# CHAPTER 6: Conclusion and Recommendations

## Introduction

This study provided an evaluation of distracted driving from three perspectives. Firstly, it found the factors contributing to crashes involving distracted driving. Secondly, it examined the changing patterns in distraction events or in the behaviors of the drivers. Thirdly, it investigated a method for the detection of distracted driving. A summary of the results of these chapters is provided in detail in the following sections. It should be noted that the results presented in the observational study support the research hypothesis; that the variation of temporal and roadway features significantly influences driver behavior and their patterns of getting distracted.

## Summary of Results

### *Crash Analysis*

The crash analysis chapter aimed at finding those factors that contributed to crashes involving cellphone distractions. The summary of its results is as follows:

- Young drivers are less prone to fatal crashes caused by distractions (13.65%) compared to older drivers (18.17%).
- An increase in the total number of vehicles involved in a distracted driving crash increases the likelihood of ‘injury’ crashes by 39%.
- An increase in the speed limit resulted in an increase in the likelihood of injury crashes by 2.8%.
- Driving under the influence of substances (e.g., alcohol or drugs) increased the probability of serious injury crashes by 4.7%
- The injury severity level is more likely to increase in crashes during the early morning (midnight to 6:30 a.m.)
- An increase in traffic volume led to a decrease in the likelihood of ‘injury’ crashes by 21.9%.
- An urban road setting has a significant impact on injury severity, with the probability of “injury” found to decrease by 43.08%

### *Event Data Analysis*

Based on the 14,835 miles of collected event data from ten different corridors in New Jersey (with variations for peak/off-peak hours, seasons, the day of the week, signalized/unsignalized roads, toll/non-toll roads, number of lanes, posted speed limit, and median type), the following findings were reached:

- “Handheld cellphone” is the leading type of distraction, irrespective of time, type of roadway, season, and the geometric properties of the roadway.
- “Receiving calls” significantly increased during the weekdays, on the unsignalized and non-toll roads compared to the weekends, signalized corridors, and toll roads.
- The behavior of “eating/drinking” significantly increased on the signalized road compared to unsignalized road, and during the summer compared to the spring.
- The “fidgeting/grooming” distraction events significantly increased during the summer season compared to the spring.
- The “radio/reaching object” distraction event significantly increased on the weekdays compared to the weekend.
- Summer (24.4%) has a higher rate of distractions than spring (20.8%).
- An increase in speed limit significantly increased most of the distraction events.
- A “curbed” median encountered a significant reduction in distraction events compared to “unprotected” and “positive” median.
- An increase in median width significantly decreased distraction events.
- An increase in shoulder width significantly increased distraction events.
- An increase in the number of lanes significantly decreased the distraction events.

### *Distracted Driving Detection*

Based on the detection of driver behavior from observational video recording data using YOLOv5, the following findings were reached:

- Detection of driver behavior from cameras outside the car requires well-organized data preprocessing.
- The YOLOv5 model accurately predicted 85.01% of the driver distraction behaviors.
- The model prediction produces fewer false-positive values than false negatives.

## Recommendations and Future Work

The findings of the cellphone-related crash analysis could help engineers and policymakers to take appropriate countermeasures as a way of minimizing the severity of such crashes. Increasing road divisions, speed controls, increasing road friction by innovative materials, offering incentives for not using a cellphone while driving, and raising awareness through campaigns like *U-drive, U text, U Pay* can decrease crash severity due to cellphone distractions. In terms of future work, researchers could conduct further research by extending the years of crash data analyzed. A comparison of the findings through the analysis of crash data using various statistical logit and probit models could also be further investigated.

The findings from the observational data can also help practitioners and policymakers in various ways. For example, enforcement could be made stricter during the summer, during off-peak hours, and on weekdays because of the higher prevalence of distractions during these times. These results also emphasize the importance of restricting cellphone use, as it was the primary source of distraction for the majority of drivers. Various incentives could be provided by insurance companies to restrict driver use of cellphones, including texting and receiving calls.

Future research on longitudinal observational data analysis could be conducted on a more diverse dataset. For instance, the fall and winter seasons could form part of an extended study. Variations in event frequency could then be seen by performing more observations on more corridors. Finally, the event data could be collected during different years to find a yearly comparison of distraction and eliminate temporal bias. Cross-sectional observation of such events could also be done, and this method of data collection compared with the findings of the longitudinal data.

Detection of driver behavior from cameras outside the car is a promising avenue for future research. As there is no publicly available dataset on driver behavior from outside the car, efforts could go into preparing just such a training dataset for future researchers. The implementation of modern image-processing techniques with advanced deep learning and machine learning algorithms is also recommended for future studies. The video data from a static camera mounted on roadside polls or at toll plazas would likely also provide valuable images of driver behavior. Additionally, extracting images from more video data could help to obtain better performance in the detection of distracted driving, and the number of training images could be increased by various preprocessing techniques like data augmentation and glare removal. Further improvements to detection are also possible by automating the process of extracting frames from the observational videos. These improvements in the efficacy of such detection models will ultimately help practitioners and professionals to detect driver distractions in real-time.

# REFERENCES

Abouelnaga, Y., Eraqi, H. M., & Moustafa, M. N. (2017). Real-time distracted driver posture classification. *arXiv preprint arXiv:1706.09498*.

Aceable (2014). T*exting while driving penalties by state.* Retrieved from https://www.aceable.com/blog/texting-and-driving-fines-by-state.

Adebisi, A., Ma, J., Masaki, J., & Sobanjo, J. (2019). Age-related differences in motor-vehicle crash severity in California. *Journal of Safety*, *5*(3), 48.

Adanu, E. K., Smith, R., Powell, L., & Jones, S. (2017). Multilevel analysis of the role of human factors in regional disparities in crash outcomes. Accident Analysis & Prevention, 109, 10-17.

Ahangari, S., Jeihani, M., & Dehzangi, A. (2019). A machine learning distracted driving prediction model. In *Proceedings of the 3rd International Conference on Vision, Image and Signal Processing*, 1-6.

Al-Bdairi, N. S., & Hernandez, S. (2020). Comparison of contributing factors for injury severity of large truck drivers in run-off-road crashes on rural and urban roadways: accounting for unobserved heterogeneity. *International Journal of Transportation Science and Technology, 9*(2), 116-127.

Alotaibi, M., & Alotaibi, B. (2020). Distracted driver classification using deep learning. *Signal, Image and Video Processing*, *14*(3), 617-624.

Asbridge, M., Brubacher, J. R., & Chan, H. (2013). Cell phone use and traffic crash risk: a culpability analysis. *International Journal of Epidemiology, 42*(1), 259-267.

Astrain, I., Bernaus, J., Claverol, J., Escobar, A., & Godoy, P. (2003). Prevalence of mobile phone use while driving vehicles. *Gaceta Sanitaria, 17*(1), 66-69.

Baheti, B., Gajre, S., & Talbar, S. (2018). Detection of distracted driver using convolutional neural network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 1032-1038.

Baheti, B., Talbar, S., & Gajre, S. (2020). Towards computationally efficient and realtime distracted driver detection with MobileVGG network. *IEEE Transactions on Intelligent Vehicles*, *5*(4), 565-574.

Bakhit, P. R., Guo, B., & Ishak, S. (2018). Crash and near-crash risk assessment of distracted driving and engagement in secondary tasks: a naturalistic driving study. *Transportation Research Record, 2672*(38), 245-254.

Beanland, V., Fitzharris, M., Young, K. L., & Lenné, M. G. (2013). Driver inattention and driver distraction in serious casualty crashes: data from the Australian national crash in-depth study. *Accident Analysis & Prevention, 54*, 99–107.

Behnood, A., Modiri-Gharehveran, M., & Moradkhani Roshandeh, A. (2016). The effects of drivers' behavior on driver-injury severities in Iran: an application of the mixed-logit model. *Scientia Iranica, 23*(6), 2429-2440.

Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.

Bommer, W. (2018).*Observational study of distracted driving due to electronic device useamong California drivers for 2018: summary report.* California Office of Traffic Safety.

Braitman, K. A., & Braitman, A. L. (2017). Patterns of distracted driving behaviors among young adult drivers: exploring relationships with personality variables. *Transportation Research Part F: Traffic Psychology and Behaviour*, *46*, 169-176.

Cabrall, C., Janssen, N., Goncalves, J., Morando, A., Sassman, M., & De Winter, J. (2016). Eye-based driver state monitor of distraction, drowsiness, and cognitive load for transitions of control in automated driving. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 001981-001982. IEEE.

Caird, J. K., Willness, C. R., Steel, P., & Scialfa, C. (2008). A meta-analysis of the effects of cell phones on driver performance. *Accident Analysis & Prevention, 40*(4), 1282–1293.

Cambridge Mobile Telematics (2020). *The harsh realities of phone distraction*. Retrieved from https://www.cmtelematics.com/phone_distraction_report/

Cao, H., Zhang, Z., Song, X., Wang, H., Li, M., Zhao, S., & Wang, J. (2020). An investigation on the link between driver demographic characteristics and distracted driving by using the SHRP 2 naturalistic driving data. *Journal of Intelligent and Connected Vehicles*.

Cardell, S., & Dunbar, F. (1980). Measuring the societal impacts of automobile downsizing. *Transportation Research Part A: General, 14*(5-6), 423-434.

Carney, C., Harland, K. K., & McGehee, D. V. (2016). Using event-triggered naturalistic data to examine the prevalence of teen driver distractions in rear-end crashes. *Journal of Safety Research, 57*, 47-52.

Casado-Sanz, N., Guirao, B., & Attard, M. (2020). Analysis of the risk factors affecting the severity of traffic accidents on Spanish crosstown roads: The driver's perspective. *Sustainability, 12*(6), 2237.

Charlton, S. G. (2009). Driving while conversing: Cell phones that distract and passengers who react. *Accident Analysis & Prevention, 41*(1), 160-173.

Chen, L., Chen, C., Srinivasan, R., McKnight, C., Ewing, R., & Roe, M. (2012). Evaluating the safety effects of bicycle lanes in New York City. *American Journal of Public Health, 102*(6), 1120-1127.

Chen, Z., & Lym, Y. (2021). The influence of built environment on distracted driving related crashes in Ohio. *Transport Policy, 101*, 34-45.

Claveria, J. B., Hernandez, S., Anderson, J. C., & Jessup, E. L. (2019). Understanding truck driver behavior with respect to cell phone use and vehicle operation. *Transportation Research Part F: Traffic Psychology and Behavior, 65*, 389-401.

Covington, T (2021). *Distracted driving statistics: research and facts in 2021.* The Zebra. Retrieved from https://www.thezebra.com/resources/research/distracted-driving-statistics/

Cramer, S., Mayer, J., & Ryan, S. (2007). College students use cell phones while driving more frequently than found in government study. *Journal of American College Health, 56*(2), 181-184.

Curry, A. E., Hafetz, J., Kallan, M. J., Winston, F. K., & Durbin, D. R. (2011). Prevalence of teen driver errors leading to serious motor vehicle crashes. *Accident Analysis & Prevention*, *43*(4), 1285-1290.

D'Souza, K. (2012). An analysis of transit bus driver distraction using multinomial logistic regression models. *53rd Annual Transportation Research Forum*, 1-20.

De Castro, M. J. C., Medina, J. R. E., Lopez, J. P. G., de Goma, J. C., & Devaraj, M. (2018). A non-intrusive method for detecting visual distraction indicators of transport network vehicle service drivers using computer vision. In *2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, 1-5. IEEE.

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), pp. 2636–2641.

Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., ... & Knipling, R. R. (2006). *The 100-car naturalistic driving study, phase II-results of the 100-car field experiment* (No. DOT-HS-810-593). United States Department of Transportation. National Highway Traffic Safety Administration.

Donmez, B., Boyle, L. N., & Lee, J. D. (2007). Safety implications of providing real-time feedback to distracted drivers. *Accident Analysis & Prevention, 39*(3), 581-590.

Eby, D. W., Vivoda, J. M., & Louis, R. M. S. (2006). Driver hand-held cellular phone use: A four-year analysis. *Journal of Safety Research*, *37*(3), 261-265.

Elliott, A. C., & Hynan, L. S. (2011). A SAS® macro implementation of a multiple comparison post hoc test for a Kruskal–Wallis analysis. *Computer Methods and Programs in Biomedicine, 102*(1), 75-80.

Elqattan, Y., Moustafa, M. N., & El-Shafey, M. H. (2019). System for detecting and reporting cell phone distracted drivers. In *2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA)*, 215-221.

EndDD (2021). *Together we can end distracted driving*. Retrieved from https://www.enddd.org/end-distracted-driving/.

Eraqi, H. M., Abouelnaga, Y., Saad, M. H., & Moustafa, M. N. (2019). Driver distraction identification with an ensemble of convolutional neural networks. *Journal of Advanced Transportation*.

Failla, Zak (2019). *Distracted driving crackdown starts in Connecticut*. Stratford Daily Voice. Retrieved from https://dailyvoice.com/connecticut/stratford/police-fire/distracted-driving-crackdown-startsin-connecticut/773168/

Farmer, C. M., & Braitman, K. A. (2010). Cell phone use while driving and attributable crash risk. *Traffic Injury Prevention, 11*(5), 466–470.

Fatality Analysis Reporting System. (2020). *Drivers involved in fatal crashes, by previous driving record and license compliance - state: USA, year: 2019.* Retrieved from https://www-fars.nhtsa.dot.gov/People/PeopleDrivers.aspx

Ferguson, S. A. (2003). Other high-risk factors for young drivers—how graduated licensing does, doesn't, or could address them. *Journal of Safety Research, 34*(3), 71-77.

Fitch, G. M., Grove, K., Hanowski, R. J., & Perez, M. A. (2014). Compensatory behavior of drivers when conversing on a cell phone: Investigation with naturalistic driving data. *Transportation Research Record*, *2434*(1), 1-8.

Ford Driving Skills for Life. (2018). Retrieved from https://www.drivingskillsforlife.com/

Foss, R.D. and Goodwin, A.H., (2014). Distracted driver behaviors and distracting conditions among adolescent drivers: findings from a naturalistic driving study. *Journal of Adolescent Health*, *54*(5), 50-60.

Gallahan, S. L., Golzar, G. F., Jain, A. P., Samay, A. E., Trerotola, T. J., Weisskopf, J. G., & Lau, N. (2013). Detecting and mitigating driver distraction with motion capture technology: distracted driving warning system. In *2013 IEEE Systems and Information Engineering Design Symposium*, 76-81. IEEE.

Gao, M., Cai, Q., Zheng, B., Shi, J., Ni, Z., Wang, J., & Lin, H. (2021). A hybrid YOLO v4 and particle filter based robotic arm grabbing system in nonlinear and non-gaussian environment. *Electronics*, *10*(10), 1140.

Gliklich, E., Guo, R., & Bergmark, R. W. (2016). Texting while driving: A study of 1211 US adults with the Distracted Driving Survey. *Preventive Medicine Reports*, *4*, 486-489.

Gordon, C. P. (2009). *Crash studies of driver distraction*. Boca Raton, FL: CRC Press.

Governors Highway Safety Association (2019). *Distracted driving*. Retrieved from https://www.ghsa.org/state-laws/issues/Distracted-Driving.

Gras, M. E., Planes, M., Font-Mayolas, S., Sullman, M. J., Jiménez, M., & Prat, F. (2012). Driving distractions in Spain. In *Driver Behaviour and Training*, 299-305. Ashgate Publishing Ltd.

Haigney, D. E., Taylor, R. G., & Westerman, S. J. (2000). Concurrent mobile (cellular) phone use and driving performance: task demand characteristics and compensatory processes. *Transportation Research Part F: Traffic Psychology and Behaviour, 3*(3), 113-121.

Haleem, K., & Gan, A. (2013). Effect of driver's age and side of impact on crash severity along urban freeways: A mixed logit approach. *Journal of safety research*, *46*, 67-76.

Hanowski, R. J., Perez, M. A., & Dingus, T. A. (2005). Driver distraction in long-haul truck drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 8*(6), 441–458.

Hao, W., Kamga, C., & Wan, D. (2016). The effect of time of day on driver's injury severity at highway-rail grade crossings in the United States. *Journal of Traffic and Transportation Engineering (English edition), 3*(1), 37-50.

Hasan, A. S., Orvin, M. M., & Jalayer, M. (2021). *Analysis of distracted driving crashes in New Jersey using mixed logit model* (No. TRBAM-21-02932).

Hashim, H. H., Sarani, R., Sharifah Allyana, S. M. R., Mohamed, N., Jamilah, M. M., & Mohd Sadullah, A. F. (2011). Patterns of anatomical injury severity to different types of road users in road crashes. In *Proceedings of 3rd International Conference on Road Safety and Simulation.*

Hickman, J. S., & Hanowski, R. J. (2012). An assessment of commercial motor vehicle driver distraction using naturalistic driving data. *Traffic Injury Prevention, 13*(6), 612-619.

Hoffman, J. I. (2015). *Biostatistics for medical and biomedical practitioners*. Academic Press.

Horberry, T., Anderson, J., Regan, M. A., Triggs, T. J., & Brown, J. (2006). Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. *Accident Analysis & Prevention*, *38*(1), 185-191.

Huang, C., Wang, X., Cao, J., Wang, S., & Zhang, Y. (2020). HCF: a hybrid CNN framework for behavior detection of distracted drivers. *IEEE Access*, *8*, 109335-109349.

Insurance Information Institute (2018). *Facts + statistics: distracted driving.* Retrieved from https://www.iii.org/fact-statistic/facts-statistics-distracted-driving.

Insurance Information Institute (2020). *Fatality facts 2019: yearly snapshot.* Retrieved from https://www.iihs.org/topics/fatality-statistics/detail/yearly-snapshot#distraction

Islam, S., & Mannering, F. (2006). Driver aging and its effect on male and female single-vehicle accident injuries: some additional evidence. *Journal of Safety Research, 37*(3), 267–276.

Jashami, H., Abadi, M. G., & Hurwitz, D. S. (2017). Factors contributing to self-reported cell phone usage by younger drivers in the Pacific Northwest. In *Proceedings of the Ninth International Driving Symposium on Human Factors in Driver Assessment*, *Training and Vehicle Design*, 144-150.

Jenkins, S., Codjoe, J., Alecsandru, C., & Ishak, S. (2017). Exploration of the SHRP 2 NDS: development of a distracted driving prediction model. In *Advances in Human Aspects of Transportation*, 231-242.

Jiang, Z., Zhao, L., Li, S., & Jia, Y. (2020). Real-time object detection method based on improved YOLOv4-tiny. *arXiv preprint arXiv:2011.04244*.

Johnson, M. B., Voas, R. B., Lacey, J. H., McKnight, A. S., & Lange, J. E. (2004). Living dangerously: driver distraction at high speed. *Traffic Injury Prevention*, *5*(1), 1-7.

Jurewicz, C., Sobhani, A., Woolley, J., Dutschke, J., & Corben, B. (2016). Exploration of vehicle impact speed–injury severity relationships for application in safer road design. *Transportation Research Procedia, 14*, 4247-4256.

Khorashadi, A., Niemerier, D., Shankar, V., & Mannering, F. L. (2005). Differences in rural and urban driver-injury severities in accidents involving large-trucks: An exploratory analysis. *Accident Analysis & Prevention, 37*(5), 910–921.

Kidd, D. G., & Chaudhary, N. K. (2019). Changes in the sources of distracted driving among Northern Virginia drivers in 2014 and 2018: a comparison of results from two roadside observation surveys. *Journal of Safety Research*, *68*, 131-138.

Kim, J. K., Ulfarsson, G. F., Shankar, V. N., & Mannering, F. L. (2010). A note on modeling pedestrian-injury severity in motor-vehicle crashes with the mixed logit model. *Accident Analysis & Prevention, 42*, 1751-1758.

Kim, J.-K., Ulfarsson, G. F., & S. Kim, a. V. (2013). Driver-injury severity in single-vehicle crashes in California: A mixed logit analysis of heterogeneity due to age and gender. *Accident Analysis & Prevention, 50*, 1073–1081.

Kitch, A. (2018). State and federal efforts to reduce distracted driving. *In National Conference of State Legislatures*.

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.* National Highway Traffic Safety Administration.

Kong, X., Das, S., Zhou, H., & Zhang, Y. (2021). Characterizing phone usage while driving: safety impact from road and operational perspectives using factor analysis. *Accident Analysis & Prevention*, *152*, 106012.

Kusano, K., & Gabler, H. (2013). Characterization of opposite-direction road departure crashes in the united states. *Transportation Research Record, 2377*(1), 14-20.

Lambert, C. M. (2005). *Does birth order affect intelligence?* Rowan University.

Lansdown, T. C. (2009). Frequency and severity of in-vehicle distractions: A self-report survey. In *1st International Conference on Driver Distraction and Inattention (DDI 2009)* Chalmers University of Technology, Swedensafer Vehicle And Traffic Safety Centreinrets-ARCUEIL, France.

Leekha, M., Goswami, M., Shah, R. R., Yin, Y., & Zimmermann, R. (2019, September). Are you paying attention? detecting distracted driving in real-time. In *2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM)*,171-180). IEEE.

Li, H. R., Pickrell, T. M., & KC, S. (2016). *The 2015 national survey of the use of booster seats* (No. DOT HS 812 309).

Liang, Y., & Lee, J. D. (2014). A hybrid Bayesian Network approach to detect driver cognitive distraction. *Transportation Research Part C: Emerging Technologies*, *38*, 146-155.

Liang, Y., Horrey, W. J., & Hoffman, J. D. (2015). Reading text while driving: Understanding drivers' strategic and tactical adaptation to distraction. *Human Factors*, *57*(2), 347-359.

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.

Liang, Y., Reyes, M. L., & Lee, J. D. (2007). Real-time detection of driver cognitive distraction using support vector machines. *IEEE transactions on Intelligent Transportation Systems*, *8*(2), 340-350.

Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014, September). Microsoft coco: common objects in context. In *European Conference on Computer Vision*, 740-755. Springer, Cham.

Liu, C., & Subramanian, R. (2009). *Factors related to fatal single-vehicle run-off-road crashes* (No. HS-811 232).

Mannering, F. L., & Bhat, C. R. (2014). Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research, 1*, 1-22.

Manski, C. F., & McFadden, D. (Eds.). (1981). *Structural analysis of discrete data with econometric applications*, 2-50. Cambridge, MA: MIT press.

Marchese, C. (2019). *Distracted driving and crash responsibility in fatal USA collisions 1991-2015* (Doctoral dissertation).

Mase, J. M., Chapman, P., Figueredo, G. P., & Torres, M. T. (2020). A hybrid deep learning approach for driver distraction detection. In *2020 International Conference on Information and Communication Technology Convergence (ICTC)*,1-6.

McEvoy, S., & Stevenson, M. (2007). An exploration of the role of driver distraction in serious road crashes. In *International Conference on the Distractions in Driving, 2005,* Sydney, New South Wales, Australia.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics, 15*(5), 447-470.

Milton, J. C., Shankar, V. N., & Mannering, F. L. (2008). Highway accident severities and the mixed logit model: an exploratory empirical analysis. *Accident Analysis & Prevention, 40*(1), 260-266.

Murray P., Lin, J., Rojanasakul, M., & Stock, K. (2019). *Drivers refuse to put down their phones: people keep dying.* Bloomberg. Retrieved from https://www.bloomberg.com/graphics/2019-distracted-drivers/.

Narine, S., Walter, L. K., & Charman, S. C. (2009). Mobile phone and seat belt usage rates in London 2009. *Trl Published Project Report*.

National Highway Traffic Safety Administration (2019). Distracted driving 2019: summery of statistical findings.DOT HS 813 111. U.S. Department of Transportation, Washington, DC, 2021.

National Highway Traffic Safety Administration (2020). *Distracted driving*. NHTSA. Retrieved from https://www.nhtsa.gov/risky-driving/distracted-driving

National Security Council (2019). *Technology solutions*. Retrieved from https://www.nsc.org/road-safety/safety-topics/distracted-driving/technology-solutions

National Traffic Law Center (2017). *Investigation and prosecution of distracted driving cases (Report No. DOT HS 812 407)*. Washington, DC: National Highway Traffic Safety Administration.

New Jersey State Police. (2019). *Fatal Motor Vehicle Crash*. Retrieved from https://www.njsp.org/information/pdf/fcr/2019_fatal_crash_report.pdf

Neyens, D. M., & Boyle, L. N. (2007). The effect of distractions on the crash types of teenage drivers. *Accident Analysis & Prevention, 39*(1), 206-212.

NHTSA (2012). *Blueprint for ending distracted driving.* Retrieved from https://www.automotivesafetycouncil.org/wp-content/uploads/2017/01/NHTSA-Blue-Print-for-Distracted-Driving.pdf

NHTSA (2014). *Distracted driving 2012*. Retrieved from https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812012

NHTSA (2018). *Distracted driving 2016*. National Highway Traffic Safety Adminidtration. Retrieved from https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812517

NHTSA (2019). *Distracted driving*. National Highway Traffic Safety Adminidtration. Retrieved from https://www.nhtsa.gov/risky-driving/distracted-driving.

NHTSA (2019). *U drive. U text. U pay.* Retrieved from https://www.nhtsa.gov/campaign/distracted-driving

NHTSA (2020). *Overview of motor vehicle crashes in 2019.* National Highway Traffic Safety. Retrieved from https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813060

NJDHTS (2021). *Numetric crash query.* Retrieved from https://njdhts.numetric.com/crash-query

NJDOT (2018). *Weekday average speed data.* Retrieved from https://www.state.nj.us/transportation/refdata/roadway/pdf/wim/18_Classdata_spd.pdf

NJDOT (2021). *NJDOT Web SLD Data Browser.* Retrieved from https://njsld.org/NJDOT/SLD/DataBrowser

Olson, R. L., Hanowski, R. J., Hickman, J. S., & Bocanegra, J. (2009). *Driver distraction in commercial vehicle operations* (No. FMCSA-RRT-09-042). United States. Department of Transportation. Federal Motor Carrier Safety Administration.

Oviedo-Trespalacios, O., Haque, M. M., King, M., & Washington, S. (2017). Effects of road infrastructure and traffic complexity in speed adaptation behaviour of distracted drivers. *Accident Analysis & Prevention, 101*, 67–77.

Oviedo-Trespalacios, O., King, M., Vaezipour, A., & Truelove, V. (2019). Can our phones keep us safe? A content analysis of smartphone applications to prevent mobile phone distracted driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *60*, 657-668.

Owens, J. M., McLaughlin, S. B., & Sudweeks, J. (2011). Driver performance while text messaging using handheld and in-vehicle systems. *Accident Analysis & Prevention*, *43*(3), 939-947.

Patel, D. (2020). *Evaluating the effectiveness of the pedestrian safety intervention program: behavioral and observational approach*. Rowan University.

Pickrell, T. M., & Jianqiang Ye, T. (2009). *Driver electronic device use in 2008* (No. DOT HS 811 184). United States. Department of Transportation. National Highway Traffic Safety Administration.

Prat, F., Planes, M., Gras, M. E., & Sullman, M. J. M. (2015). An observational study of driving distractions on urban roads in Spain. *Accident Analysis & Prevention*, *74*, 8-16.

Pucher, J., & Dijkstra, L. (2003). Promoting safe walking and cycling to improve public health: lessons from the Netherlands and Germany. *American Journal of Public Health*, *93*(9), 1509-1516.

Qi, Y., Vennu, R., & Pokhrel, R. (2020). *Distracted driving: a literature review.* FHWA-ICT-20-004.

Qin, L., Li, Z. R., Chen, Z., Bill, M. A., & Noyce, D. A. (2019). Understanding driver distractions in fatal crashes: An exploratory empirical analysis. *Journal of Safety Research, 69*, 23-31.

Ratner, B. (2009). The correlation coefficient: Its values range between+ 1/− 1, or do they?. *Journal of Targeting, Measurement and Analysis for Marketing*, *17*(2), 139-142.

Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.

Regan, M. (2007). Driver distraction: Reflections on the past, present and future. *Distracted driving. Sydney, NSW: Australasian College of Road Safety*, 29-73.

Regan, M. A., Lee, J. D., & Young, K. (2008). *Driver distraction: theory, effects, and mitigation*. CRC press.

Retting, R. A., Ferguson, S., & McCartt, A. T. (2003). A review of evidence-based traffic engineering measures designed to reduce pedestrian–motor vehicle crashes. *American Journal of Public Health, 93*(9), 1456-1463.

Revelt, D., & Train, K. (1997). Mixed logit with repeated choices: households' choice of appliance efficiency level. *Review of Economics and Statistics. 80* (4), 1997, pp.., 80(4), 647–657.

Revelt, D., & Train, K. (1999). Customer-specific taste parameters and mixed logit. *University of California, Berkeley*.

Roque, C., Jalayer, M., & Hasan, A. S. (2021). Investigation of injury severities in single-vehicle crashes in North Carolina using mixed logit models. *Journal of Safety Research*, *77*, 161-169.

Roy, U., & Dissanayake, S. (2011). Comparison of factors associated with run-off-road and non-run-off road crashes in kansas. *Journal of the Transportation Research Forum, 50*, 69-86.

Royal, D. (2003). *National survey of speeding and unsafe driving attitudes and behaviors: 2002. Volume ii-findings report* (No. HS-809 688).

Sack, R., Foreman, C., Forni, S., Glynn, R., Lehrer, A., Linthicum, A., & Perruzzi, A. (2019). *Social media practices in traffic safety* (No. DOT HS 812 673). United States. National Highway Traffic Safety Administration.

Salkind, N. J. (Ed.). (2010). *Encyclopedia of research design* (Vol. 1). sage.

Savolainen, P. T., Mannering, F. L., Lord, D., & Quddus, M. A. (2011). The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. *Accident Analysis & Prevention*, *43*(5), 1666-1676.

Schroeder, P., Wilbur, M., & Peña, R. (2018). *National survey on distracted driving attitudes and behaviors-2015* (No. DOT HS 812 461). United States. National Highway Traffic Safety Administration.

Shah, A. (2013). *Text messaging and distracted driving: using voice dictation to make roads safer*. Ross School of Business, University of Michigan.

Skousen, Gulbrandsen, & Patience PLC. (2019). *4 ways local police fight against distracted driving*. HG.org. Legal Sources. Retrieved from https://www.hg.org/legal-articles/4-ways-local-police-fightagainst-distracted-driving-38203.

Smalheiser, N. (2017). *Data literacy: How to make your experiments robust and reproducible*. Academic Press.

State Farm (2016). *State farm distracted driver detection.* Retrieved from https://www. kaggle. com/c/state-farm-distracted-driver-detection/data

Statistics Solutions. (2022). Kruskal-Wallis Test. Statistics Solutions. Retrieved from https://www.statisticssolutions.com/kruskal-wallis-test/

Stavrinos, D., Jones, J. L., Garner, A. A., Griffin, R., Franklin, C. A., Ball, D., & Fine, P. (2013). Impact of distracted driving on safety and traffic flow. *Accident Analysis & Prevention, 61*, 63–70.

Stim, R. (2010). *Cell Phone and Texting Laws.* Nolo. Retrieved from https://www.drivinglaws.org/indexhf.php

Stimpson, J. P., Wilson, F. A., & Muelleman, R. L. (2013). Fatalities of pedestrians, bicycle riders, and motorists due to distracted driving motor vehicle crashes in the US, 2005–2010. *Public Health Reports, 128*(6), 436-442.

Streiffer, C., Raghavendra, R., Benson, T., & Srivatsa, M. (2017). Darnet: a deep learning solution for distracted driving detection. In *Proceedings of the 18th acm/ifip/usenix middleware conference: Industrial track,* 22-28.

Stutts, J. C., & Hunter, W. W. (2003). Driver inattention, driver distraction and traffic crashes. Institute of Transportation Engineers. *ITE Journal, 73*(7), 34.

Sullman, M. J. (2012). An observational study of driver distraction in England. Transportation research part F: *Traffic Psychology and Behaviour*, 15(3), 272-278.

Sullman, M. J., & Baas, P. H. (2004). Mobile phone use amongst New Zealand drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*, *7*(2), 95-105.

TeenDrive365: In School (2018). *Prepare lessons.* Retrieved from http://www.teendrive36Sinschool.com/educators/preparing-lessons.

Texas Department of Public Safety (2018). *Impact Texas Drivers (ITD) Program.* Retrieved from https://www.dps.texas.gov/section/driver-license/impact-texas-drivers-itd-program.

Thapngam, T., Yu, S., Zhou, W., & Beliakov, G. (2011). Discriminating DDoS attack traffic from flash crowd through packet arrival patterns. In *2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 952-957. IEEE.

Tison, J., Chaudhary, N., Cosgrove, L., & Preusser Research Group. (2011). *National phone survey on distracted driving attitudes and behaviors* (No. DOT HS 811 555). United States. National Highway Traffic Safety Administration.

Train, K. (2003). *Discrete Choice Methods with Simulation (1st ed.)*. Cambridge, UK: Cambridge University Press.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

Tran, D., Do, H. M., Lu, J., & Sheng, W. (2020). Real-time detection of distracted driving using dual cameras. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2014-2019.

Tzutalin (2017). *LabelImg*. Retrieved from https://github.com/tzutalin/labelImg

Ulfarsson, G., & Mannering, F. L. (2004). Differences in male and female injury severities in sportutility vehicle, minivan, pickup and passenger car accidents. *Accident Analysis & Prevention, 36*(2), 135-147.

Ultralytics (2020). *YOLOv5*. Retrieved from https://github.com/ultralytics/yolov5

Victor, T., Dozza, M., Bärgman, J., Boda, C. N., Engström, J., Flannagan, C., ... & Markkula, G. (2015). *Analysis of naturalistic driving study data: Safer glances, driver inattention, and crash risk* (No. SHRP 2 Report S2-S08A-RW-1).

Walter, L. (2010). Seatbelt and mobile phone usage surveys: England and Scotland 2009. *Londres, Royaume-Uni: Department for Transport*.

Wang, C. Y., Liao, H. Y. M., Wu, Y. H., Chen, P. Y., Hsieh, J. W., & Yeh, I. H. (2020). CSPNet: A new backbone that can enhance learning capability of CNN. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 390-391. IEEE.

Wang, J., Wu, Z., Li, F., & Zhang, J. (2021). A Data augmentation approach to distracted driving detection. *Future Internet*, *13*(1), 1.

Wang, K., Liew, J. H., Zou, Y., Zhou, D., & Feng, J. (2019). Panet: few-shot image semantic segmentation with prototype alignment. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 9197-9206. IEEE.

Wang, L., Li, R., Wang, C., & Liu, Z. (2021). Driver injury severity analysis of crashes in a western China's rural mountainous county: taking crash compatibility difference into consideration. *Journal of Traffic and Transportation Engineering (English edition)*, *8*(5), 703-714.

Wang, Y., Bao, S., Du, W., Ye, Z., & Sayer, J. R. (2017). Examining drivers' eye glance patterns during distracted driving: Insights from scanning randomness and glance transition matrix. *Journal of Safety Research*, *63*, 149-155.

Washington, S., Karlaftis, M., Mannering, F., & Anastasopoulos, P. (2020). *Statistical and econometric methods for transportation data analysis*. Chapman and Hall/CRC.

Winston, C., Maheshri, V., & Mannering, F. (2006). An exploration of the offset hypothesis using disaggregate data: the case of airbags and antilock brakes. *Journal of Risk and Uncertainty*, *32*(2), 83-99.

World Health Organization (2011). *Mobile phone use: a growing problem of driver distraction*. Retrieved from https://apps.who.int/iris/bitstream/handle/10665/44494/9789241500890_eng.pdf

Wu, Q., Chen, F., Zhang, G., X.C., L., Wang, H., & Bogus, S. M. (2014). Mixed logit model-based driver injury severity investigations in single- and multi-vehicle crashes on rural two-lane highways. *Accident Analysis & Prevention, 72*, 105–115.

Xie, Y., Zhang, Y., & Liang, F. (2009). Crash injury severity analysis using bayesian ordered probit models. *Journal of Transportation Engineering, 135*(1), 18-25.

Xu, R., Lin, H., Lu, K., Cao, L., & Liu, Y. (2021). A forest fire detection system based on ensemble learning. *Forests*, *12*(2), 217.

Yan, X., Radwan, E., & Mannila, K. K. (2009). Analysis of truck-involved rear-end crashes using multinomial logistic regression. *Advances in Transportation Studies an International Journal, 17*, 39-52.

Yannis, G., Laiou, A., Papantoniou, P., & Christoforou, C. (2014). Impact of texting on young drivers' behavior and safety on urban and rural roads through a simulation experiment. *Journal of Safety Research*, *49*, 25-e1.

Young, K. L., & Lenné, M. G. (2010). Driver engagement in distracting activities and the strategies used to minimise risk. *Safety Science, 48*(3), 326-332.

Yucek, Tevfik (2020). *Counter - Tally Counter (2021.6.3)*. Retrieved from https://apps.apple.com/us/app/counter-tally counter/id1144352844.

Zendrive (2018). *Zendrive's Distracted Driving Snapshot*. Retrieved from https://d1x6dm64pjo2h2.cloudfront.net/casestudies/Zendrive_Distracted_Driving_2018.pdf

Zhou, M., & Chin, H. C. (2019). Factors affecting the injury severity of out-of-control single-vehicle crashes in Singapore. *Accident Analysis & Prevention, 124*, 104-112.

Zutobi. (2022). *The US Distracted Driving Report: A Major Cause of Road Accidents*. Zutobi Drivers Ed. Retrieved from https://zutobi.com/us/driver-guides/distracted-driving-report