APPROVAL SHEET

This

Thesis

is submitted in partial fulfillment of the requirements
for the degree of

Author: ARIK SMITH

Advisor: Dr. Arsalan Heydarian

Advisor:

Committee Member: Dr. William Scherer

Committee Member: Dr. Reid Bailey

Committee Member:

Committee Member:

Committee Member:

Committee Member:

Accepted for the School of Engineering and Applied Science:

Jennifer L. West, School of Engineering and Applied Science

Framework for Evaluating Roadway Impacts on E-Scooter Safety through Computer Vision and Human Sensing Techniques

Arik Capers Smith

A Thesis presented to the Graduate Faculty of the University of Virginia in Candidacy for the Degree of Master of Science

Master's Thesis Committee Members:
Dr. Arsalan Heydarian (Advisor)
Dr. William Scherer (Chair)
Dr. Reid Bailey

Department of Systems and Information Engineering

University of Virginia
May, 2024

Table of Contents

List of Figures

List of Tables

Acknowledgement

First, I would like to thank Jesus Christ for giving me the opportunity to complete this thesis and my master's program. Next, I would like to express my sincere gratitude to my thesis advisor, Dr. Arsalan Heydarian, for his invaluable guidance, support, and encouragement throughout the entire process of completing this thesis. His expertise, patience, and dedication have been indispensable in shaping the direction of my research and helping me overcome various challenges along the way.

I am also deeply thankful to the members of my thesis committee, Dr. Bill Scherer and Dr. Reid Bailey, for their valuable feedback, insightful comments, and scholarly input that have significantly contributed to the quality and depth of this work. Their expertise and commitment to academic excellence have been instrumental in refining my ideas and strengthening my argumentation.

I would also like to extend my appreciation to the University of Virginia (UVA) for providing the necessary resources, facilities, and support that enabled me to conduct this research effectively. Additionally, I am grateful to the UVA Systems Engineering Department for their financial assistance, which allowed me to focus on my studies and pursue my academic goals.

My heartfelt thanks go to my family for their unwavering love, encouragement, and understanding throughout this journey. Their constant support and belief in my abilities have been a source of motivation and inspiration.

Finally, I am indebted to my friends, colleagues, and mentors who have offered their encouragement, advice, and camaraderie during this challenging yet rewarding endeavor. Your friendship, intellectual exchange, and shared experiences have enriched my graduate journey in countless ways.

To all those who have contributed to this thesis, whether directly or indirectly, I am sincerely grateful for your support and encouragement. This achievement would not have been possible without your collective contributions and belief in my potential.

- Arik Smith

Executive Summary

Micro-mobility usage has exploded over the last decade. Based on data from the National Association of City Transportation Officials (NACTO), between 2010-2022, there have been over 730 million trips on shared bikes and e-scooters in the United States and Canada alone [1]. These means of transportation appeal to riders because they provide mobility that is flexible, sustainable, and cost-effective. Additionally, this alternative mode of transportation has been coined as the solution to the first- and last-mile problem in urban areas because they provide increased connectivity to public transportation hubs [2]. No vehicle within this growing realm of micro-mobility has had as steep of an adoption curve as e-scooters. In 2022, e-scooter trips accounted for about 45% of all trips taken on shared micro-mobility [1]. Along with a steep increase in usership comes an increase in accidents on this transportation platform [3]. Unfortunately, due to the novelty of e-scooters, there is a lack of policy and infrastructure governing their existence in our transportation systems [4]. The focus of the research and analysis herein is to introduce and validate a novel framework to evaluate the impact of roadway features and conditions on e-scooter rider behaviors. This novel framework integrates advanced computer vision and human sensing techniques to identify where objects and conditions on the road may impact the e-scooter rider's physiological and/or behavioral responses. Specifically, this research demonstrates the merits of the proposed framework through beginning to analyze how certain road situations, such as passing a pedestrian, passing a bus, encountering an occupied crosswalk, etc. impact e-scooterists' behaviors using established metrics such as gaze entropy, variability, and road center fixations.

In the context of e-scooters, gaze entropy provides insight into the level of visual exploration and attention distribution exhibited by e-scooter riders, reflecting their engagement with the environment and potential implications for safety and situational awareness [5]. In the same context, gaze variability indicates how consistently or erratically e-scooter riders distribute their visual attention across their surroundings, potentially reflecting levels of distraction or situational awareness during riding [6]. Road center fixations indicate how consistently riders visually prioritize the center of the road while navigating, reflecting their focus on maintaining a safe path of travel and potential awareness of surrounding hazards [7]. In summary of our analysis, which included 10 hours of e-scooter data, we find that the situation wherein a rider

switches from the "bike lane to [the] crosswalk", which commonly occurs at intersections, the gaze transition entropy and gaze variability in their eye-tracking data are the highest among all situations at 28.68 bits and 209.96 pixels respectively. This is most likely due to the difficulty of switching roadway infrastructures and navigating an intersection at the same time. We also find that the situation "road fixture," which includes the navigation of speed bumps/tables, manhole covers, and potholes, has the highest percentage of road center fixations at 76%. This can be attributed to the rider's need to focus on the obstacle positioned within their path of travel to safely navigate it.

1. Introduction

Micro-mobility usage has exploded over the last decade. Based on data from the National Association of City Transportation Officials (NACTO), between 2010-2022, there have been over 730 million trips on shared bikes and e-scooters in the United States and Canada alone [1]. These means of transportation appeal to riders because they provide mobility that is flexible, sustainable, and cost-effective [2]. However, as micro-mobility usage continues to soar, so do concerns regarding rider safety. Between 2017 and 2021, injuries connected to micro-mobility vehicles spiked 127% to 77,200 [8]. Within this time-period, e-scooter users experienced the greatest increase in injuries and fatalities [9]. The Centers for Disease Control and Prevention (CDC) concluded that 20 out of every 100,000 e-scooter trips result in an injury [10]. The motivation for the research in this thesis is to close the gaps in knowledge that exist in our understanding of what parts of an e-scooterist's environment affect their behavioral and physiological responses and therefore their safety. Specifically, this thesis will begin to close the gap in knowledge of how e-scooterists' behaviors relate to differences in roadway situations and conditions by using human sensing and computer vision technologies to create a framework for e-scooter data collection.

A successful framework for behavioral e-scooter data collection is characterized by being able to collect data that comprehensively captures the environment around the rider and the way the rider is behaving within that environment. Therefore, the proposed framework utilizes the capabilities of computer vision to capture the environment around the rider, and sensory devices like eye-tracking and GPS to monitor the rider's behavioral response to the environment.

In addition to the behavioral data collection framework, this thesis will contribute the first e-scooter dataset tailored to be useful in the analysis of e-scooter rider behavior. As e-scooters become the preferred mode of micro-mobility travel, it is important to have a functional framework for data collection and a dataset that is comprehensive of the components needed to conduct complex analysis. The vision for the final result of the work done in this thesis is the development of a safety system for e-scooters that, given the environment and state of the rider, can act as a responsive and alert system to protect the rider.

1.1 Injuries Are Higher Using E-Scooters

Studies show that a large reason for the increased injury statistics associated with e-scooters is due to the architectural nature of the e-scooter itself, a lack of experience among riders, and the proclivity of riders to take risks [11]. Due to a decreased wheel radius, and an often lack of suspension, when compared to a bike, the vibration levels experienced by the rider are greater than when compared to the same. For example, someone riding an e-scooter will encounter a far greater disturbance to their intended path of travel when encountering roadway fixtures such as manhole covers and speed-humps than someone riding a bicycle. Additionally, the lessened turning radius of an e-scooter makes a quick adjustment of path to avoid a collision extremely unsafe, as the weight of the person operating the scooter can easily cause them to be thrown from the e-scooter.

In addition to the threats to rider safety that an e-scooter's architecture poses, the demographic that e-scooters attract are more prone to getting injured [12]. Shared e-scooters, which account for the most rides, are often perceived as convenient and intuitive to use, which attracts riders of varying skill levels and experience [3]. However, the simplicity of operating e-scooters can contribute to overconfidence among inexperienced riders, who may underestimate the risks associated with riding in urban environments, resulting in higher rates of accidents and injuries. Accordingly, a study conducted in Brisbane Australia concluded that illegal behaviors in terms of the maneuvering of the e-scooter and lack of safety equipment like helmets were more prevalent among those riding shared e-scooters [13].

Another contributing factor to higher injury rates with riding e-scooters is the risk-taking behaviors of riders. This observation could arguably be explained by the age demographics of those riding e-scooters, as data shows that e-scooter usage skews toward the younger demographic (20-39 years old) [14]. Accordingly, those who ride e-scooters are more likely to be injured due to risky behaviors such as riding while intoxicated or using their cell phone when compared to other micro-mobility options with different rider-age distributions [3].

Operating an e-scooter under the influence of drugs or alcohol is significantly more dangerous than operating a car or other micro-mobility options. Even at lower blood alcohol contents (.21-.6 g/kg), a rider's performance is decreased to 60% of its original level [15]. The steep performance drop-off is due to the increased dependence on the rider's physical performance, as safely navigating a path of travel on an e-scooter requires higher levels of proprioception and other motor skills [16]. These are primarily and significantly impacted when under the influence.

1.2 Infrastructure Design Lacks Proper Support For E-Scooters

A prevailing issue surrounding the use of e-scooters on present roadway infrastructures, is that most roadways were not originally designed to accommodate e-scooters or any other form of micro-mobility, excluding non-motorized bicycle [17]. Therefore, there are several issues that present themselves as threats to rider safety.

The first being the fact that the e-scooter is the transportation mode that does not quite have an infrastructural home [18]. There are multiple infrastructures that exist for different types of travelers; the most common ones are sidewalks for those who are walking, roads for those who are driving, and bike lanes for those who are cycling. The e-scooter rider is forced to navigate them all, which means that the e-scooter rider must be aware of all of the laws and customs concerning each infrastructure type to navigate them safely. This is not always the case and yet the typical e-scooter trip is split between sidewalks (18 %), bike lanes (11 %), and roadways (33 %), with 38 % being completed on other unclassified infrastructures [19].

In addition to the lack of dedicated road infrastructure for e-scooters, the same lack of infrastructure exists in relation to designated lanes or parking areas. Many cities are having trouble organizing and getting riders to comply with where e-scooters can be parked, which further exacerbates the challenges riders face [20]. Without designated spaces, e-scooters are often left cluttering sidewalks or obstructing pedestrian pathways, posing a hazard to both riders and pedestrians alike. Moreover, the absence of clear guidelines and regulations regarding e-scooter usage on roads adds to the confusion and potential for conflicts between different road users.

Lastly, the condition of roads themselves presents a significant obstacle to e-scooter riders. Potholes, uneven surfaces, and other road defects can pose serious safety risks to e-scooter users, who lack the protection afforded by enclosed vehicles [21]. The impact of these road imperfections is amplified for e-scooters due to their smaller wheels and less robust suspension systems compared to cars, making navigation through urban environments a challenging and potentially hazardous endeavor [22]. In order to fully integrate e-scooters into existing transportation networks, significant upgrades and modifications to road infrastructure are necessary to ensure the safety and accessibility of all road users.

1.3 Limited Studies on Micro-Mobility Behaviors Compared to Vehicles

Despite the increasing popularity of e-scooters, the research and availability of behavioral data on micro-mobility modes lag significantly behind that of cars. This lack of attention is further emphasized by the stark difference in the number of articles published on cars compared to those mentioning micro-mobility. This knowledge gap hinders our understanding of how e-scooter riders interact with other road users, navigate different types of infrastructure, and adhere to traffic laws [23].

One key challenge contributing to the dearth of research and behavioral data for micro-mobility modes is the relatively recent emergence of these transportation options compared to cars [24]. While cars have been the dominant mode of transportation for over a century, micro-mobility options have only gained traction in the past decade, with shared e-scooter and bike sharing schemes becoming prominent in urban landscapes [25]. As a result, there has been insufficient time for researchers to conduct comprehensive studies and gather longitudinal data on micro-mobility behaviors and impacts.

Additionally, the data collection infrastructure for micro-mobility modes is often less developed compared to that of cars, posing significant challenges for researchers. For instance, while cars are typically equipped with advanced tracking systems and onboard sensors that record detailed information about travel behavior, micro-mobility vehicles often lack standardized data collection mechanisms [26]. This makes it difficult to gather comprehensive data on factors such as trip duration, route choice, and user demographics, which are essential for understanding micro-mobility patterns and informing policy decisions.

Another factor contributing to the lack of research on micro-mobility is the fragmented nature of the industry and regulatory landscape. Unlike the automotive sector, which is characterized by well-established manufacturers and regulatory frameworks, the micro-mobility sector comprises a diverse array of stakeholders, including startups, ride-sharing companies, local governments, and advocacy groups [27]. This fragmentation can lead to inconsistencies in data collection practices, data sharing agreements, and research priorities, further hindering efforts to generate comprehensive insights into micro-mobility behaviors.

1.4 Summary of The Gaps in Research Related to The Impact of Roadway Design on E-Scooterists

1.4.1 Situational Categories

Based on the gaps observed in literature, and that there is a distinct need for research on the behavioral characteristics of e-scooterists, we have developed situational categories that are important to be studied in relation to how e-scooterist behave in different environments. The situational categories listed in Table 1 were curated based on their theoretical ability to cause discomfort and/or a heightened level of awareness among riders. Table 1 was developed as a means to help organize and explain these situational categories and the associated objects of interest that computer vision must be able to detect to recognize them.

Table 1: The selected situational categories, their descriptions, and the associated objects needing to be detected by computer vision to recognize the situations through computer vision.

Situational Category	Description	Object of Interest
Bike Lane or Road to Crosswalk	Rider transitions from a bike lane or a road shared by both bikes and cars to the crosswalk either at intersections or	Bike Lane Sign, Crosswalk Sign
Bike Lane to Road	Rider transitions from road bike lane to road	Road Surface, Bike Lane Sign
Close Proximity	Rider passes or encounters vehicles passing them at a close proximity	Cars, Trucks
High Speed Downhill	Rider encounters a visibly steep hill traveling downwards	
Intersection	Rider encounters an intersection (i.e. stop sign traffic light)	Stop Light, Stop Sign, Cars, Trucks, Bicycles, E-Scooters, People

Occupied Crosswalk	Rider travels through an occupied crosswalk	People, Crosswalk, Crosswalk Sign
Passing Bus	Rider passes a bus that is stopped to either drop-off or gather people	Bus
Passing Pedestrian on Sidewalk	Rider is traveling on the sidewalk and must pass pedestrians while doing so	People, Sidewalk
Road Fixtures	Rider encounters a road fixture such as a speed bump, speed table, or pothole	Speed Bump, Pothole

In general, when transitioning from one road or road infrastructure to another, danger exists because the type, speed, and rules of travel are typically different between infrastructure types [18]. Traveling through an intersection as an e-scooterist is inherently dangerous as there are far more points of conflict with several lanes of traffic intending to go in opposing directions [28]. Also, as an e-scooterist, one has little to no protection if involved in a collision [29]. Traveling from a bike lane or the road to a crosswalk can be dangerous because this is one of the locations e-scooterists come into close contact with pedestrians and are at a greater risk of getting into a collision, so increased awareness is warranted in this area [30]. When transitioning from a bike lane to a road, whether it be due to the bike lane ending or to avert an obstacle located in the bike lane, the rider must travel into a lane with much bigger and faster vehicles, which puts the rider at a much higher risk of injury [31]. When going from the road to a bike lane, a rider must be aware of other cyclist or scooterist already in the bike lane, therefore awareness must be high to prevent a collision with another traveler.

The next grouping of situational categories involves situations that serve as a threat to e-scooterists due to a high reliance on the e-scooterist being able to control the scooter and the architecture of the scooter not lending itself to being able to do so. An e-scooterist must exercise increased control of the e-scooter when in close proximity to other vehicles while on the road, when passing over a road fixture (speed bump, pothole, etc.), and when traveling at high speeds downhill to avoid collisions or being thrown from the e-scooter. These situations are worthy of heightened awareness due to the architecture of the e-scooter, spoken about in section 1.1. The relatively smaller wheels and wheelbase make it easy to deviate from a rider's path of travel due to errors by the rider themselves trying to avoid a vehicle while passing or something like a pothole or uneven surface causing a path-altering vibration to change the path of the e-scooter [32]. In these listed situations, a small error in path can result in a collision. The same problem exists with passing pedestrians on the sidewalk or traveling through an occupied crosswalk. Many times, the path of a pedestrian can be hard to predict when approaching from the front or behind, and as aforementioned the architecture of the scooter does not yield to fast adjustments, which can also easily throw the rider from the e-scooter, causing injuries. Lastly, when passing a bus that stopped to collect or drop-off passengers, it may depart from its stop unexpectedly which forces the rider to respond quickly by shifting further into the lane of oncoming traffic or slowing down abruptly, which can both cause injuries.

Within the objects of interest in table 1 are the components of an e-scooterists environment that computer vision adapted to e-scooters must be able to detect to determine whether the rider is encountering one of the situations also found in table 1. Therefore, the computer vision algorithm must be trained to recognize these objects as we will demonstrate in the methodology section.

1.4.2 Lack of Dataset on E-Scooter Behaviors

In addition to a lack of behavioral research for e-scooters, there is currently not a dedicated dataset containing e-scooter data for research, whereas integrated datasets encompassing various transportation modes have played a crucial role in enhancing research on urban mobility and transportation safety [33]. For instance, in the automotive sector, integrated datasets often combine vehicle telemetry, traffic flow data, and road infrastructure information to analyze driver behavior, identify congestion hotspots, and develop intelligent transportation systems [34]. Similarly, in public transportation, integrated datasets may incorporate passenger boarding and alighting patterns, vehicle occupancy rates, and service reliability metrics to optimize route planning and improve the overall passenger experience [35]. Additionally, integrated datasets involving cycling and pedestrian movements, coupled with urban design and land use data, contribute to the development of pedestrian-friendly infrastructure and promote active transportation initiatives in cities [36]. Overall, the integration of diverse datasets across different transportation modes facilitates comprehensive analyses and enables evidence-based decision-making to address the complex challenges of urban mobility and transportation planning.

Analysis of integrated datasets for e-scooters, containing variables such as speed, GPS coordinates, gaze activity, and object detection, presents a multifaceted approach to understanding rider behavior and enhancing e-scooter safety [37]. By merging these diverse data sources, researchers can gain comprehensive insights into the various aspects of e-scooter usage, including rider navigation patterns, interaction with the surrounding environment, and the potential hazards encountered during rides. For example, integrating GPS data with gaze activity and object detection can facilitate the identification of high-risk zones, such as intersections with heavy traffic or areas prone to pedestrian congestion [38]. Also, analyzing speed data in conjunction with gaze activity may reveal correlations between visual attentional shifts and changes in riding behavior, offering insights into the cognitive processes underlying rider decision-making and situational awareness [39].

Overall, the integration of diverse datasets across different dimensions of data for e-scooters will enable evidence-based decision-making that will help to address the complex challenges of achieving safety for those who ride e-scooters.

2. Literature and background: The Advancement of Technology and Analytical Methods and How They Increase Road Safety

2.1 Computer Vision data

Technologies like computer vision have already been proven to be making our roads safer for vehicles and pedestrians alike when applied to common transportation modes. For example, reported a 30% reduction in fatal road accidents in the decade between 2009-2019 [40]. Much of this decline is attributed to computer vision technology like advanced driver assistance systems (ADAS), which use real-time visual data to enhance situational awareness and mitigate the risk of accidents [41]. One way ADAS does this is through collision avoidance systems, which leverage cameras and sophisticated algorithms to detect potential collision hazards such as vehicles, pedestrians, and obstacles in the vehicle's path. When a threat is identified by the system, it then issues warnings to the driver and can even initiate automatic emergency braking to prevent or minimize the impact. This type of system has been proven to significantly reduce the likelihood of accidents caused by human error or inattention [42].

ADAS has also greatly improved lane-keeping assistance systems (LKA), which help drivers maintain proper lane positioning and reduce the risk of lane departure, which can lead to accidents [43]. A study done on fatal head-on and single vehicle crashes in Finland concluded that 27% of the accidents could have been prevented if currently available (LKA) systems were deployed [44]. These types of systems alert drivers when they are unintentionally leaving their lane of origin by analyzing video feeds from onboard cameras and identifying lane markings. Some advanced systems can actively intervene by applying gentle steering corrections to guide the vehicle back into the lane, thereby enhancing vehicle control and safety, especially on highways and during long-distance drives [45].

Another feature of ADAS is pedestrian detection technology. This feature is a crucial safety feature in modern vehicles, particularly in urban environments where interactions between vehicles and pedestrians are frequent. A study conducted in Germany considered pedestrian fatalities from 1999-2007 and concluded that if the vehicles involved were outfitted with pedestrian avoidance technology, the number of resulting fatalities would have decreased by 40% [46]. This life-saving technology uses visual data from cameras to recognize pedestrians near vehicles and issues warnings to the driver if a potential collision risk is detected [47]. This capability is instrumental in preventing accidents involving multiple types of vulnerable road users like pedestrians, cyclists, and e-scooterists and improving overall road safety. Overall, computer vision technology continues to play a pivotal role in advancing vehicle safety standards, paving the way for safer and more secure transportation systems for drivers, passengers, and pedestrians alike.

2.1.1 Computer Vision: Object Detection Algorithms

In the realm of autonomous and smart vehicles alike, object detection algorithms have become immensely significant for ensuring safe driving. Specifically, deep learning-based algorithms have emerged as pivotal tools due to their capability to achieve high detection accuracy while demanding fewer computing resources, thereby becoming indispensable in autonomous and smart driving systems [48].

Implementing current object detection algorithms on e-scooters presents challenges owing to various factors [49]. First, the frequent vibrations generated by e-scooter movements disrupt sensor data, posing difficulties in real-time motion artifact mitigation. Second, e-scooters operate in diverse environments with varying lighting conditions, weather, and road surfaces, making it challenging for object detectors to adapt consistently. Additionally, reconciling algorithmic demands with the limited capacity of e-scooters introduces further complexity. Therefore, a crucial consideration lies in striking a balance between the effectiveness and efficiency of object detectors.

There are two primary types of deep learning object detectors: two-stage detectors, which involve a preprocessing step for object proposal generation in the initial stage, followed by object classification and bounding box regression in the subsequent stage. On the other hand, single/one-stage detectors are end-to-end, eliminating the necessity for the region proposal process. Compared to two-stage detectors like Faster-RCNN [50] and Mask-RCNN [51], one stage detectors are more computationally efficient, faster in inference, and particularly suitable for real-time applications, especially on resource-constrained embedded devices like e-scooters. A notable example of a one-stage detector is You Only Look Once (YOLO), originally developed by Redmon et al. (2016) and further developed into YOLOv3 [52]. YOLOv3 strikes a balance between accuracy and speed, making it Proceedings Paper Formatting Instructions – 3 – Rev. 10/2015 one of the most widely used object detectors. Following YOLOv3, architectural modifications have been introduced to enhance accuracy and/or speed, resulting in versions like YOLOv4 [53], YOLO v5, YOLOv6, YOLOv7, and the latest YOLOv8 (Ultralytics 2023). These YOLO-derived object detectors can be configured with varying levels of model complexity, leading to different implementation variants.

2.2 Behavioral and Physiological Data

Despite the growing popularity of e-scooters as a mode of urban transportation, there remains a lack of studies focused specifically on the behavioral and physiological data that can be collected from e-scooter riders. While research on traditional modes of transportation such as cars has accumulated over decades, studies examining the behavioral and physiological characteristics associated with e-scooter use are relatively scarce. This gap in the literature is underscored by the small fraction of transportation research that is dedicated to micro-mobility modes, including e-scooters.

Additionally, the limited availability of behavioral and physiological data on e-scooter riders further accentuates this research gap. Compared to cars, which are equipped with sophisticated tracking systems and onboard sensors, e-scooters often lack standardized data collection mechanisms, posing challenges for researchers aiming to understand factors such as rider behavior and physiological interactions with the urban environment [54]. Therefore, there is a pressing need for more data-comprehensive studies focusing on e-scooter riders to address critical questions related to safety, usability, and the integration of e-scooters into urban transportation systems.

2.2.1 Eye-Tracking

In addition to computer vision, the collection of eye-tracking data is another feature that has become important to improving road safety. Eye-tracking systems often utilize infrared sensors or cameras to monitor the driver's eye movements and gaze patterns in real-time [55]. By analyzing where the driver is looking, these systems can assess the driver’s attention and focus on the road, providing valuable insights into their cognitive state and level of engagement with driving tasks [56]. One currently significant application of eye-tracking in vehicle safety is distraction detection. By detecting instances where the driver's gaze is diverted away from the road, such as towards a mobile phone or an in-vehicle display, these systems can issue warnings or alerts to bring the driver's attention back to driving, thereby reducing the risk of accidents caused by distracted driving [57].

Additionally, eye-tracking can also be utilized in fatigue detection systems, which monitor the driver's eye movements and blink patterns to assess their level of alertness and drowsiness [58]. Prolonged periods of inattention or decreased blink rates may indicate fatigue or drowsiness, which are major contributors to accidents on the road [59]. By analyzing eye-tracking data in conjunction with other physiological indicators, such as heart rate variability, these systems can accurately detect signs of driver fatigue and prompt the driver to take breaks or rest periods, ultimately enhancing safety during long-distance drives or late-night journeys [60].

Lastly, eye-tracking data can provide valuable insights into the driver's situational awareness and hazard perception skills [61]. By tracking the driver's gaze behavior, such as scanning for potential hazards or checking blind spots, these systems can assess their ability to anticipate and respond to changing road conditions [62]. By providing feedback and coaching based on eye-tracking data analysis, these systems can help improve the driver's hazard perception skills and promote safer driving behaviors, ultimately reducing the likelihood of accidents and improving overall road safety. Overall, the integration of eye-tracking into vehicle safety systems offers a proactive approach to mitigating the risks associated with driver distraction, fatigue, and lack of situational awareness, ultimately enhancing the safety and well-being of drivers and passengers alike.

Gaze Entropy

Entropy is generally a measure of randomness in a system. However, in literature, a principal factor in the analysis of eye-tracking is gaze entropy. Specifically, gaze transition entropy (GTE) and stationary gaze entropy (SGE) have correlations to human states such as workload, stress level, and emotions [63]. A fixation is generally defined as when gaze becomes stationary. The shift from one point of fixation to another rapidly is defined as a saccade. It is the different sequences and variations of saccades and fixations that have been correlated to stress level and workload [63].

GTE, specifically, refers to the measure of randomness or unpredictability in gaze transitions between different points of interest in a visual scene. This metric provides insights into the flexibility and efficiency of visual exploration strategies employed during tasks requiring dynamic shifts in attention, which is crucial for road safety [64]. Studies investigating transition gaze entropy have demonstrated its relevance in understanding gaze behavior and cognitive processing in the context of driving [65].

Alternatively, SGE is a metric used to quantify the randomness or unpredictability of gaze points while an individual fixates on a stationary point of interest. This measure provides valuable insights into the stability and consistency of attention allocation during tasks requiring focused visual attention, which is crucial for understanding driver behavior and road safety [66]. Studies in the field of road safety and driver behavior have demonstrated the utility of stationary gaze entropy in various contexts. For example, research by Land and Lee (1994) [67] used stationary gaze entropy to characterize fixation patterns during steering tasks. The study found that higher stationary gaze entropy was associated with increased steering variability and reduced lane-keeping performance, indicating that erratic gaze patterns impair the ability to maintain vehicle control and trajectory stability, thus impacting road safety.

Analyzing e-scooterist data using GTE and SGE can reveal important patterns in gaze behavior. Additionally, insights gained from such analyses can inform future studies and even technological innovations aimed at enhancing scooterist safety. Overall, leveraging stationary gaze entropy and gaze transition entropy in e-scooterist research represents an avenue for advancing our understanding of human factors in relation to e-scooters and greater micro-mobility.

Road Center Fixations

Generally, a fixation is characterized as when an individual's eyes stabilize on a specific point. More specifically, Road center fixations refer to when the driver is fixated on the center of the road and whatever may occupy that space [68]. Road center fixations are typically measured using eye-tracking technology. Such technology allows researchers to monitor and analyze drivers' eye movements in real time. By identifying the percentage of fixations that are towards the center of the road in different driving situations, researchers can gain valuable insights into the factors that influence attention allocation on the road [69].

In literature, road center fixations are closely connected to various aspects of driver behavior, including hazard detection, lane keeping, and navigation. Several studies have investigated the relationship between road center fixations and driving performance. For example, the study done by Land and Lee (1994) also found that drivers frequently fixate on the center of the road while steering, suggesting that road center fixations play a critical role in guiding steering movements and maintaining vehicle position within the lane [67]. Additionally, road center fixations have been linked to hazard detection and collision avoidance. Research by Underwood et al. (2003) demonstrated that experienced drivers tend to fixate on potential hazards located near the center of the road, allowing them to detect and respond to threats more effectively than novice drivers [70]. This finding underscores the importance of road center fixations in facilitating hazard perception and reducing crash risk.

Road center fixations are a fundamental aspect of driver gaze behavior that can be adapted to e-scooters to gain information on rider performance, including hazard detection, lane keeping, and navigation. Understanding the patterns of road center fixations can provide valuable insights into e-scooterists behavior and inform future studies to make the e-scooter itself and the surrounding infrastructure safer.

2.2.2 Smart Watch and Wearable

Heart rate data is increasingly becoming useful for enhancing safety in vehicles, particularly in the context of driver monitoring systems. By integrating biometric sensors into vehicles, these systems can continuously monitor the driver's heart rate, providing insights into their physiological state and overall well-being [71]. One significant application of heart rate data in vehicle safety is fatigue detection. By analyzing fluctuations in heart rate patterns, these systems can identify signs of driver fatigue or drowsiness, which are major contributors to accidents on the road [72]. Upon detecting such indicators, the system can issue warnings to the driver, prompting them to take a break or rest, thereby reducing the risk of accidents caused by impaired alertness [72].

Heart rate data can also be utilized in stress detection systems, which assess the driver's stress levels based on variations in their heart rate [73]. High levels of stress can impair cognitive function and decision-making abilities, increasing the likelihood of errors and accidents while driving. By monitoring heart rate data in real-time, these systems can detect elevated stress levels and provide interventions such as calming prompts or adaptive vehicle settings to help alleviate stress and maintain driver composure, ultimately enhancing safety on the road.

Lastly, heart rate data can complement other driver monitoring metrics, such as eye movements and facial expressions, to provide a comprehensive assessment of the driver's physiological and cognitive state [74]. By integrating multiple biometric inputs, vehicle safety systems can better understand the driver's overall condition and tailor interventions accordingly. This holistic approach to driver monitoring enables proactive safety measures, such as adjusting the vehicle's driving dynamics or alerting emergency services in case of a medical emergency, ultimately contributing to a safer driving experience for all road users.

Heart Rate & Heart Rate Variability

A common combination of metrics used in analyzing a human's state is heart rate and heart rate variability. Heart rate variability is a measure of heart rate fluctuations around the mean heart rate. More specifically, heart rate variability represents signal properties that are calculated through methods like the root mean squared of the successive difference (RMSSD). Decreased levels of RMSSD and increased heart rate have been linked to stress [75]. In addition, many studies have been conducted to link heart rate to perceptions of risk. For example, a study conducted on cyclists in Cork, Ireland, concluded that in situations wherein cyclists self-reported the perception of risk, recorded heart rates were higher in a statistically significant manner [76].

There are two common methods of recording heart rate data, and these are electrocardiogram (ECG) and Photoplethysmogram (PPG). ECG uses contact electrodes to measure the electrical activity of the heart, and PPG measures blood volume in the veins using infrared technology and then approximates heart rate. These methods of heart rate data collection can be used to measure the heart rate of e-scooterists in similar fashion as the study in Cork, Ireland. This would provide valuable information about various human states of the rider, in connection to changes in their environment.

2.3 Research Objective

The objective of the research herein is to create the foundations and test the preliminary framework for a system that ultimately integrates computer vision, eye-tracking, and heart rate data in real-time to enhance the safety of e-scooters. By equipping e-scooters with advanced sensor technologies, operators can monitor the rider's physiological state and cognitive engagement in real-time [37]. Eye-tracking systems can analyze the rider's eye movements and focus, providing insights into their attention and awareness of the surrounding environment [77]. Additionally, heart rate sensors can detect signs of fatigue or stress, alerting the rider to take breaks or rest periods when necessary to prevent accidents caused by impaired cognitive function or decreased alertness [78].

Additionally, computer vision technology can augment safety measures by providing a comprehensive view of the rider's surroundings and identifying potential hazards on the road. Cameras mounted on e-scooters can capture video data of the rider's environment, allowing intelligent algorithms to detect obstacles, pedestrians, and other vehicles in real-time. By analyzing this visual information alongside gaze and heart rate data, the system can assess the rider's situational awareness and provide proactive warnings or alerts about potential collision risks or hazardous road conditions, empowering them to make informed decisions and navigate safely through traffic [74].

Moreover, the integration of computer vision, eye-tracking, and heart rate data can enable personalized safety interventions tailored to the individual rider's needs and preferences. By correlating biometric and visual data with patterns of riding behavior, the system can dynamically adjust the e-scooter's speed, responsiveness, or route guidance to optimize safety. Additionally, in the event of an emergency, such as a sudden increase in heart rate or a distracted gaze, the system can automatically alert nearby riders, pedestrians, and emergency services, facilitating rapid assistance and intervention. Overall, the synergy between computer vision, eye-tracking, and heart rate data offers a multifaceted approach to improving e-scooter safety, promoting responsible riding behavior, and enhancing the overall riding experience for users.

3. Methodology

The overall aim of the research displayed in this thesis is to create a framework for an e-scooter data collection system. The proposed system is intended to integrate computer vision, gaze, and heart rate data in a manner that makes it possible to do complex analysis on the data gathered from e-scooter riders. The purpose behind the creation of this system is to learn more about e-scooter rider's behaviors. Specifically, how the behavior of riders change based on traffic situations, road conditions, speeds, etc. Computer vision provides a fast and efficient means of identifying features of an e-scooterist's environment, while based on available research, physiological responses such as gaze and heart rate data allow for the objective analysis of the way a rider is attentive and engaged with their surroundings. The general goal for the analysis included in this thesis is to test riders' behavioral and physiological responses (dependent variable) as the environment they are riding in changes (independent variable(s)). However, it is important to note that this thesis is based on a snapshot in time of a large and rather complicated research process intended to result in a refined and integrated data system for e-scooters. As a result, and as is demonstrated in figure 1, the methodology, results, and analysis sections will focus primarily on the computer vision and gaze integration techniques as these are the areas of this research that have progressed the furthest to date.

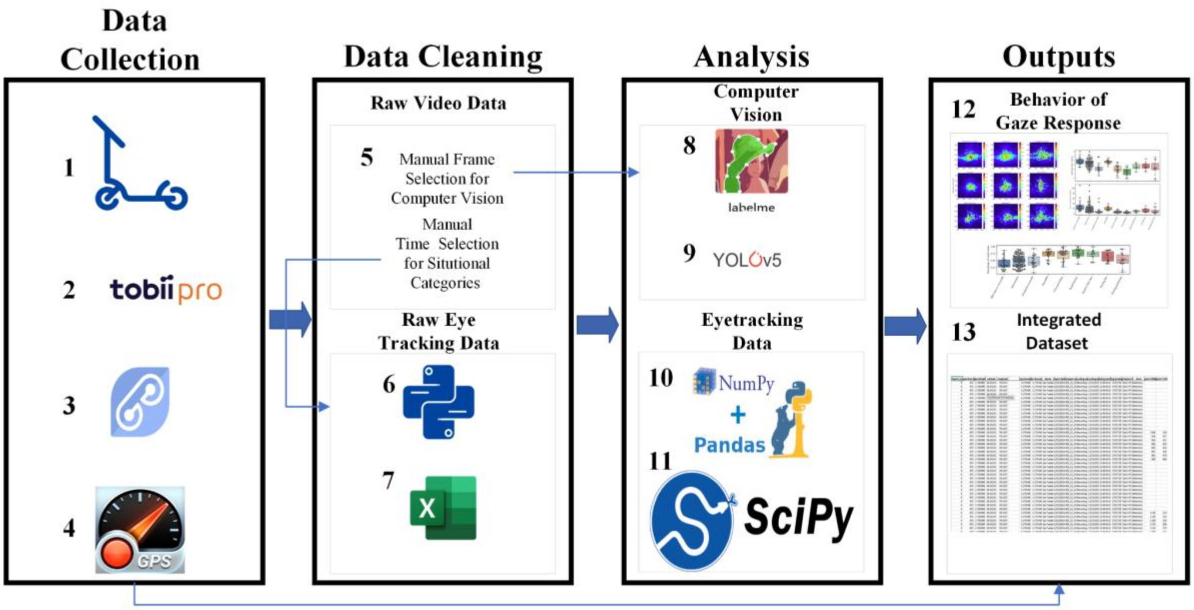

Figure 1: System architecture of data collection system. Data collection: 1) the Segway Ninebot E-scooter; 2) Tobii Pro Glasses 3; 3) Tobii Pro Glasses 3 controller; 4) SpeedTracker App. Data Cleaning: 5) Manual frame selection for computer vision and manual time selection for situational categories; 6) python for separation of eye-tracking data by time selection; 7) eye-tracking data separated by situation in separate excel files. Analysis: 8) labelme software for the creation of bounding boxes for computer vision training; 9) You Only Look Once (YOLO) software for computer vision training and analysis; 10) NumPy and Pandas python packages for entropy, variability, road center fixation analysis; 11) SciPy for optimization and added functionality to 10. Outputs: 12) results of gaze behavior analysis; 13) Integrated Dataset including speed, location, object detection, and gaze.

In the following subsections, the details of equipment selection, the data collection process, and the data cleaning process for each data stream, and the methods of analysis will be expounded upon.

3.1 Equipment Selection

The following table provides a description of the device and software being used within the cyber-physical data collection system we have developed. The table outlines how the devices and software are being used, the data type of data they output, the format in which they export the data, and their pertinent specifications:

Table 2: Data collection equipment specifications, outputs, use(s), and formatting.

Data Collection Devices and Softwares	Specifications	Outputs	Use Description	Format
Tobii Pro Glasses 3	Technique: Corneal reflection, dark pupil, stereo geometry Sample Rate: 100Hz	Time (Computer Timestamp) Gaze Point X Gaze Point Y Fixation Point X Fixation Point Y	The eyetracking glasses are being used to measure gaze, and fixation points (x-y coordinate format) of the e-scooter rider. The glasses are also being used to record video footage of what the e-scooter rider is seeing as they ride (mp4 format)	CSV
SpeedTracker App	Technique: GPS Sample Rate: 1Hz	Time (Mobile Phone Timestamp) Speed Longitude Latittude Elevation	The SpeedTracker app is being used to measure the location, speed, and elevation of the rider throughout the duration of the trip to be used alongside the eye-tracking data for further analysis	CSV

Segway Ninebot Kickscooter Max G30LP	Max Speed: 18.6 MPH Tire size:10” Range: 25 mi.	*	This is the e-scooter that was used to collect data	*
Tobbi Glasses 3 Controller	*	*	This software allows us to turn the recording apparatus that collects eye-tracking data within the Tobbi pro glasses 3 on and off.	*

3.2 Data Collection Process

As the research included in this thesis is intent upon creating a framework for data collection, we focused on collecting data from two participants. The first participant, the author of this thesis, collected 9 hours of e-scooter data. The second participant collected 1 hour of e-scooter data. We asked the participants to collect this data as a part of their normal daily routine. The data collected from the participants was useful in checking the effectiveness of the data collection system and contributing to the first comprehensive dataset for e-scooterists.

To collect the e-scooter data, we begin by setting up the Tobii Pro Glasses 3. We do this by connecting the glasses and the recording device to the Tobii Glasses 3 software via ethernet. We then enter the participant’s name and the applicable day to mark the file for future recall. Next, we calibrate the glasses using the calibration card provided in the Tobii Pro Glasses 3 kit, and then press the record button. The ethernet is then disconnected from the Glasses 3 assembly and the computer. The glasses memory device is then attached to the waist of the participant via belt clip. Now the glasses are operational and ready for use. Secondly, we must set up the SpeedTracker app. To do this, we simply open the app and press the “Start” button. The smartphone is then placed in the participant pocket. Now both the eye-tracking glasses and the SpeedTracker app are ready to collect data.

Next, the participant begins to ride the e-scooter on a self-chosen route throughout University of Virginia's grounds and downtown Charlottesville, Virginia; an example of a route taken while collecting data is shown in figure 2. Upon completion of a trip, the rider returns to the computer where Tobii Glasses 3 software is running. We then reconnect the Glasses 3 assembly to the computer via ethernet. Once the Glasses 3 software recognizes the glasses, we press stop on the recording. The Tobbi Glasses 3 software automatically saves the trip data to an SD card located within the memory device in the Glasses 3 assembly. This process was used to collect the 10 hours of e-scooter data, which was then used as a sample data set for the integration of the proposed system.

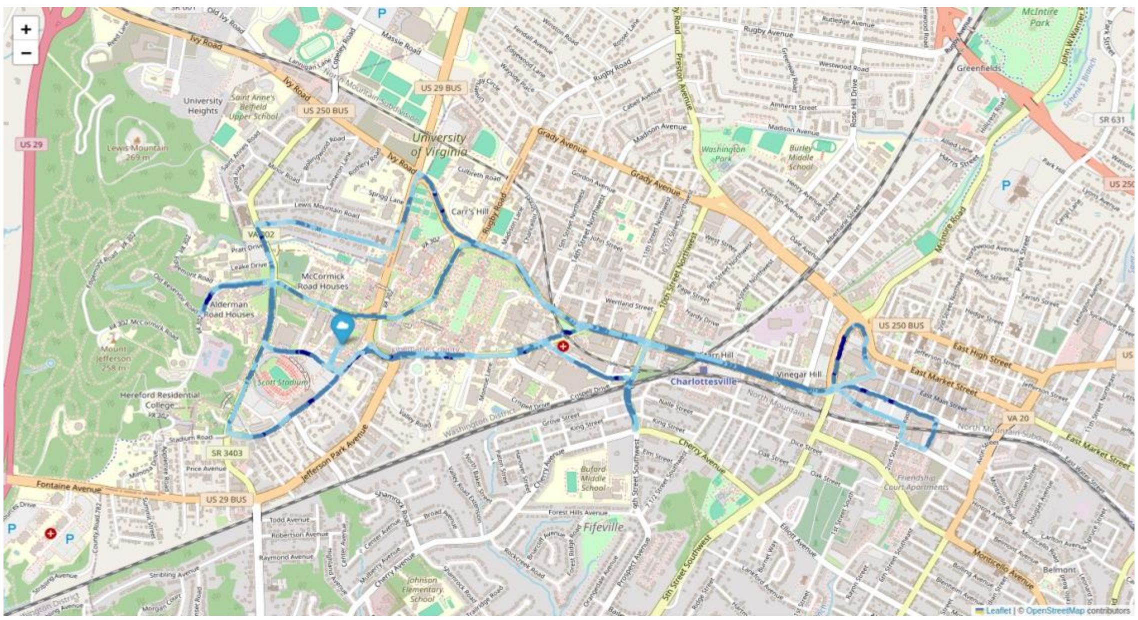

Figure 2: Map of route taken to collect data on December 11th, 2023, showing the location of data collection being downtown Charlottesville, VA and the University of Virginia campus.

3.3 Data Cleaning and Processing

We must clean the raw data that was collected during the rides to create a dataset that can be effectively operated on. However, each stream of data requires a different cleaning process.

3.3.1 Video Data

The video data is saved by the Tobbi Glasses 3 software as an mp4 file. Currently, we must manually watch and record the times wherein the video data corresponds with one of the situational categories. Next, we create and employ a python code that will use the times that we manually recorded to clip the full-length videos of the e-scooter trips into shorter videos, only showing the parts of the trips that correspond with one of the situational categories. Then, we aggregate the videos into separate files based on the category they apply to. Through computer vision developed specifically for e-scooters, as is expanded upon in section 3.4, we can conduct analysis with this data alone and alongside other data collected during the e-scooter rides.

3.3.2 Eye-tracking Data

The eye-tracking data is saved by the Tobii Glasses 3 software as a CSV that contains 54 columns. As a result, the excel file is rather large. However, we only need 5 of the columns (Timestamp, Gaze Point x, Gaze Point y, Fixation Point x, Fixation Point y). The other 49 columns are deleted, and the file is resaved. This data operation makes the excel file considerably smaller and, therefore, easier to conduct analysis on. Next, just as all the videos corresponding to a single situational category were grouped together into a file during the video data cleaning, all of the timestamps corresponding to those videos are manually entered into a CSV file. The next thing we must do is separate the eye-tracking data based on the timestamps collected from the videos. This allows us to view the eye-tracking data that corresponds with the established situational categories. However, this is not an easy task, as each CSV file for a trip of 1 hour or more can contain over 1 million rows of data due to the sample rate of the glasses being 100 Hz. To do this efficiently, we create and employ a python code that separates the CSV using timestamps we manually found within the video data. The video data and eye-tracking data operate on the same clock as they are collected on using the same device. The output of the python code is multiple CSVs that individually represent all the eye-tracking data collected for one situational category.

Gaze Entropy

As discussed in section 2.2.1, there are two measurements for gaze entropy, stationary gaze entropy (SGE) and gaze transition entropy (GTE). In the analysis of the collected data, both metrics are calculated. To do this, we create and employ python code to calculate both SGE and GTE. SGE is based on the predictability of fixation locations and can be a proxy for gaze dispersion [63]. If we assign fixation point to spatial bins (p_i), calculating SGE is as follows:

$$SGE = -\sum_{i=1}^{n} p_i \log_2 p_i$$

Gaze transition entropy (GTE) is a measure of the predictability of a future fixation location based on the current fixation location [63]. For a sequence of transitions between different spatial bins of i and j along with a probability of p_{ij}, the GTE can be calculated as follows:

$$GTE = -\sum_{i=1}^{n} p_i \sum_{j=1}^{n} p_{ij} \log_2 p_{ij}$$

Percentage of Road Center Fixations:

To conduct an analysis of road center fixations with the situational categories, we developed a Python code to automatically identify and quantify fixations within the eye-tracking data that are directed towards the center of the road. This code utilizes algorithms to process the spatial distribution of fixation points relative to the road center. The main mathematical equation behind this code is Euclidean distance as shown in equation 3, which calculates the Euclidean distance between fixation point (x_f, y_f) and the road center (x_c, y_c) in a two-dimensional Cartesian coordinate system.

$$\text{Euclidean Distance} = \sqrt{(x_f - x_c)^2 + (y_f - y_c)^2}$$

After determining whether the gaze points are located in the center of the road, the percentage of such gaze points is calculated to provide the analysis metric of percentage of road center fixations (PRC). This meaningful metric is indicative of rider focus and engagement with the surrounding environment.

3.4 Road Object Detection through Computer Vision Algorithms

The computer vision capability used in this study was developed in a collaborative effort in an earlier study testing different object detection algorithms and measuring their efficiency and effectiveness. Sections 3.4.1-3.4.3 are excerpts from the methodology section of this study.

3.4.1 Video Dataset

The traffic scene portion of the dataset being created within this study is also collected with Tobii Pro Glasses 3. The collection route spans from the University of Virginia campus to the urban area of Charlottesville, VA, USA. To ensure a diverse set of images for robust model performance, the dataset is collected under natural lighting conditions. Image frames are extracted from the recorded videos taken by the Tobii Pro Glasses. Subsequently, data cleaning procedures are applied to remove low-quality images and those lacking relevant objects of interest.

The meticulously curated dataset is then labeled by trained personnel with bounding boxes around traffic objects in the images using the LabelMe tool (https://github.com/labelmeai/labelme). This study focuses on 11 specific objects: i.e., “person”, “car”, “truck”, “bus”, “traffic light”, “fire hydrant”, “stop sign”, “bench”, and “scooter”. The resulting annotations, stored in JavaScript Object Notation (JSON) file format, then undergo visualization and double-checking by experts to ensure annotation accuracy and quality. This meticulous process results in a dataset comprising 2013 images covering 11 traffic object classes, with a total of 11,011 bounding box annotations, publicly accessible in the Zenodo repository (https://zenodo.org/records/10578641). This dataset will be continually updated with additional traffic scene images collected and labeled for future experiments.

3.4.2 Identifying Optimal YOLO Object Detectors

In this particular study, five versions of YOLO object detectors—YOLOv3, YOLOv5, YOLOv6, YOLOv7, and YOLOv8—are selected to develop traffic object detection models for the e-scooter dataset. Fig. 3 illustrates the modeling pipeline for traffic object detection, starting from data preparation to model training. The conversion process from the original image annotations in JSON format (from LabelMe) to YOLO format labels is a crucial initial step, ensuring compatibility with the YOLO training framework. After converting the formats, the dataset is then randomly divided into training, validation, and test subsets. This division follows a partition ratio of 60%, 20%, and 20%, corresponding to 1207, 402, and 404 images for each subset, respectively.

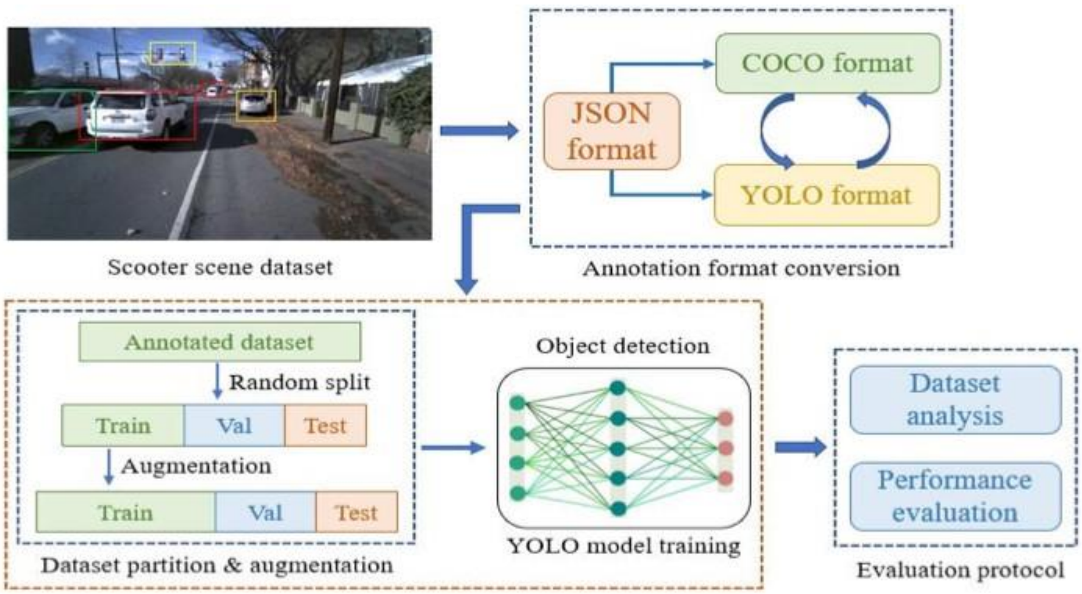

Figure 3: The proposed pipeline of object detection by YOLO object detectors.

To enhance the training process, all YOLO object detection models undergo a phase of transfer learning, as described by Weiss et al. (2016) [79], which involves refining the preexisting weights acquired from training on the COCO dataset, as detailed by Lin et al. (2014) [80]. The original images are adjusted to a uniform size of 640 × 640 pixels to meet the input requirements of the YOLO architecture. These models are then trained using a batch size of 16 for 100 epochs within the PyTorch framework (version 2.1.2) [81]. To optimize the learning rate throughout the training period, a cosine annealing strategy is applied across all YOLO models, a method proposed by He et al. (2019) [51]. The entire process of model training and evaluation is carried out on a server running Ubuntu 20.04, which is configured with an AMD 7502 32-core processor (128 GB of RAM) and dual GeForce RTX 3090Ti GPUs (each with 24 GB of GDDR6X RAM).

3.4.3 Performance Evaluation Metrics

The assessment of YOLO object detectors for identifying e-scooters involved evaluating detection accuracy, inference speeds, and model complexity. The quantity of model parameters indicates the complexity of the model, an important aspect for deploying models in practical settings. Generally, models with more parameters require additional memory for deployment and affect both computational costs and inference speeds. Detection accuracy, a crucial metric for object detection highlighted by Padilla et al. (2020), encompasses precision (P), recall (R), and mean average precision (mAP, notably mAP@0.5), with mAP being the key indicator for gauging the performance of object detectors in multi-class scenarios. Computational cost and inference time are assessed by metrics like floating point operations (FLOPs), which quantify the computational effort required to process a single instance, with FLOPs calculation facilitated by the THOP library. Inference time, the time it takes for a model to predict outcomes on an input image, is crucial for applications demanding real-time processing. This time was measured as the mean duration needed to analyze all images in the test dataset.

4. Results and Discussion

In this section we will be testing the novel framework created to collect e-scooter rider behavioral data by analyzing the gaze entropy attributes of the data. We will test the quality of the system by completing entropy analysis on data sets that have been selected from the larger data set based on metrics obtained through computer vision. Displaying the data collection system's ability to acquire, organize, and provide data that can be analyzed in a complex multidimensional manner will prove the merits of the system.

4.1 Computer Vision

The results and discussion for the computer vision portion of this study (sections 4.1.1-4.1.2) are a part of the collaborative study referenced in section 3.4.

4.1.1 Performance of YOLO models

Table 3 summarizes the performance of various YOLO models on a test dataset for object detection, emphasizing their effectiveness in identifying 11 classes of traffic-related objects. Notably, the table includes different versions of YOLOv3—such as YOLOv3-tiny and YOLOv3, which incorporate feature pyramid networks (FPN), and YOLOv3-SPP, which utilizes spatial pyramid pooling (SPP) [82]. These models display remarkable accuracy levels, with mean average precision (mAP) at a threshold of 0.5 ranging from 27.4% for YOLOv7-E6E to a high of 86.8% for YOLOv5s. While most models (excluding YOLOv7-W6, YOLOv7-E6, YOLOv7-D6, and YOLOv7-E6E, which fall below 70%) achieve mAP@0.5 accuracies between 72.1% and 86.8%, with six models surpassing 85%. The lower performance of YOLOv7 variants is attributed to overfitting, suggesting the need for further exploration of solutions like data augmentation and generation techniques to enhance future model performance. The table also reveals that the majority of YOLO models tend to have higher precision than recall, indicating effective object detection capabilities. However, challenges remain in detecting smaller objects at a distance from the e-scooter, leading to missed detections and lower recall rates.

Model complexity and inference times are important for real-world application, especially in scenarios with limited resources like e-scooters. Fig. 4 displays the correlation between GFLOPs and inference times against the total number of parameters across all evaluated YOLO detectors, revealing a linear augmentation in GFLOPs and inference times as model parameters increase. YOLOv8x stands out for having the highest GFLOPs and the longest inference time at 29.5 milliseconds. Conversely, YOLOv5 variants—specifically YOLOv5n and YOLOv5s—and YOLOv3-tiny are highlighted for their superior computational efficiency and swift inference times (under 5 milliseconds). Additionally, Fig. 4 also illustrates model inference times versus mAP@0.5, indicating potential compromises in choosing models based on accuracy versus inference speed, where increased accuracy is often linked with longer inference durations. Despite these variances, all tested YOLO detectors are capable of real-time object detection, achieving processing rates of dozens or even hundreds of frames per second. Notably, YOLOv5 and YOLOv8 models exemplify an optimal balance between accuracy and efficiency. It is crucial to acknowledge these assessments are performed on high-end computing setups, with the performance on embedded systems yet to be determined.

Table 3: Object detection performance of 22 YOLO detectors on the testing dataset.

Index	YOLO models	Precision	Recall	mAP@0.5
1	YOLOv3-tiny	0.747	0.680	0.721
2	YOLOv3-v3	0.854	0.842	0.857
3	YOLOv3-SPP	0.841	0.844	0.855
	Average	0.814	0.789	0.811
4	YOLOv5n	0.673	0.789	0.797
5	YOLOv5s	0.912	0.812	0.868
6	YOLOv5m	0.855	0.812	0.849
7	YOLOv5l	0.871	0.850	0.866
8	YOLOv5x	0.826	0.841	0.846
	Average	0.827	0.821	0.845
9	YOLOv6n	0.832	0.821	0.841
10	YOLOv6s	0.822	0.814	0.841
11	YOLOv6m	0.842	0.833	0.857
	Average	0.832	0.823	0.846
12	YOLOv7	0.830	0.802	0.808
13	YOLOv7x	0.857	0.876	0.862
14	YOLOv7-W6	0.705	0.517	0.583
15	YOLOv7-E6	0.680	0.543	0.601
16	YOLOv7-D6	0.478	0.516	0.470
17	YOLOv7-E6E	0.561	0.241	0.274
	Average	0.509	0.491	0.457
18	YOLOv8n	0.838	0.736	0.804
19	YOLOv8s	0.841	0.784	0.818
20	YOLOv8m	0.842	0.767	0.795
21	YOLOv8l	0.869	0.752	0.797
22	YOLOv8x	0.843	0.732	0.809
	Average	0.847	0.754	0.805

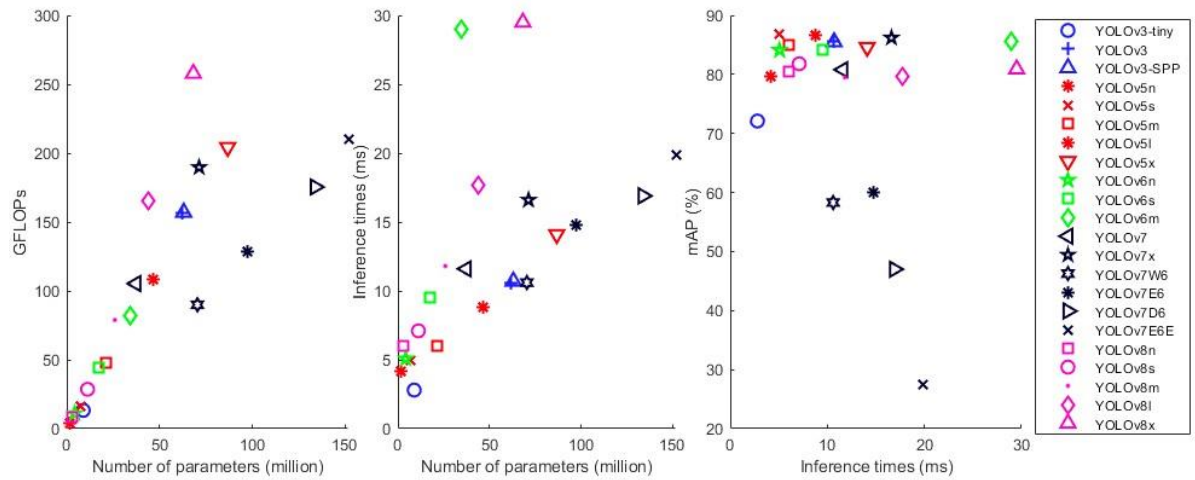

Figure 4: presents illustrative examples of image predictions made by YOLOv5s, showing the model's capability to generate visually accurate predictions across a range of scenarios, including those with diverse and cluttered backgrounds typical of densely populated urban environments.

Figure 5: Examples of traffic scene images with predicted bounding boxes.

These outcomes collectively highlight the effectiveness of the chosen YOLO object detectors in accurately identifying multiple classes of traffic objects in the context of e-scooter use. The ability of these detectors to maintain high accuracy in complex urban settings underscores their potential utility in enhancing the safety and navigational efficacy of e-scooters, further demonstrating their contribution to the field of machine vision-based safety systems for micro-mobility solutions.

4.1.2 Class-wise performance of selected YOLO models

In this subsection, in-depth analysis of detection accuracies for specific traffic object classes is conducted, focusing solely on YOLOv5s and YOLOv8s due to limitations in space, as detailed in Table 4. The accuracy in detecting traffic objects can be influenced by various factors, including the quantity and dimensions of bounding boxes, class variability, and similarities among object classes. Within the testing dataset, which included 2,160 instances, the "car" category has the most annotations (1,408) and shows high mAP@0.5, achieving 92.4% for YOLOv5s and 92.1% for YOLOv8s. Other categories such as "bicycle", "motorcycle", "bus", "fire hydrant", and "stop sign" also yield mAP@0.5 scores above 90%. Despite having nearly 200 bounding box annotations, the "person" class is detected with comparatively lower accuracy by both YOLOv5s and YOLOv8s, with mAP@0.5 scores of 79.7% and 77.2%, respectively. This lower accuracy is likely due to the small and blurred appearances of persons in the images, complicating accurate localization. Furthermore, the "bench" and "scooter" classes are particularly challenging for both detectors, with mAP scores falling below 80%. The limited number of annotations for these classes may contribute to their lower detection accuracies. To improve detection accuracy for these more challenging classes, incorporating a greater variety of training samples and refining training methodologies may be beneficial.

Table 4: Class-wise performance of YOLOv5s and YOLOv8s. P, R and mAP represent precision, recall and mean average precision, respectively.

Index	Object Class	# Instances	YOLOv5s-P	YOLOv5s-R	YOLOv5s-mAP@0.5	YOLOv8s-P	YOLOv8s-R	YOLOv8s-mAP@0.5
1	person	224	0.576	0.691	0.797	0.821	0.772	0.772
2	bicycle	14	0.902	0.829	0.946	0.879	0.766	0.766
3	car	148	0.907	0.926	0.944	0.93	0.902	0.902
4	motorcycle	11	0.867	0.818	0.906	0.842	0.965	0.965
5	bus	34	0.977	0.971	0.989	0.949	0.963	0.963
6	truck	184	0.959	0.864	0.895	0.873	0.887	0.887
7	traffic light	146	0.935	0.783	0.898	0.884	0.817	0.817
8	fire hydrant	14	0.918	0.844	0.902	0.813	0.814	0.814
9	stop sign	1	1	0.79	0.95	0.901	0.905	0.905
10	bench	18	0.84	0.778	0.776	0.759	0.692	0.692

All		2160	0.912	0.812	0.868	0.851	0.784	0.815

4.2 Analyzing Gaze and Behavioral Data

In the analysis of the eye-tracking data, we begin by visualizing the data through various descriptive techniques. Each visualization will be based on the 9 situational categories that we identified in section 1.4.1, the first of which being a heat map consisting of the participants field of view, which reveals the distribution of the participant's gaze points.

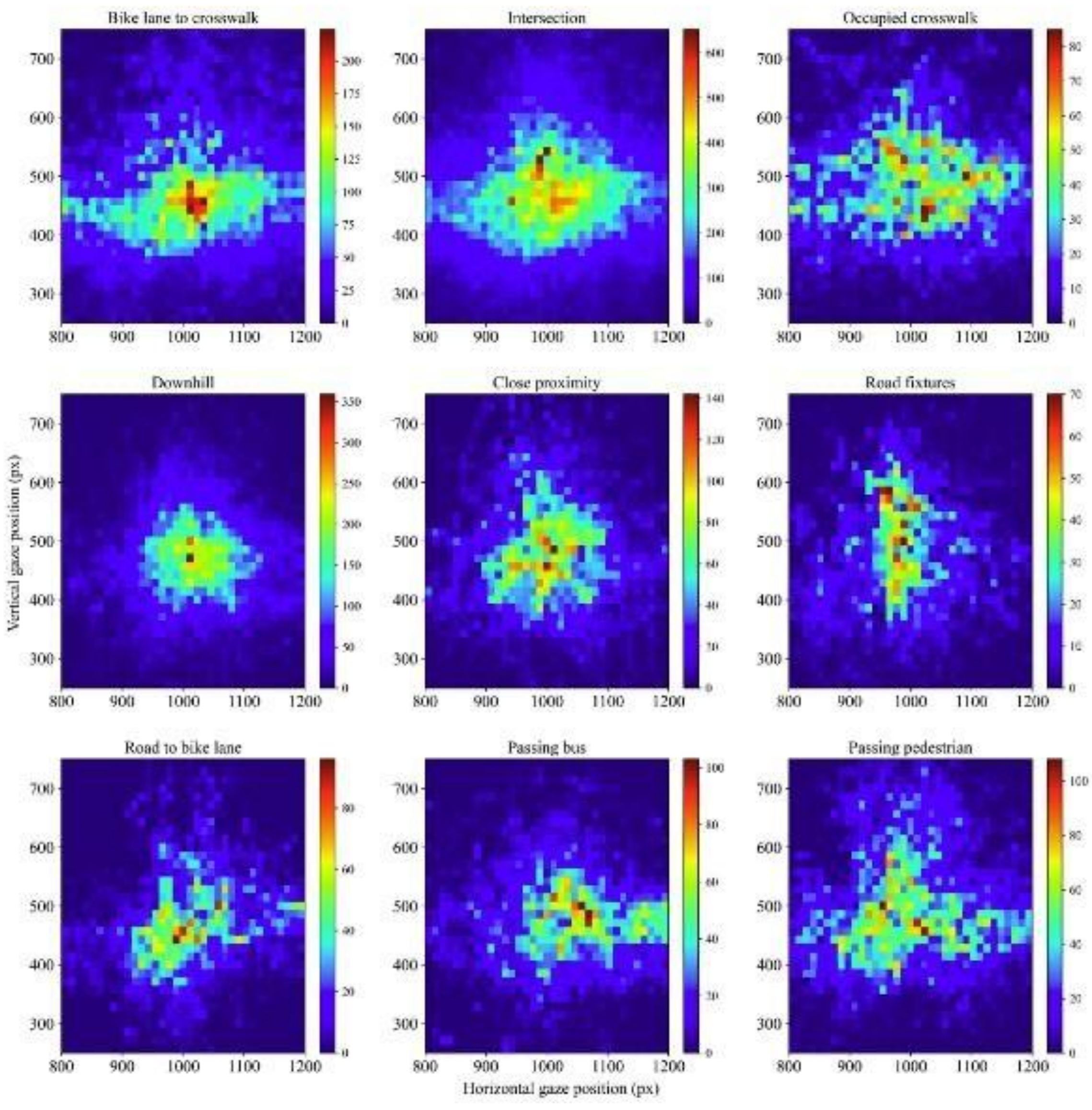

Figure 6: Gaze density heat map for the different situational categories identified in table 1.

In figure 6, we visually observe that the distribution and density of gaze points vary amongst the different situational categories. For example, in “Intersection”, the participant’s gaze points are far more equally distributed when compared to “Bike lane to crosswalk”. The differences in gaze distribution indicates that the areas of the participants' field of view that require attention most frequently changes depending on the situation that they are in. To explore these differences more deeply, we extracted some descriptive statistics from each situational dataset as can be seen in Table 5. In our descriptive analysis we determined the mean, median, range, and standard deviation of the x-y coordinates in each data set. For reference, the statistical measurements included in Table 5 are in the unit of pixels. The field of view for this study was 1920x1080 pixels.

Table 5: Descriptive Statistics for Gaze Points (X and Y)

Situational Category	Coordinate (X or Y)	Statistical Measurements (pixels)
Bike lane to Crosswalk	x	Mean: 960.66 Standard Deviation: 288.59 Range: 2101.00 Median: 994.00 Interquartile Range: 301.00
	y	Mean: 482.82 Standard Deviation: 141.14 Range: 1977.00 Median: 465.00 Interquartile Range: 106.00
Close Proximity	x	Mean: 983.85 Standard Deviation: 164.84 Range: 1935.00 Median: 997.00 Interquartile Range: 115.00
	y	Mean: 488.35 Standard Deviation: 108.12 Range: 1741.00 Median: 478.00 Interquartile Range: 112.00
High Speed Downhill	x	Mean: 1008.18 Standard Deviation: 155.58 Range: 2043.00 Median: 1016.00

		Interquartile Range: 113.00
	y	Mean: 475.82 Standard Deviation: 97.88 Range: 1974.00 Median: 471.00 Interquartile Range: 94.00
Intersection	x	Mean: 967.55 Standard Deviation: 246.32 Range: 2103.00 Median: 990.00 Interquartile Range: 228.00
	y	Mean: 493.46 Standard Deviation: 148.39 Range: 1993.00 Median: 472.00 Interquartile Range: 117.0
Occupied Crosswalk	x	Mean: 946.73 Standard Deviation: 217.52 Range: 2066.00 Median: 981.00 Interquartile Range: 247.00
	y	Mean: 492.63 Standard Deviation: 114.34 Range: 1547.00 Median: 485.00 Interquartile Range: 109.00
Passing Bus	x	Mean: 1010.86 Standard Deviation: 268.07 Range: 1986.00 Median: 1042.00 Interquartile Range: 201.00
	y	Mean: 482.69 Standard Deviation: 103.17 Range: 1929.00 Median: 479.00 Interquartile Range: 93.00
Passing Pedestrians	x	Mean: 987.03 Standard Deviation: 152.98 Range: 1948.00 Median: 986.00

		Interquartile: 151.00
	y	Mean: 493.02 Standard Deviation: 95.31 Range: 1487.00 Median: 478.00 Interquartile: 110.00
Road Fixtures	x	Mean: 979.85 Standard Deviation: 141.01 Range: 1727.00 Median: 980.00 Interquartile: 89.00
	y	Mean: 481.96 Standard Deviation: 118.92 Range: 1561.00 Median: 481.00 Interquartile: 139.00
Road to Bike Lane	x	Mean: 1004.44 Standard Deviation: 204.36 Range: 2098.00 Median: 1011.00 Interquartile Range: 122.00
	y	Mean: 472.96 Standard Deviation: 80.15 Range: 695.00 Median: 467.00 Interquartile Range: 94.00

The descriptive statistics in Table 5 provides insights on how this participant's gaze responded to different situations. For example, through the standard deviation metric, we observe that in the horizontal direction of gaze movement (x), the "Intersection" eye-tracking data is spread out further from its mean, with a standard deviation of 246.32 pixels, than in "High Speed Downhill" data, with a standard deviation of 155.57 pixels. With all the situational categories having mean x-y observations close to the center of the field of view (960, 540), this indicates that in the "Intersection" situation the participant's attention is drawn away from the center of the field of view more often than when the participant is in the "High Speed Downhill" situation. In another example, we observe that the range of gaze points where the majority of observations lie (interquartile range) for the vertical direction of gaze movement (y) is less in the "Road to Bike Lane" category, at 94.00 pixels, when compared to the "Road Fixtures" category, at 139.00

pixels. This indicates that the majority of the participants' attention is focused in a smaller section of the total field of view when traveling from the road to a bike lane in comparison to when the participant is traveling over or around a road fixture.

Additionally, we chose to analyze the data using metrics that are familiar to eye-tracking data such as mean fixations per second, mean fixation length, percentage of road center gaze, gaze variability, stationary gaze entropy, and gaze transition entropy. Herein, we will be visualizing the data by these metrics through the use of boxplots. The boxplots are constructed using each individual occurrence of a situation within the situational categories defined in section 1.4.1. Each point displayed represents an occurrence of that situation.

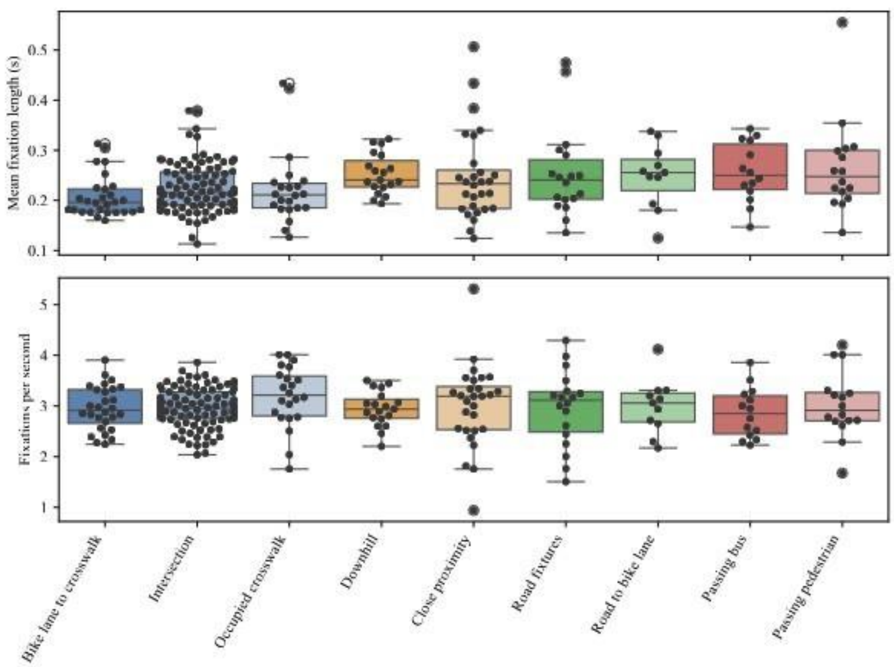

Figure 7: Mean fixation length (s) and number of fixations per second for all occurrences of events matching the situational categories listed in table 1.

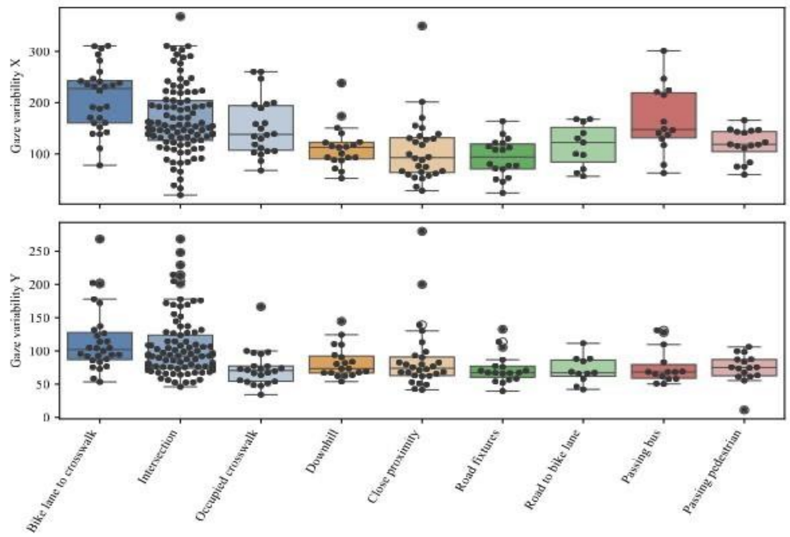

Figure 8: Gaze variability (pixels) for the x and y coordinate respectively of all occurrences of events matching the situational categories listed in table 1.

In figure 8, we see that in the horizontal (x-direction) eye-tracking data, “bike lane to crosswalk” varies the most at a mean of 209.96 pixels, while “road fixtures” varies the least at a mean of 92.65 pixels. In the vertical (y-direction) eye-tracking data, “bike lane to crosswalk” maintains the highest variability at a mean of 115.82 pixels, while “road fixtures” maintains the lowest at a mean of 73.72 pixels.

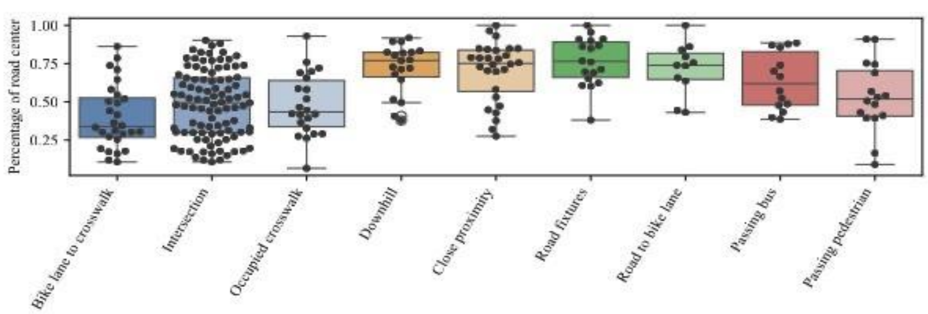

Figure 9: Percentage of road center fixations (%) for all occurrences of events matching the situational categories listed in table 1.

In figure 9, we see that “road fixtures” have the highest mean percentage of road center (PRC) fixations at 76% while, “bike lane to crosswalk” has the lowest mean PRC at 40%.

In figure 10, we see that for stationary gaze entropy the “bike lane to crosswalk” situational category has the highest mean score at 5.77 bits, while “road fixtures” has the lowest mean score at 3.39 bits. For gaze transition entropy “bike lane to crosswalk” also has the highest mean score at 28.68 bits, while “road fixtures” has the lowest score at 5.17 bits.

In summary, based on the results of the analysis of the eye-tracking data for the participant in this study, we successfully validate our data collection system as the metrics collected across the 9 situations included in this study are consistently aligned with theoretical frameworks on road user behavior. For example, in the heatmap, we see that the gaze points for the “Intersection” situation are distributed far more broadly than “Downhill” as one might expect the attention to be more distributed when at an intersection to ensure safe navigation, whereas one might expect an e-scooterist traveling downhill, most likely at a high speed, to be focused on what is directly in front of them to avoid obstacles. When we look at the descriptive statistics for the eye-tracking data, we see similar trends backed by numbers. Within the standard deviation metric, we observe that in the “intersection” category the standard deviation of the gaze points in both the x and y directions are higher than situations like “downhill” and “road fixtures”. This indicates that there is an increased variability in attention allocation while in an intersection compared to other situations.

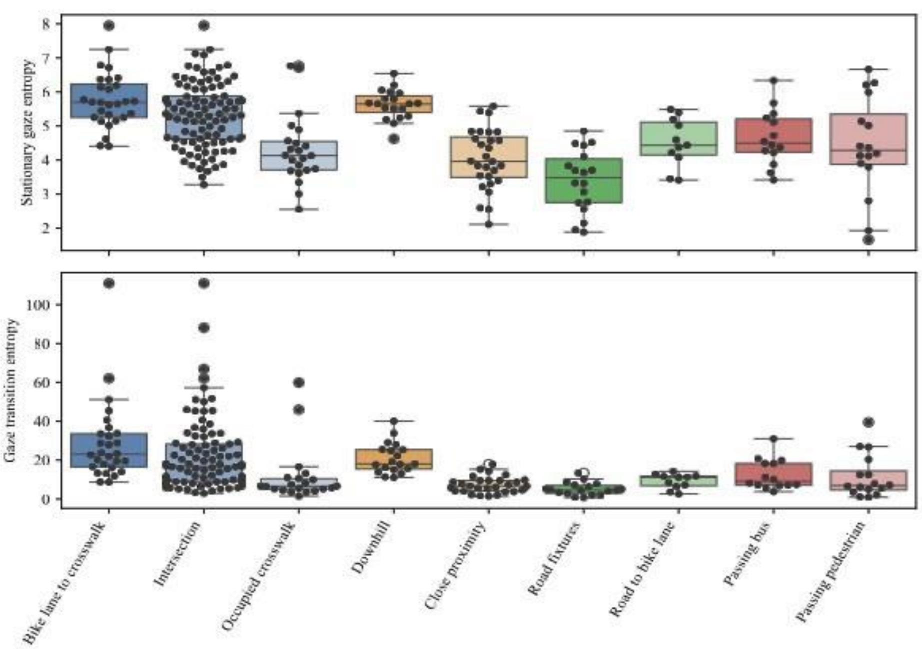

Figure 10: Gaze transition entropy and stationary gaze entropy (bits) for all occurrences of events matching the situational categories listed in table 1.

When we consider the PRC metric, we observe that “road fixtures” has the highest mean PRC at 76%, which when compared with “bike lane to crosswalk”, at 40%, indicates that this participant consistently was operating the e-scooter in an attentive manner and demonstrated a high level of focus on avoiding or navigating the obstacle in their path of travel. As the obstacle that an e-scooterist is attempting to avoid or navigate is typically center-of-the-road in relation to the rider, one would expect the rider's attention to be primarily on the obstacle.

The distribution of gaze points and the frequency of specific fixations only help inform us on how attention is allocated. Other metrics like entropy allow us to determine how focused this attention is. In the data we collected from our participants we see that “bike lane to crosswalk” has both the highest mean stationary gaze entropy score and gaze transition entropy score, indicating a lack of predictability in fixations and therefore high cognitive loading. The bike lane to crosswalk category typically occurs in intersections and therefore includes two cognitively complex tasks, one to assess and navigate the intersection appropriately, but also to switch between two different road infrastructures.

5. Limitations & Future Work

The most prevalent limitation of the data collection framework developed in this thesis is that it was designed to specifically study e-scooterists' behavioral attributes in relation to different environments. Future studies relating to e-scooters, but not necessarily the behavioral aspect of the riders, may require a system with different components to capture the data needed to address the question at hand. However, even in light of this limitation, this thesis provides the fundamental framework and process for how to develop a system that fits any research question in that both the system architecture and coding included herein can be adapted to address other research questions relating to e-scooters.

Beyond the framework for data collection presented in the thesis, the study we conducted to test the framework has limitations as well. However, the recognition of these limitations helps to inform future work. The first limitation is that we cannot make generalizations or establish causal relationships based on how e-scooterists behave in different situations due to the small sample size of our study. The lack of our ability to establish cause and effect with the study herein serves as a threat to internal validity. However, this study is setting the groundwork for future researchers to conduct larger studies with more participants that would overcome sample size issue within. Future research could involve longitudinal studies to observe e-scooterists' behavior over extended periods, comparative studies across various cities to assess the impact of infrastructure and regulatory environments, safety assessments to identify risk factors and mitigate hazards associated with e-scooter use, and environmental studies to determine the impacts on sustainability that e-scooter use presents.

Furthermore, there are threats to the external validity of the created framework. The current framework requires participants to wear eye-tracking glasses which may change how the participant behaves if they have not worn glasses in the past or feel differently knowing that the glasses are recording their gaze movements. Additionally, the current framework does not provide for conducting studies during the nighttime as the low light exposure provides low quality images for the computer vision algorithm to process. This issue can be resolved by adding higher quality cameras to the e-scooter for environmental feature extraction through computer vision. Lidar can also be added to the e-scooter to more accurately interpret the environment around the participant.

Another limitation that yields way to potential future work is the crucial limitation within the length of use of the system. Currently, the data collection system has a battery life of approximately 1.5 hours. Additionally, the batteries for the Tobii Pro Glasses, the smartphone, and the e-scooter itself must be charged separately. Future work might include integrating the data collection system within the e-scooter itself, possibly utilizing an NVIDIA Jetson as the computing power. This integration could streamline data collection processes, reduce the need for additional external devices, and enhance the portability and usability of the system. By embedding the data collection system within the e-scooter, future researchers could overcome limitations related to battery life and ease of recharging, ultimately improving the efficiency and effectiveness of data collection in real-world settings.

By addressing these limitations and incorporating these future research directions, future studies can build upon the foundation laid by this thesis, further advancing our understanding of e-scooter behavior and enhancing the efficacy of data collection methodologies in micro-mobility research.

6. Conclusion:

Based on the work presented in this thesis, we see that adding computational and sensory capabilities to the e-scooter can help improve our understanding of rider behavior. The main objective of the work included in this thesis was to develop and validate an e-scooter data collection system that is customized to collect data to help future researchers learn more about the behaviors of e-scooterists. Through the development of our system's computer vision and data analysis capabilities in conjunction with our test of the system with eye-tracking data collected from two participants, we have achieved this goal. A Secondary goal of the work in this thesis was to provide the first comprehensive dataset for e-scooter research. Through combining gaze, GPS, and object detection data in a time-synced manner, we have also achieved this goal.

Additionally, the system developed in this thesis is suitable to be tailored to other micro-mobility modes of transportation. With a continued effort to understand e-scooters and other micro-mobility options, the safety and comfort of riders and other road users alike are likely to increase. Furthermore, the findings of this research have significant implications for future studies in the field. By highlighting potential avenues for further research, such as investigating the impact of infrastructure design on e-scooter usage patterns or exploring the effectiveness of interventions aimed at promoting safer riding practices, we can continue to advance our understanding of micro-mobility systems and inform evidence-based policies and interventions.

Moreover, the technological advancements achieved through the development and implementation of the data collection system have broader applications beyond academic research. The integration of e-scooter data collection systems with smart city initiatives could revolutionize urban mobility, optimize traffic flow, and contribute to sustainability goals. Furthermore, collaborative partnerships between academia, industry, and government agencies are essential for translating research findings into real-world solutions and fostering innovation in the field of micro-mobility.

Finally, this thesis represents a significant contribution to the study of e-scooter behavior and data collection methodologies. By addressing key research objectives, providing valuable insights, and laying the foundation for future studies, this research has the potential to drive positive change in the realm of urban transportation, ultimately improving the safety, efficiency, and sustainability of e-scooter and micro-mobility systems.

References

[1] "Shared Micromobility in 2022, NACTO".

[2] S. Shaheen and N. Chan. "Mobility and the Sharing Economy: Potential to Overcome First- and Last-Mile Public Transit Connections". Federal Reserve Bank of St. Louis. Jan. 2016. 10.7922/g2862dn3.

[3] N. K. Namiri, H. Lui, T. Tangney, I. E. Allen, A. J. Cohen and B. N. Breyer, "Electric Scooter Injuries and Hospital Admissions in the United States, 2014-2018".

[4] H. Kleinertz, D. Ntalos, F. Hennes, J. Nüchtern, K. Frosch and D. M. Thiesen, "Accident Mechanisms and Injury Patterns in E-Scooter Users".

[5] H. Jeong, Z. Kang and Y. Liu. "Driver glance behaviors and scanning patterns: Applying static and dynamic glance measures to the analysis of curve driving with secondary tasks". Human factors and ergonomics in manufacturing (Print). vol. 29. no. 6. pp. 437-446. Aug. 2019. 10.1002/hfm.20798.

[6] S. Martin, S. Vora, K. Yuen and M. M. Trivedi. "Dynamics of Driver's Gaze: Explorations in Behavior Modeling and Maneuver Prediction". IEEE transactions on intelligent vehicles. vol. 3. no. 2. pp. 141-150. Jun. 2018. 10.1109/tiv.2018.2804160.

[7] J. Engström, E. Johansson and J. Östlund, "Effects of visual and cognitive load in real and simulated motorway driving".

[8] "Micromobility Products-Related Deaths, Injuries, and Hazard Patterns: 2017-2022". [9] "Injuries Using E-Scooters, E-Bikes and Hoverboards Jump 70% During the Past Four Years".

[10] "DOCKLESS Electric SCOOTER-RELATED INJURIES STUDY".

[11] M. Paudel, F. F. Yap and A. K. Bastola. "Safety assessment of personal mobility devices with different wheel size based on their dynamic stability performance". vol. 3. no. 4. pp. 227227. Jan. 2020. 10.1504/ijsdes.2020.112094.

[12] M. Toofany, S. Mohsenian, L. K. Shum, H. C. Chan and J. R. Brubacher. "Injury patterns and circumstances associated with electric scooter collisions: a scoping review". Injury prevention. vol. 27. no. 5. pp. 490-499. Mar. 2021. 10.1136/injuryprev-2020-044085. [13] "Comparing the risky behaviours of shared and private e-scooter and bicycle riders in downtown Brisbane, Australia.".

[14] E. Karpinski, E. Bayles, L. Daigle and D. Mantine. "Comparison of motor-vehicle involved e-scooter fatalities with other traffic fatalities". Journal of Safety Research. vol. 84. pp. 61-73.

Feb. 2023. 10.1016/j.jsr.2022.10.008.

[15] K. Zube et al.. "E-scooter driving under the acute influence of alcohol—a real-driving fitness study". International journal of legal medicine (Print). vol. 136. no. 5. pp. 1281-1290.

Feb. 2022. 10.1007/s00414-022-02792-3.

[16] https://doi.org/10.1016/j.iatssr.2020.08.006.

[17] M. D. Mura, S. Failla, N. Gori, A. Micucci and F. Paganelli. "E-Scooter Presence in Urban Areas: Are Consistent Rules, Paying Attention and Smooth Infrastructure Enough for Safety?".

Multidisciplinary Digital Publishing Institute. vol. 14. no. 21. pp. 14303-14303. Nov. 2022. 10.3390/su142114303.

[18] W. Zhang, R. Buehler, A. Broaddus and T. E. Sweeney. "What type of infrastructures do e-scooter riders prefer? A route choice model". Elsevier BV. vol. 94. pp. 102761-102761. May.

2021. 10.1016/j.trd.2021.102761.

[19] N. Zuniga-Garcia, N. R. Juri, K. A. Perrine and R. B. Machemehl. "E-scooters in urban infrastructure: Understanding sidewalk, bike lane, and roadway usage from trajectory data". Elsevier BV. vol. 9. no. 3. pp. 983-994. Sep. 2021. 10.1016/j.cstp.2021.04.004. [20] N. J. Klein, A. Brown and C. Thigpen. "Clutter and Compliance: Scooter Parking Interventions and Perceptions". vol. 3. no. 1. Jan. 2023. 10.16997/ats.1196.

[21] S. Cafiso, A. D. Graziano, V. Marchetta and G. Pappalardo. "Urban road pavements monitoring and assessment using bike and e-scooter as probe vehicles". Elsevier BV. vol. 16. pp. e00889-e00889. Jun. 2022. 10.1016/j.cscm.2022.e00889.

[22] J. D. Cano-Moreno, M. Islán, F. Blaya, R. D'Amato, J. A. Juanes and E. R. Soriano. "E-scooter Vibration Impact on Driver Comfort and Health". Springer Science+Business Media. vol. 9. no. 6. pp. 1023-1037. Jan. 2021. 10.1007/s42417-021-00280-3.

[23] N. Zuniga-Garcia, N. R. Juri, K. A. Perrine and R. B. Machemehl, "E-scooters in urban infrastructure: Understanding sidewalk, bike lane, and roadway usage from trajectory data". [24] D. L. Marques and M. C. Coelho. "A Literature Review of Emerging Research Needs for Micromobility—Integration through a Life Cycle Thinking Approach". Multidisciplinary Digital

Publishing Institute. vol. 2. no. 1. pp. 135-164. Feb. 2022. 10.3390/futuretransp2010008. [25] L. Zhang and J. Song. "The periodicity and initial evolution of micro-mobility systems: a case study of the docked bike-sharing system in New York City, USA". Springer Science+Business Media. vol. 14. no. 1. Jun. 2022. 10.1186/s12544-022-00549-y. [26] A. V. Prabu, R. Tian, S. Chien, L. Li, Y. Chen and R. Sherony, "Risk assessment and mitigation of e-scooter crashes with naturalistic driving data".

[27] K. Button, H. Frye and D. P. Reaves, "Economic regulation and E-scooter networks in the USA".

[28] S. Das, A. Hossain, M. A. Rahman, A. Sheykhfard and B. Kutela. "Case Study on the Traffic Collision Patterns of E-Scooter Riders". SAGE Publishing. Jul. 2023.
10.1177/03611981231185770.

[29] L. Kobayashi et al.. "The e-merging e-pidemic of e-scooters". Trauma surgery & acute care open. vol. 4. no. 1. pp. e000337-e000337. Jul. 2019. 10.1136/tsaco-2019-000337.

[30] A. Maiti, N. Vinayaga-Sureshkanth, M. Jadliwala, R. Wijewickrama and G. Griffin. "Impact of E-Scooters on Pedestrian Safety: A Field Study Using Pedestrian Crowd-Sensing". Mar. 2022.
10.1109/percomworkshops53856.2022.9767450.

[31] S. Gößling. "Integrating e-scooters in urban transportation: Problems, policies, and the prospect of system change". Elsevier BV. vol. 79. pp. 102230-102230. Feb. 2020.
10.1016/j.trd.2020.102230.

[32] M. Paudel and F. F. Yap. "Front steering design guidelines formulation for e-scooters considering the influence of sitting and standing riders on self-stability and safety performance".
SAGE Publishing. pp. 095440702199217-095440702199217. Feb. 2021.
10.1177/0954407021992176.

[33] A. I. Torre-Bastida, J. D. Ser, I. Laña, M. Ilardia, M. N. Bilbao and S. Campos-Cordobés. "Big Data for transportation and mobility: recent advances, trends and challenges". Iet Intelligent Transport Systems. vol. 12. no. 8. pp. 742-755. Jul. 2018. 10.1049/iet-its.2018.5188.

[34] "Challenges and Opportunities in Transportation Data".

[35] N. G. Harris, F. D. Simone and B. Condry. "A Comprehensive Analysis of Passenger Alighting and Boarding Rates". Springer Science+Business Media. vol. 8. no. 1. pp. 67-98. Mar.
2022. 10.1007/s40864-021-00161-8.
[36] K. Lee and I. N. Sener. "Emerging data for pedestrian and bicycle monitoring: Sources and applications". Elsevier BV. vol. 4. pp. 100095-100095. Mar. 2020. 10.1016/j.trip.2020.100095.
[37] M. McCue, "Exploring Stress Levels of E-Scooter Riders – New Jersey State Policy Lab".
[38] Q. Ma, H. Yang, A. Mayhue, Y. Sun, Z. Huang and Y. Ma. "E-Scooter safety: The riding risk analysis based on mobile sensing data". Elsevier BV. vol. 151. pp. 105954-105954. Mar.
2021. 10.1016/j.aap.2020.105954.
[39] T. Zhang et al.. "Physiological Measurements of Situation Awareness: A Systematic Review". SAGE Publishing. vol. 65. no. 5. pp. 737-758. Nov. 2020.
10.1177/0018720820969071.
[40] ec.europa.eu, (Accessed: 1 Apr. 2024).
[41] "Assessment of the safety benefits of vehicles' advanced driver assistance, connectivity and low level automation systems.".
[42] A. M. Svancara, W. J. Horrey, B. C. Tefft and M. A. Benson. "Potential Reduction in Crashes, Injuries and Deaths from Large-Scale Deployment of Advanced Driver Assistance Systems". AAA Foundation for Traffic Safety.. Sep. 2018.
[43] R. Utriainen, M. Pollanen and H. Liimatainen, "The Safety Potential of Lane Keeping Assistance and Possible Actions to Improve the Potential".
[44] R. Utriainen, M. Pöllänen and H. Liimatainen. "The Safety Potential of Lane Keeping Assistance and Possible Actions to Improve the Potential". Institute of Electrical and Electronics Engineers. vol. 5. no. 4. pp. 556-564. Dec. 2020. 10.1109/tiv.2020.2991962.
[45] "Vehicle Active Safety Monitoring Technology".
[46] E. Rosen, J. Källhammer, D. Eriksson, M. Nentwich, R. Fredriksson and K. Smith. "Pedestrian injury mitigation by autonomous braking". Elsevier BV. vol. 42. no. 6. pp. 19491957. Nov. 2010. 10.1016/j.aap.2010.05.018.
[47] D. Chang, "Automated Vehicles and Pedestrian Safety: Exploring the Promise and Limits of Pedestrian Detection".

[48] Y. Cai et al.. "YOLOv4-5D: An Effective and Efficient Object Detector for Autonomous Driving". Institute of Electrical and Electronics Engineers. vol. 70. pp. 1-13. Jan. 2021. 10.1109/tim.2021.3065438.

[49] H. Alai and R. Rajamani. "Multi-Stage Estimation Algorithm for Target Vehicle Trajectory Tracking with Applications to E-Scooter Protection*". Institute of Electrical and Electronics Engineers. pp. 1-12. Jan. 2023. 10.1109/tiv.2023.3307973.

[50] S. Ren, K. He, R. Girshick and J. Sun. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks". IEEE Computer Society. vol. 39. no. 6. pp. 11371149. Jun. 2017. 10.1109/tpami.2016.2577031.

[51] K. He, G. Gkioxari, P. Dollár and R. Girshick. "Mask R-CNN". Oct. 2017. 10.1109/iccv.2017.322.

[52] J. Redmon, S. Divvala, R. Girshick and A. Farhadi. "You Only Look Once: Unified, RealTime Object Detection". Jun. 2016. https://doi.org/10.48550/arXiv.1804.02767.

[53] A. Bochkovskiy, C. Wang and H. Liao. "YOLOv4: Optimal Speed and Accuracy of Object Detection". Apr. 2020.

[54] "E-Scooter safety: The riding risk analysis based on mobile sensing data.".

[55] A. Ledezma, V. Zamora, Ó. Sipele, M. P. Sesmero and A. Sanchis. "Implementing a Gaze Tracking Algorithm for Improving Advanced Driver Assistance Systems". Electronics. vol. 10.
no. 12. pp. 1480-1480. Jun. 2021. 10.3390/electronics10121480.

[56] I. Dua, T. A. John, R. Gupta and C. V. Jawahar. "DGAZE: Driver Gaze Mapping on Road". Oct. 2020. 10.1109/iros45743.2020.9341782.

[57] A. Fernández, R. Usamentiaga, J. L. Carús and R. Casado. "Driver Distraction Using Visual-Based Sensors and Algorithms". Sensors (Basel). vol. 16. no. 11. pp. 1805-1805. Oct. 2016. 10.3390/s16111805.

[58] A. Sahayadhas, K. Sundaraj and M. Murugappan. "Detecting Driver Drowsiness Based on Sensors: A Review". Sensors (Basel). vol. 12. no. 12. pp. 16937-16953. Dec. 2012. 10.3390/s121216937.

[59] M. Chakraborty and A. N. H. Aoyon. "Implementation of Computer Vision to detect driver fatigue or drowsiness to reduce the chances of vehicle accident". Apr. 2014. 10.1109/iceeict.2014.6919054.

[60] T. Dang, S. Bhattacharya and J. Crumbley. "A Review Study on the Use of Oculometry in the Assessment of Driver Cognitive States". Mar. 2021.
10.1109/southeastcon45413.2021.9401905.

[61] P. K. Sharma and P. Chakraborty. "A Review of Driver Gaze Estimation and Application in Gaze Behavior Understanding". arXiv (Cornell University). Jul. 2023.
10.48550/arxiv.2307.01470.

[62] I. Kotseruba and J. K. Tsotsos, "Attention for Vision-Based Assistive and Automated Driving: A Review of Algorithms and Datasets".

[63] B. Shiferaw, L. A. Downey and D. P. Crewther. "A review of gaze entropy as a measure of visual scanning efficiency". Neuroscience & Biobehavioral Reviews. vol. 96. pp. 353-366. Jan. 2019. 10.1016/j.neubiorev.2018.12.007.

[64] "Examining drivers' eye glance patterns during distracted driving: Insights from scanning randomness and glance transition matrix.".

[65] F. Schieber and J. Gilland, "Visual Entropy Metric Reveals Differences in Drivers' Eye Gaze Complexity across Variations in Age and Subsidiary Task Load".

[66] P. K. Sharma and P. Chakraborty, "A Review of Driver Gaze Estimation and Application in Gaze Behavior Understanding".

[67] M. F. Land and D. N. Lee. "Where we look when we steer". Nature Portfolio. vol. 369. no. 6483. pp. 742-744. Jun. 1994. 10.1038/369742a0.

[68] Y. Ren, X. Li, X. Zheng, Z. Li and Q. Zhao. "Analysis of Drivers' Eye-Movement Characteristics When Driving around Curves". Discrete Dynamics in Nature and Society. vol. 2015. pp. 1-10. Jan. 2015. 10.1155/2015/462792.

[69] D. Vetturi, M. Tiboni, G. Maternini and M. Bonera. "Use of eye tracking device to evaluate the driver's behaviour and the infrastructures quality in relation to road safety". Transportation Research Procedia. vol. 45. pp. 587-595. Jan. 2020. 10.1016/j.trpro.2020.03.053. [70] G. Underwood, P. Chapman, N. Brocklehurst, J. Underwood and D. Crundall. "Visual attention while driving: sequences of eye fixations made by experienced and novice drivers". Taylor & Francis. vol. 46. no. 6. pp. 629-646. Jan. 2003. 10.1080/0014013031000090116. [71] G. Sinnapolu and S. Alawneh. "Intelligent wearable heart rate sensor implementation for invehicle infotainment and assistance". Internet of things (Amsterdam. Online). vol. 12. pp. 100277-100277. Dec. 2020. 10.1016/j.iot.2020.100277.

[72] H. Wang, "Application Heart Rate Variability to Driver Fatigue Detection of Dangerous Chemicals Vehicles".

[73] L. Chen, Y. Zhao, P. Ye, J. Zhang and J. Zou, "Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers".

[74] M. Q. Khan and S. Lee. "A Comprehensive Survey of Driving Monitoring and Assistance Systems". Sensors (Basel). vol. 19. no. 11. pp. 2574-2574. Jun. 2019. 10.3390/s19112574.

[75] R. McCraty and F. Shaffer. "Heart Rate Variability: New Perspectives on Physiological Mechanisms, Assessment of Self-regulatory Capacity, and Health Risk". SAGE Publishing. vol. 4. no. 1. pp. 46-61. Jan. 2015. 10.7453/gahmj.2014.073.

[76] R. Doorley, V. Pakrashi, E. Byrne, S. Comerford, B. Ghosh and J. A. Groeger. "Analysis of heart rate variability amongst cyclists under perceived variations of risk exposure". Elsevier BV. vol. 28. pp. 40-54. Jan. 2015. 10.1016/j.trf.2014.11.004.

[77] M. Lohani, B. R. Payne and D. L. Strayer. "A Review of Psychophysiological Measures to Assess Cognitive States in Real-World Driving". Frontiers in Human Neuroscience. vol. 13. Mar. 2019. 10.3389/fnhum.2019.00057.

[78] P. Avinash et al., "A Wearable Data Collection System for Studying Micro-Level E-Scooter Behavior in Naturalistic Road Environment".

[79] K. Weiss, T. Khoshgoftaar and D. Wang. "A survey of transfer learning". Journal of Big Data. pp. 1-40. May. 2016.

[80] T. Lin et al.. "Microsoft COCO: Common Objects in Context". Springer Science+Business Media. pp. 740-755. Jan. 2014. 10.1007/978-3-319-10602-1_48.

[81] A. Paszke et al.. "PyTorch: An Imperative Style, High-Performance Deep Learning Library". Cornell University. Dec. 2019. 10.48550/arxiv.1912.01703.

[82] J. Redmon and A. Farhadi. "YOLOv3: An Incremental Improvement". Cornell University. Apr. 2018. 10.48550/arxiv.1804.02767.