Roadway Contextual Risk Assessment Using Dynamic Traffic Conditions Data Obtained from Autonomous Vehicles

Vijay G. Bendigeri[1]; Fengjiao Zou[2]; Jennifer H. Ogle, Ph.D.[3]; and Kushal Kusram[4]

[1]Glenn Dept. of Civil Engineering, Clemson Univ., Clemson, SC.
Email: vbendig@g.clemson.edu
[2]Glenn Dept. of Civil Engineering, Clemson Univ., Clemson, SC. Email: fengjiz@g.clemson.edu
[3]Glenn Dept. of Civil Engineering, Clemson Univ., Clemson, SC. Email: ogle@clemson.edu
[4]Edgewater, CO. Email: kushalbkusram@outlook.com

ABSTRACT

Traditional road safety assessment methodologies do not recognize the driving environment's fast-changing dynamics that influence the contextual complexity and, ultimately, its risk. This paper proposes a method to use diverse open-source sensor data (LiDAR) collected by Waymo autonomous vehicles to estimate the road environment's complexity considering dynamic traffic conditions. The proposed contextual risk factor (CRF) model estimates the driving scene's complexity using the density and proximity of the objects around the vehicle. The data was analyzed frame-by-frame, and contextual risk categories of high, medium, and low were assigned. The results revealed the objects in the scene well represent the contextual complexity. However, what the driver sees in front of them and within their forward reaction space is not directly representative of the complexity of the scene and vice versa.

INTRODUCTION

Traditional road safety assessment methodologies depend heavily on historical crash data and static roadway characteristics in the absence of rich contextual information regarding the dynamic driving environment (i.e. rapidly changing interactions with vehicles, pedestrians, bicyclists that can influence the risk level of the driving environment). As a proxy for the contextual information from the dynamic driving environment at crash sites, researchers have conducted case-control studies by returning to the site at the same time of the day, day of the week, and under similar weather conditions. The observed conditions are used as surrogate for the dynamic environmental and operating conditions at the time of the crash, but the observed conditions could be vastly different from those experienced at the actual time of the crash. The advent of autonomous vehicles (AV) open datasets have created new opportunities to measure dynamic complexity and incorporate dynamic interactions metrics into risk estimates and safety assessments. Identifying and predicting high-risk environments can significantly benefit safety research, driver education, auto-insurance risk assessment, AV route planning, and many more. For example, this research could allow driving instructors and driving rehabilitation specialists to score dynamic complexity during training and testing to ensure that the driver is competent in all situation levels.

From the literature review, the key variables for measurement of visual-clutter and cognitive load are the density of the objects and their proximity to the vehicle. As the number of objects in the driving environment increases, the amount of information that needs to be processed by the driver also increases, in turn increasing the cognitive load on the driver. Further, near objects

present a greater risk to the driver compared to distant objects. To measure these two important parameters, a Contextual Risk Factor (CRF) was estimated for each frame.

Everyday routine trips expose drivers to massive amounts of input that is either static (i.e., roadway configuration and traffic control devices), or dynamic (i.e., surrounding vehicles, traffic signals, and vulnerable road users) (Olson and Farber, 1996). An important concept related to driver information processing is the useful field of view (UFOV). The UFOV is defined as "the total visual field from which target characteristics can be acquired when the head and eye movements are excluded"; and the extent of the UFOV differs between drivers, depending on how well they select and process relevant information from the environment (Dewar and Olson, 2002). So, the drivers scan the whole driving environment on occasion while driving, but the focus is the forward visual field, the UFOV.

This paper aims to understand the dynamic scene complexity from a driver's perspective using the density and proximity of objects surrounding the vehicle. The authors developed a Context Risk Factor (CRF) model to estimate the driving scene's complexity and classify contextual risk. The output is a continuous heatmap that classifies the driving environment's complexity into high, medium, and low categories. The researchers use LiDAR data collected by Waymo AV to estimate frame-by-frame road complexity considering dynamic traffic conditions. The LiDAR data provides rich real-world activity information around the vehicle, including stationary and non-stationary objects such as vehicles, pedestrians, and signs. This study considered the density of the objects (stationary and non-stationary) in the entire driving environment and within the driver's cone of vision (COV) as a measure of the driving environment complexity. The authors contrasted the differences between the total environment complexity and the complexity within the driver UFOV. The following literature review covers the concepts of CRF as well as UFOV and stopping sight distance (SSD) that are critical to development of the analytical methodology.

LITERATURE REVIEW

A few prior research articles were reviewed which propose contextual risks for the driving environment. Jiang et al., estimated safety risk by understanding vehicle queue status, historical crash data and simulated driver-based data (Jiang et al., 2020). Their proposed risk model is sound in estimating risk profiles, however, it does not include true data from the context. Another study by Hu et al., used smart phone GPS data, geographical network information and historical crash data to build safety risk model (Hu et al., 2018). The use of telematics data for modelling risk paints the picture of context partially and does not represent the true context.

While several papers associated with CRF models were identified, all were defined based on historic traffic operations data (average daily traffic or peak/off-peak), static roadway characteristics (number of lanes, route type, urban/rural), and either simulated or recorded acceleration and velocity information. None took into consideration the potential interactions with other road users and elements. This paper develops an analysis framework using LiDAR sensor data from autonomous vehicle to develop metrics to explain the rapidly changing dynamic conditions of the driving environment (i.e., the proximity and density of all vehicles, pedestrians, bicyclists, and signs with respect to the subject vehicle). It was essential to understand the background research related to a driver's visual field and ability to process components of the driving environment. Therefore, we have included a review of literature on driver UFOV and SSD. Described below are key features from historical research on these concepts.

Researchers define all the information that a driver must process to operate a vehicle as the visual demand, including traffic on the road, roadway environment, information in the vehicle, and other inputs (Dewar and Olson, 2002). Human factors experts generally believe that when the visual demand increases, the risk of traffic crashes increases as well (Dewar and Olson, 2002). Abdel-Aty and Radwan modeled crash occurrence and involvement and found that heavy traffic increases the likelihood of crashes (Abdel-Aty and Radwan, 2000). One of the reasons for the crashes increasing with the traffic complexity or object-density is that the 'driver's cognitive load increases. Cognitive load is defined as "a multidimensional construct that represents the load that performing a particular task imposes on the cognitive system of a learner" (Paas and Van Merriënboer, 1994). It is believed that cognitive load plays a vital role in the process of learning complex tasks (Paas et al., 2003), such as driving. Cognitive load has been proven to impair drivers' ability to detect safety-critical events (Lee et al., 2007).

Because the object-density or visual clutter created by objects around the vehicle affect the driver information processing and cognitive load, this paper provides a framework to understand what is in and outside of UFOV to portray a richer picture of the driving scene. The driver's field of view contains objects that the driver can see through the vehicle's windows, or from the mirrors (Olson and Farber, 1996). Olson and Farber (Olson and Farber, 1996) summarized the information that a driver must see, including objects/obstacles that need to be avoided on the road, and information like traffic signs, markings, and traffic signals to keep the vehicle positioned in the lane along the direction of travel. Dewar and Olson (Dewar and Olson, 2002) mentioned that stationary human UFOV extends to around 180 degrees, but it decreases to near 140 degrees when the driver is 70 years old. It has also been found that older drivers tend to drive more slowly to compensate for their vision and cognitive deficiencies (Szlyk et al., 1995). UFOV also reduces with increase in vehicle speed, traffic congestion, rain, and any other high-demand tasks (Dewar and Olson, 2002). Researchers (Rogé et al., 2004) tested UFOV for young and older drivers driving at high speeds (80mph) and lower speeds (55mph). The results show that UFOV deteriorates with the increase of age and with the increase of the vehicle's speed. Other researchers estimated that when the drivers are travelling at 30 mph, they can see targets in a visual field of 150 degrees; however, when speed is doubled (60mph), drivers can only see targets in half of the visual field (only 75 degrees) (Dewar and Olson, 2002). As speeds increase, the distance required to perceive and react appropriately increases because drivers need to look further down the road for objects in the potential collision zone. As the UFOV narrows, it also expands in length due to an increase in SSD. SSD enables drivers to stop before reaching obstacles in the path, and it is the sum of the brake reaction distance and braking distance (AASHTO, 2018). Although there were variations in UFOV produced by different researchers, the main conclusion is consistent: UFOV decreases in all directions with increases in driving speed. The reasoning is that drivers tend to focus further ahead when speed is higher and the ability to process information decreases, leaving the driver with ‘tunnel vision’. Research also reveals that the UFOV decreases when the quantity of information to be processed in the driver's peripheral area increases (Mackworth, 1976) (Ikeda and Takeuchi, 1975), meaning that the level of object density is high and the road scene is complex.

This paper adopts the UFOV recommendation from Dewar and Olson's (2002) and uses it to measure contextual complexity within drivers' view/control and compare it to the complexity of the entire driving scene. The methodology seeks to create a CRF model using computer-driven autonomous vehicle data obtained from Waymo open dataset (waymo_open_dataset, 2019) to estimate complexity of driving scene and classify contextual risk within driver’s UFOV and the

total driving environment. The methodology section describes how the authors applied the varying UFOVs at different speeds and SSD to define the COV and extract the number of objects within the cone from the entire scene.

METHODOLOGY

The methodology aims to develop a framework to better understand the dynamic risk of the driving environment by creating a CRF Model Using AV data obtained from Waymo.

Data Source. This study uses the open-source autonomous vehicle perception dataset published by Waymo in 2019 (waymo_open_dataset, 2019). The dataset consists of high-quality LiDAR and video data obtained from multiple sensors mounted on the vehicle. LiDAR data was opted over image data because of the additional components that were available (i.e. coordinates of point cloud, object bounding boxes, etc.) which are important for accurate measurements of objects around the vehicle. The LiDAR dataset included high-quality 3D bounding boxes with object type annotations which were manually checked for accuracy by trained labelers. The object types that were tracked include vehicles, pedestrians, bicyclists, and traffic signs. Additionally, vehicle speed vectors in 3-dimensional space for each frame were provided. The data was collected in San Francisco, Phoenix and Mountain View (Sun et al., 2020) under low speed environments.

A total of 798 trip segments of perception data, comprised of 158,090 LiDAR frame samples, were analyzed for this study. The maximum accurate range of the long-range LiDAR mounted on the vehicle was 250 feet which corresponds with a max safe operating speed of 35 mph based on human SSD requirements. Objects beyond that range are less likely to be detected or classified. Thus, frames with vehicle speeds exceeding 35 mph were excluded from the analysis – approximately 13% of total Waymo data. Additionally, there were a substantial number of frames where the vehicle was not moving (zero speed) due to the urban and ultra-urban settings and stop-and-go traffic operations. At zero speed, the SSD and the COV were also zero, which skewed the sample towards zero. Thus, frames with speeds less than 0.1 mph were excluded from the analysis. After clipping the frames with speed greater than 35 mph and less than 0.1 mph, the sample size was reduced to 108,369 frames (68.54%).

Analytical Approach. A total of 798 scenes of perception data, each spanning a total of 20 seconds at 10 Hz/second (i.e., ~200 LiDAR frames), were analyzed for this study. The total number of objects in each LiDAR frame and its proximity to the driver was estimated as a measure of scene complexity. The speed of the vehicle was used to derive SSD and select an appropriate UFOV. The SSD and UFOV were in turn used to construct a 3-dimentional filter cone, referred to as the cone of vision (COV), to identify any objects that fall within that cone.

Figure 1 provides a pictorial representation of the SSD, UFOV, and COV in the LiDAR point cloud. The orange dots represent the LiDAR points. The red car in the center is a representation of the autonomous vehicle. The blue bounding boxes with label “c” are the locations of detected cars. The pink bounding boxes with label “p” represent the location of the pedestrians. The green bounding boxes with label “s” represent the location of the traffic signs. The white dotted line in the center represents the direction of travel of the autonomous vehicle. The UFOV is the angle “a” between the orange lines that extends from the red car. The SSD is a distance that extends from the red car to a distance “d”. COV is the 3-dimensional volume of space bounded by the UFOV and truncated at the SSD, which is represented by blue boundary. Objects that fall within this COV were identified for each frame along with the total objects in the scene. In Figure 1

below, there are a total of 9 objects in the entire scene (5 cars, 2 pedestrians, and 2 signs) and only 4 objects within COV (2 cars and 2 pedestrians).

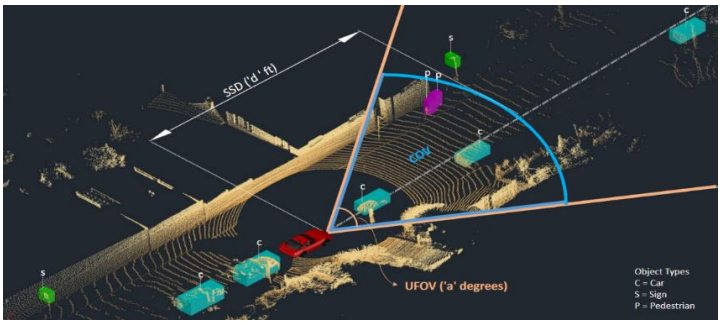

Figure 1. LiDAR point-cloud, SSD, UFOV & COV representation with object types.

Data Processing. The data processing was carried out in two stages to derive necessary attributes from the raw data (i.e., pre-processing, and post-processing).

Data Pre-Processing: All the scenes were stored in a google cloud bucket in a TensorFlow file format. The files were downloaded, and the raw data from the LiDAR data was extracted. The raw data contained segment context, LiDAR images, and LiDAR labels. The LiDAR data extraction consisted of point cloud with x, y & z coordinates in 3-dimensional space for each object referenced back to the autonomous vehicle as the origin (0, 0). The distance of the objects and their angle from the autonomous vehicle were estimated from x, y & z coordinates. Vehicle speed for each frame was obtained from the segment context metadata.

Data Post-Processing: COV is a function of SSD and UFOV. Thus, to determine the COV for each frame, first, the SSD and UFOV were computed. The SSD of the vehicle for each frame was calculated using the relationship in Equation 1. A standard driver reaction time of 2.5 seconds and a flat grade (i.e., grade = 0%) were assumed in all SSD estimations. UFOV shares an exponential relationship with speed. Following Dewar and Olson (2002), at zero speed the UFOV is 160 degrees, and as speeds rise, the UFOV decreases. UFOV is 150 degrees at 30 mph and 75 degrees at 60 mph speed. UFOV was computed using linear interpolation for all fractional speeds that fall in-between the above speed ranges (i.e. 0mph, 30 mph and 60 mph) within each scene. The COV boundary was calculated within the LiDAR point cloud using SSD & UFOV values. Objects that fall within the COV boundary were identified in addition to the total objects in the scene.

$$SSD=1.47st+s^2/(30*a/g) \qquad \text{Equation 1 (AASHTO, 2018)}$$

Where, SSD = Total stopping sight distance for the vehicle (feet), s= speed of the vehicle (mph), t= standard reaction of the driver (2.5 seconds), a = standard deceleration rate (11.2 ft/s^2) , g = acceleration due to gravity (32.2 ft/s^2)

Analysis. Figure 2 provides statistical distributions of the sample size for all the variables used in the analysis. The mean speed of the sample is 14.31 mph with a standard deviation of 8.97 mph. The mean SSD of the sample is 80.08 feet with a standard deviation of 61.5 feet. The mean COV was observed to be 146.8 degrees with a standard deviation of 7.51 degrees. The mean for total objects per frame is 62.18 with a standard deviation of 46.44. The total objects within the COV per frame is 9.62 with a standard deviation of 12.5.

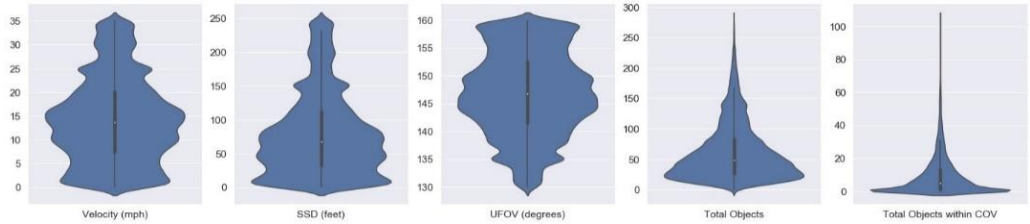

Figure 2. Statistical distribution of key variables.

From the literature review, the key variables for measurement of visual-clutter and cognitive load are the density of the objects and their proximity to the vehicle. As the number of objects in the driving environment increases, the amount of information that needs to be processed by the driver also increases, in turn increasing the cognitive load on the driver. Further, near objects present a greater risk to the driver compared to distant objects. To measure these two important parameters, a CRF was estimated for each frame using Equation 2.

$$CRF = \Sigma(1/obj_{distance}) \quad (2)$$

Where, $obj_{distance}$ = distance of the object from the autonomous vehicle (feet)

The inverse distance assignments allowed scene elements to be weighted in descending order with near objects receiving higher weights and more distant objects getting lower weights. This accounts for the limited reaction time associated with objects that are nearer to the subject vehicle. The summation of these inverse distance assignments for objects in the driving environment accounts for the total number of objects in the scene, i.e., object density. The scene CRF was estimated for each frame considering all the objects. Additionally, CRF was estimated for the COV filter in each frame. This provides an estimate of complexity within the driver's COV. Statistical quartiles for the total sample were estimated for the entire scene CRF and for CRF within COV. A frame was categorized as high if the CRF > 75th percentile, medium if CRF was in inter-quartile-range (between 25th percentile & 75th percentile), and low if CRF was less than 25th percentile, respectively. All the frames were assigned a category of high, medium, or low based upon the scene CRF, its respective quartile range.

RESULTS

The analysis provided frame by frame comparison of contextual complexity based upon the density of objects and their proximity to the autonomous vehicle as represented by the CRF. All trips were categorized as high, medium, or low risk trip based upon the statistical mode of the trip's CRF category.

Figure 3 provides an example of three such trips categorized as low, medium, and high-contextual risk trips. The figure consists of a 2x3 matrix of risk plots. Each column contains two graphs pertaining to an individual trip. The left most column is for a low-risk trip, the middle column is of a medium risk trip, and the right most column is of a high-risk trip. The x-axis represents time in seconds. The y-axis describes the CRFs. The top row illustrates CRF for the entire scene and the bottom row displays CRF within COV. The corresponding video of each of these trips is provided in the respective hyperlinks (High, Medium, Low).

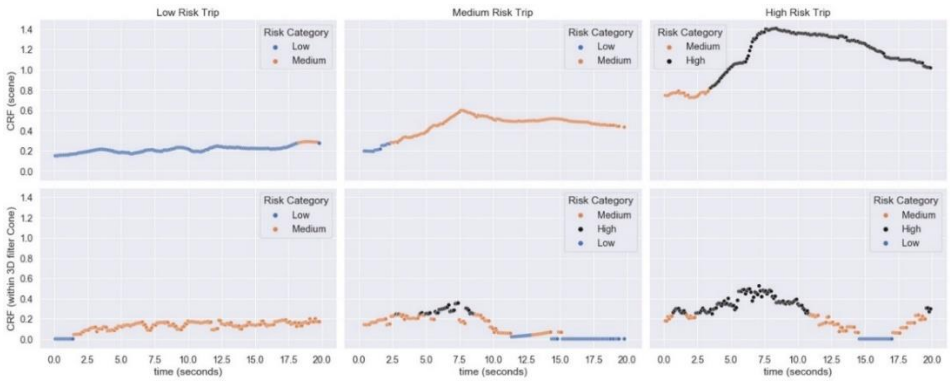

Figure 3. CRF plots for high, medium, and low risk trips (speed<= 35 mph & >0.1 mph)

The upper right plot shows a high-risk trip on a 2-lane urban road situated in an ultra-urban area. The trip predominantly consists of a high density of objects in close proximity to the vehicle. The trip starts with a medium-risk context for 3 seconds and transitions into a high-risk context for the remainder of the trip. After 3 seconds, the vehicle enters an intersection that has a lot of vehicles, pedestrians, and bicyclists, thus elevating the CRF. After traversing the intersection, the vehicle enters another 2-lane urban road with curb-side parking and moving vehicles and pedestrians in close proximity maintaining the elevated CRF. The bottom right plot shows the resulting CRF within the driver's COV. The CRF within driver's COV (bottom right plot in Figure 3) and the overall CRF of the scene (top right plot in Figure 3) vary greatly. This is because a lot of objects fall within the driver's COV at the start of the trip as the vehicle traverses the intersection, making it a high risk for the driver. The contextual complexity in the driver's COV later diminishes to medium and then to low risk as the vehicle decelerates and comes to standstill (between 15-17 seconds).

The medium-risk trip (middle top and bottom plots in Figure 3) consists of an urban multi-lane highway with a center two-way-left-turn lane. At the start of the trip, there are a few objects in the overall scene, making it a low-risk environment. At the 2-second time frame, pedestrians and bicyclists preparing to cross the road are detected elevating the risk gradually to medium as the vehicle advances. This trend is noticeable in the top middle plot. The corresponding CRF within COV of the driver also intensifies to a high-risk which is represented in bottom middle plot.

The low-risk trip is comprised of vehicle driving on a local neighborhood road with no moving vehicles, pedestrians, or bicyclists. The entire scene complexity remains low for a significant part of the trip. On the contrary, the CRF within COV remains at medium risk throughout the trip except at the beginning when the vehicle accelerates from standing still.

Based on the visual inspection of the trips, the three examples provided (high, medium, and low) accurately characterize the contextual risk of the driving environment.

CONCLUSION

The aim of this study was to develop a contextual metric to assess the relative risk of the dynamic driving environment and constantly changing interactions between a driver and other dynamic components of the driving environment (i.e., other vehicles, pedestrians, bicyclists, and traffic control). This is demonstrated by estimating CRF for the entire trip utilizing LiDAR data. The CRF captures the density and proximity of the objects from the vehicle, which are key parameters influencing the trip's complexity.

The analysis also illustrated that what the driver sees in their COV is not directly representative of the complexity of the entire scene and vice versa depending upon the location of the objects. The complexity of the scene is well represented by the account of objects in the scene. The objects considered by Waymo i.e., vehicles, pedestrians, bicyclists, and signs are the ones that significantly affect the driving risk (NCHRP Report 600, 2012). Additionally, they are also the ones that the driver seeks to notice and track while driving.

REFERENCES

AASHTO. 2018. *A Policy on Geometric Design of Highways and Streets*, 7th Edition. American Association of State Highway and Transportation Officials.

Abdel-Aty, M. A., and Radwan, A. E. 2000. Modeling traffic accident occurrence and involvement. *Accid. Anal. Prev.* 32 5, 633–642.

Dewar, R. E., and Olson, P. L. 2002. *Human Factors in Traffic Accident Litigation in: Human Factors in Traffic Safety*. Lawyers & Judges Publishing Company, Inc.

Hu, X., Zhu, X., Ma, Y. L., Chiu, Y. C., and Tang, Q. 2019. Advancing usage-based insurance–a contextual driving risk modelling and analysis approach. *IET Intelligent Transport Systems.*, 13(3).

Ikeda, M., and Takeuchi, T. 1975. Influence of foveal load on the functional visual field. *Percept. Psychophys.* 18 4, 255–260.

Jiang, S., Jafari, M., Kharbeche, M., Jalayer, M., and Al-Khalifa, K. N. "Safe Route Mapping of Roadways Using Multiple Sourced Data," in *IEEE Transactions on Intelligent Transportation Systems*.

Lee, Y. C., Lee, J. D., and Boyle, L. N. 2007. Visual attention in driving: The effects of cognitive load and visual disruption. *Hum. Factors* 49 4, 721–733.

Mackworth, N. H. 1976. Stimulus density limits the useful field of view. *Eye movements Psychol. Process.* 307–321.

NCHRP Report 600. 2012. Human Factors Guidelines for Road Systems: Second Edition, Human Factors Guidelines for Road Systems: Second Edition.

Olson, P. L., and Farber, E. 1996. *Forensic aspects of driver perception and response*, Second Edi. ed. Lawyers & Judges Publishing Company, Inc.

Paas, F., Tuovinen, J. E., Tabbers, H., and Van Gerven, P. W. M. 2003. Cognitive load measurement as a means to advance cognitive load theory. *Educ. Psychol.* 38 1, 63–71.

Paas, F. G. W. C., and Van Merriënboer, J. J. G. 1994. Instructional control of cognitive load in the training of complex cognitive tasks. *Educ. Psychol. Rev.* 6 4, 351–371.

Rogé, J., Pébayle, T., Lambilliotte, E., Spitzenstetter, F., Giselbrecht, D., and Muzet, A. 2004. Influence of age, speed and duration of monotonous driving task in traffic on the driver's useful visual field. *Vision Res.* 44 23, 2737–2744.

Sun, P., Kretzschmar, H., Dotiwalla, X., Chouard, A., Patnaik, V., Tsui, P., Guo, J., Zhou, Y., Chai, Y., Caine, B., et al. 2020. Scalability in perception for autonomous driving: Waymo open dataset, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 2446–2454.

Szlyk, J. P., Seiple, W., and Viana, M. 1995. Short note relative effects of age and compromised vision on driving performance. *Hum. Factors* 37 2, 430–436.

waymo_open_dataset. 2019. Waymo Open Dataset. URL https://www.waymo.com/open.