How Effective are Identification Technologies in Autonomous Driving Vehicles?

Otto B. Piramuthu
Computer Science, UIUC
Urbana, IL 61801, USA
obp2@illinois.edu

Matthew Caesar
Computer Science, UIUC
Urbana, IL 61801, USA
caesar@cs.illinois.edu

***Abstract*—Autonomous driving necessarily involves timely awareness of surrounding environmental conditions to facilitate safe navigation. Vision is therefore of paramount importance in these vehicles. Cameras, LiDAR, RADAR, and GNSS provide a reasonable amount of necessary environmental input in a majority of current autonomous driving initiatives. Several published studies vouch for the advantages of autonomous vehicles over their human-driven counterparts, in principle. However, extant literature does not provide clear guidance on the extent of dominance, if any, of autonomous vehicles in terms of accident avoidance. We consider 'vision' inputs in autonomous vehicles and compare their performance to that of human-driven vehicles based on recent accident data and show that current state-of-the-art of vision technology in automated vehicles are grossly insufficient for truly autonomous vehicles. Specifically, our results illustrate the extent of deficit that must be addressed in state-of-the-art machine learning algorithms and vision sensors that are used in autonomous driving vehicles.**

***Index Terms*—Autonomous/smart vehicles, reliability, vision**

I. INTRODUCTION

The market for (partially) autonomous cars has been steadily growing over the last few years. Depending on the source, the estimates for this market vary. One such estimate projects the market for partially autonomous cars to reach 13% of all vehicles on the road by 2025, estimated to be worth about $42 billion [1] – with the share of completely autonomous cars at 0.5% of the cars on the road by 2025. For such estimates to become reality, these partial or completely autonomous vehicles will require better sensors and object identification technology since existing (partially) autonomous cars fare poorly when driven in rural roads, at high speeds, under signal interference conditions due to high density of autonomous vehicles nearby, under inclement weather conditions, among others. Regardless of the technologies used, the primary goal of an autonomous vehicle is to ensure safe transportation of its passengers through maintenance of safe distance from other objects in its path.

A completely autonomous vehicle navigates itself through appropriate environmental awareness, with no need for human input [2]. On the other extreme, a completely non-autonomous vehicle requires continual human input and attention for safe operation. To facilitate ease of understanding and consistency, the degree of vehicle autonomy was classified by SAE (Society of Automotive Engineers) International in an updated document (J3016_202104) [3] that was released in 2021. This classification includes six levels (numbered 0 through 5 representing respectively "No Driving Automation", "Driver Assistance", "Partial Driving Automation", "Conditional Driving Automation", "High Driving Automation", and "Full Driving Automation") of autonomy that range from manual to automated system with the levels distinguished based on the extent of different human input requirements. Although demonstration vehicles have been in existence for a few years now, only a couple of car models (e.g., Honda Legend in Japan) have so far reached level 3 in production. Given this state of affairs, vehicles that conform to higher autonomous levels (i.e., level 4, level 5) are not expected to be available in the market at least for the next several years. Regardless, there is a growing interest in further advancing the autonomy of vehicles. For example, the number of Autonomous Vehicle Testing Permit holders in California grew from the high 30s in 2019 to 55 as of 19 October 2021. Among these, three (Waymo LLC, Nuro Inc., AutoX Technologies Inc) had driverless testing permits [4].

Today's (partially) autonomous vehicles rely on a combination of a balanced mix of technologies that include cameras, RADAR (RAdio Detection And Ranging), LiDAR (Light Detection and Ranging), GNSS (Global Navigation Satellite System), and V2X (vehicle-to-everything) communications. Each of these technologies have capabilities that overlap, are used for similar purposes to provide a certain degree of redundancy, and have their own strengths and weaknesses. Wang et al. [5] show that object detection accuracies of stereo image-based camera-generated data with appropriate representation and that from LiDAR-based 3-D object recognition are comparable. In addition to today's (partially) autonomous vehicles' inability to reliably operate in heavy weather (rain, snow), on unpaved roads, or in mixed traffic conditions, these vehicles have been shown to perform worse than human drivers in some cases [6].

There is therefore a need to study the capabilities of existing technologies and systems from the perspective of their synergy and their relative effectiveness in terms of accident avoidance vs. that of previous generation technology. To this end, we evaluate the effectiveness of state-of-the-art systems that include RADAR, LiDAR, and cameras in terms of their ability to identify accident-prone situations. Specifically, we study

978-1-6654-0306-1/21/$31.00 ©2021 IEEE

the use of these systems in autonomous driving scenarios and their relative effectiveness vs. human driven automobiles with respect to accident rates. We consider the use of only RADAR, LiDAR, and cameras in autonomous driving environments and evaluate their effectiveness with the identification of specific things (bicycles, people, other automobiles, etc.) that are on the path of the automobile of interest.

We attempt to answer the following: (a) Do the current state-of-the-art of sensors, specifically camera, LiDAR, and RADAR, have the potential to bring down the accident rate associated with autonomous vehicles to below that of human-driven vehicles? (b) If not, what is the minimum required sensor performance level to reach parity?

We use (a) failure rates of RADAR, LiDAR, and cameras in identifying things (bicycles, people, other automobiles, etc.) that are directly on the path of the reference vehicle and (b) historical automobile accident rates in terms of average number of fatalities per incident, average number of injuries per incident, and number of crash incidents. With data on sensor failure rates and actual accident rates, we use principles of reliability analysis to operationalize this study. Extraneous influences on autonomous automobiles such as cyberattacks are beyond the scope of this study and are disregarded.

To facilitate appreciation of autonomous vehicles, we discuss their history, various implementation and performance challenges, and recent experiences with partially autonomous vehicles in Section II. In Section III, we develop a reliability analysis model and through sensitivity analysis consider various scenarios with existing state-of-the-art vision technologies. We summarize results from this study and discuss other possible vision sensors for autonomous vehicle use in Section IV. We conclude the paper in Section V.

II. Background and related work

The history of autonomous driving goes back to about a century when Francis Houdina demonstrated a radio-controlled 1926 Chandler that was operated by a following car on Broadway and Fifth Avenue in New York City [7]. There have been several attempts at autonomous driving since then, with a report [8] on computer controlled cars about four decades later. In addition to details on the requirements and the problems that need to be addressed to realize such a car, this report identifies visual pattern recognition as the "most intricate single problem."

From a survey of about 1500 US consumers on their willingness to buy autonomous cars, the primary concerns raised by the respondents were broadly related to reliability, cybersecurity, and questions on interactions with other vehicles on the road [1]. A recent study [9] estimated the number of miles of driving that are required to guarantee a certain level of reliability in autonomous driving vehicles. These authors also considered the confidence intervals associated with these results. They specifically considered historical data on human-driven vehicle failure rates based on the number of fatalities, injuries, and crashes per 100 million driven miles and attempted to derive the number of miles an autonomous car has to be driven to match these statistics. Their results indicate that the number of years it would take to reach a certain level of reliability is unreasonably high (e.g., 500 years "with a fleet of 100 autonomous vehicles driving 24 hours a day, 365 days a year, at an average speed of 25 miles per hour"). The take-away from this study is that we need other means to supplement actual physical driving to reach the goal. Simulation is one such approach. Waymo cars, for example, reportedly drove over 2.5 billion simulated miles in 2016 under conditions that were richly packed with "interesting scenarios than the average mile of driving" [10].

While it is not always possible to identify and recognize objects in the automobile's path that could cause an accident, not all recognized ones may necessarily be harmful. For example, an inaccurately identified object may prove to be quite harmless (e.g., plastic bag, a piece of paper, or a small clothing item on the road). The system needs to learn to distinguish such harmless objects from those in their complementary set. Sensor-related issues that result in the detection of non-existent objects is a challenge for system designers.

If the system is designed to be too careful, the ride is bound to be slow and jerky as the vehicle frequently slows down for harmless or non-existent objects. On the other hand, when the system is not designed to be careful, the ride is bound to be smooth. However, this smooth ride comes at a cost wherein a significant object on the vehicle's path has a positive probability of being ignored. This is what happened in Tempe around 9:58PM on 18 March 2018 where the object was a human being. According to the preliminary report [11], the object (person) was detected by Uber Volvo's RADAR and LiDAR about six seconds before impact when the vehicle was traveling at 43mph. As the vehicle approached the person on the road, its self-driving software classified the person first as an unknown object, then as a vehicle, and later as a bicycle. Such unclear classification has the potential to cause uncertainties in the expected travel path of the vehicle.

There have been other accidents with human fatality as per www.tesladeaths.com when on advanced Autopilot driver assistance system. These accidents clearly show the tradeoffs between false positives and false negatives – as the RADAR signal identifies a myriad of stationary objects on the road, to prevent the car from braking at every such (false positive) object, Tesla's system apparently places more weight on the cameras than on the RADAR; On the other hand, Uber's car identified the object with enough warning for the driver to be able to respond on time (false negative). The issue of liability is a concern under such circumstances. Volvo has promised to accept full liability for its autonomous mode vehicles [12].

A complex system such as an autonomous driving vehicle has an appreciable number of software and hardware components that have the potential to fail. A minor failure (e.g., false sensor reading, distorted signal, software error) can result in catastrophic outcome. For example, the ability of Volvo self-driving car's "Large Animal Detection System" to identify deer, elk, and caribou did not translate to kangaroos because their hopping behavior confounded its system [13].

Real-world driving is replete with variations related to background structures, lighting conditions, shadows, mobile and stationary objects in addition to their varied orientations and partial occlusions. Furthermore, such objects need to be detected in real-time from a moving vehicle to provide notifications that could result in a timely response. In addition to detecting objects in its path [14] including small and distant ones [15] while taking the correct path to its destination, an autonomous vehicle also needs to be aware of its instantaneous state and position [16].

It is therefore essential for autonomous vehicles to effectively and efficiently learn to navigate their path without any accident. This process requires huge volumes of representative data and appropriate learning algorithms. To this end, there have been several initiatives to gather relevant data sets and virtual testing environments [17]. Among the learning algorithms, Convolution Neural Networks (CNNs) are widely used in autonomous driving applications. However, existing CNN frameworks do not scale well when processing multiple video streams as they suffer from higher per-frame latency and lower per-stream accuracy. Studies have attempted to address these issues [18]. A recent addition to the data collection is from Waymo [19] which includes 12 million LiDAR box annotations and 12 million camera box annotations. Sun et al. [19] also provide benchmark results of several state-of-the-art 2D and 3D object detection and tracking methods such as PointPillars followed by CNN with this data. However, the reported performance is far from ideal. For example, the best and worst Multiple Object Tracking Accuracy (MOTA) they achieved on their data are 70.6 and 12.5 for vehicles and 52.5 and 22.3 for pedestrian identification.

Even when the systems in the (partially) autonomous vehicle are operating effectively as required, malicious hacking is a threat whereby such vehicles are intentionally manipulated for purposes of amusement or crime [20]. While some level of malicious threat is expected in such vehicles, it is not trivial to predict all such possible threats and to prepare or enable the system to protect itself from such threats. In addition to endemic issues that are associated with object identification through computer vision, adversarial examples [21] add yet another wrench in the works. Adversarial examples are minor modifications, that are generally imperceptible to the human eye, to the image of an object which result in the computer vision system identifying the object as something else (e.g., a Stop sign identified as an Yield sign [22]).

From the above discussion, it is clear that several significant challenges need to be overcome for level 5 autonomous vehicles to become reality. We now consider only those that are related to vision system in these vehicles, with specific emphasis on their performance strengths and limitations.

III. Model and Analysis

We develop a reliability analysis model to study the performance of vision systems in autonomous vehicles under several different scenarios through sensitivity analysis. We use historical (human-driven vehicle) accident data as a baseline for comparison. The idea is to compare the expected accident proneness of autonomous vehicles with that of human-driven vehicles to evaluate the sufficiency of current technology to meet/exceed historical (human-driven) vehicle safety records.

Tom Laux at Continental's Segment High Resolution Flash LiDAR claims that five nines (99.999%) performance requires multiple sense functions that include "a 2D sensor, like a camera, to image the color of signals, traffic signs and lane markings. Then a RADAR to determine velocity, and finally a LiDAR to give accurate angular resolution and 3D imaging in a far more precise manner." [23]. To this end, we now consider the reliability with the use of LiDAR, RADAR, and camera in autonomous vehicles. We use failure (to detect an object) rate to determine the reliability of a fully autonomous vehicle. Failure rate (λ) is defined as the number of failures per unit of operating time. For purpose of this study, an automobile accident is considered a failure. With the knowledge of the exact historical failure rate (ratio of the number of accidents to the number of hours automobiles are driven) and the failure rate of each of the vision-based technologies (RADAR, LiDAR, cameras) and associated systems (e.g., pattern recognition software), we can determine if the performance measurements of autonomous vehicles embedded with these technologies are on par or different (better or worse) than human drivers.

A. *Sensor-based failure estimation*

We use the following notation.

t - Time

λ_S - System failure rate

λ_A, λ_B, λ_O, λ_P - Camera-based failure rates for animal, bicycle, other vehicles, pedestrian

λ_L, λ_R - Failure rate of LiDAR, RADAR

$R_i, R_i(t)$ - Reliability of component i (at time t)

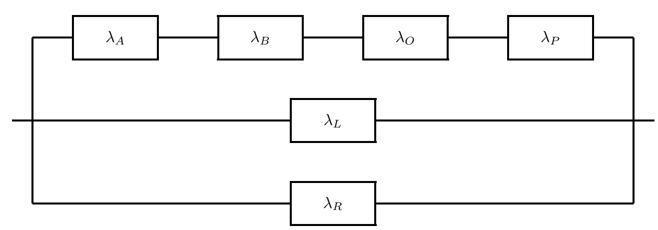

Figure 1: Block Diagram of Network of Series Elements in Parallel

For clarity, we make use of block diagrams. We assume that the camera systems in autonomous vehicles are able to recognize and identify specific objects such as animal (A), bicycle (B), other vehicle (O), and pedestrian (P) through some computer vision means. We also assume that the inputs from RADAR and LiDAR are used only to recognize the presence of an object on the path of the vehicle. It is straight-forward to include the identification of specific object, as with the case of the cameras in such systems.

We first combine the series components (A, B, O, and P) into one equivalent component (C). The failure rate of each of the components in series in Figure 1 is assumed to be constant with no burn-in or wearout periods. Here, since failures in the series system components depicted in the top part of Figure 1 are statistically independent, the reliability (R_C) of this series system representing processed inputs from the camera(s) with non-identical components is given by

$$R_C(t) = \prod_{i=1}^{n} R_i(t)$$

where n is the number of components (here, 4, comprising A, B, O, and P) and R_i is the reliability of the i^{th} component. Here, t signifies that these reliability measures are taken at time t.

Since the components have constant failure rates, as per norm we assume that the component failure time follows the exponential distribution. It follows that the component i has a constant failure rate (λ_i) with reliability $R_i(t) = e^{-\lambda_i t}$

Now,

$$R_C(t) = e^{-\sum\limits_{i=1}^{n} \lambda_i t}$$

and the mean time to failure (MTTF$_C$) is:

$$\int_0^\infty e^{-\sum\limits_{i=1}^{n} \lambda_i t} dt \;=\; \frac{1}{\sum\limits_{i=1}^{n} \lambda_i}$$

The effective failure rate of the four elements in series is: $\frac{1}{\text{MTTF}_C} = \sum\limits_{i=1}^{n} \lambda_i \;=\; \lambda_A + \lambda_B + \lambda_O + \lambda_P$

We can now redraw Figure 1 equivalently as Figure 2.

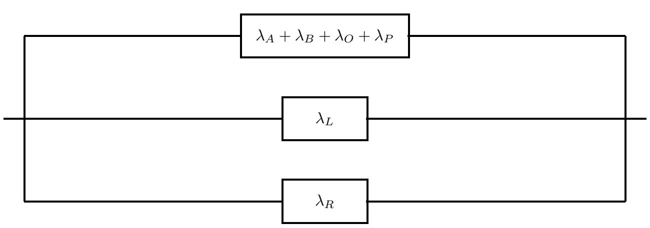

Figure 2: Equivalent Block Diagram (First Reduction) of Network of Series Elements in Parallel

Let $\lambda_C = \lambda_A + \lambda_B + \lambda_O + \lambda_P$

MTTF of the entire system of six elements (MTTF$_S$) is:

$$\int_0^\infty \Big\{1 - \prod_{j=1}^{m} (1 - e^{-\lambda_j t})\Big\} dt =$$

$$\frac{1}{\lambda_C} + \frac{1}{\lambda_L} + \frac{1}{\lambda_R} - \frac{1}{\lambda_C + \lambda_L} - \frac{1}{\lambda_C + \lambda_R} - \frac{1}{\lambda_L + \lambda_R} + \frac{1}{\lambda_C + \lambda_L + \lambda_R}$$

The failure rate of the entire system (λ_S) $= \frac{1}{\text{MTTF}_S} =$

$$\frac{1}{\frac{1}{\lambda_C} + \frac{1}{\lambda_L} + \frac{1}{\lambda_R} - \frac{1}{\lambda_C+\lambda_L} - \frac{1}{\lambda_C+\lambda_R} - \frac{1}{\lambda_L+\lambda_R} + \frac{1}{\lambda_C+\lambda_L+\lambda_R}} \tag{1}$$

Since λ_A is not readily available, we assume $\lambda_A = \lambda_P$. We also assume $\lambda_L = \lambda_R$ for now. From Wang et al. [5], we use the best case scenario values for λ_L $(= 1 - 88.7 = 11.3)$, λ_P $(= 1 - 13.5 = 86.5)$, λ_B $(= 1 - 31.3 = 68.7)$, and λ_O $(= 1 - 66.8 = 33.2)$. From equation 1, we determine the failure rate of the system to be 7.532714 failures per frame.

B. Actual failures in human-driven vehicles

We now discuss the actual numbers corresponding to human-driven vehicles in the United States during 2017. As per NHTSA [24], the total number of automobile crashes in 2017 was 6452000, which includes cases with fatalities, injuries, and only property damage. According to AAA [25], the number of drivers in the US in 2017 was about 225.848 million (87.2% of the 259 million population of age 16+). The average number of hours per year that was driven per driver is 310 [25]. From these, we calculate the number of automobile crashes per million hours of driving as $\frac{6452000}{225.848*310} = 92.15447$

The National Association of City Transportation Officials (NACTO) developed a second version of its Blueprint for Autonomous Urbanism, which encourages cities to deploy "automated vehicles [that] should be programmed for low speeds (25 mph or less) on city streets, and programmed to automatically detect and yield to people outside of the vehicle" [26]. Conservatively, we assume the autonomous vehicle speed to be 25mph (= 0.011176km/sec). For analysis, we assume that the autonomous vehicle looks ahead about 100 meters. A vehicle traveling at 25mph traverses 100 meters in $\frac{0.1}{0.011176} = 8.947745$ seconds. While autonomous vehicles process images at several frames per second, it is likely that the objects in consecutive frames do not change by a lot. Let us assume that just a single frame observation is used to traverse this 100m distance. With a hardware failure rate of 7.532714 failures per frame, it is easy to see (vs. the 92.15447 vehicle crashes per million hours of driving as determined in the previous paragraph) that the current state-of-the-art of sensors (LiDAR, RADAR, camera) in autonomous vehicles are far from sufficient for truly autonomous driving. We now consider these statistics in detail, with specific emphasis on the number of miles between accidents, the total number of accidents, the probability of proactive moves by objects that are in the path of the ego vehicle, and appropriate and timely response by the ego vehicle to an accident prone situation. An ego vehicle is the vehicle of interest that perceives its environment through sensors.

C. Failures in human-driven and autonomous vehicles

We first consider the expression for the probability (C) that an autonomous vehicle with reliability R experiences k failures when it is driven N miles. This is given by the following expression [27] [9]

$$C \;=\; 1 - \sum_{i=0}^{k} \frac{N!}{i!(N-i)!} R^{N-i} (1-R)^i$$

For $k = 0$, we have $C \;=\; 1 - R^N$. While this expression is useful to determine the failure rate for a given confidence level,

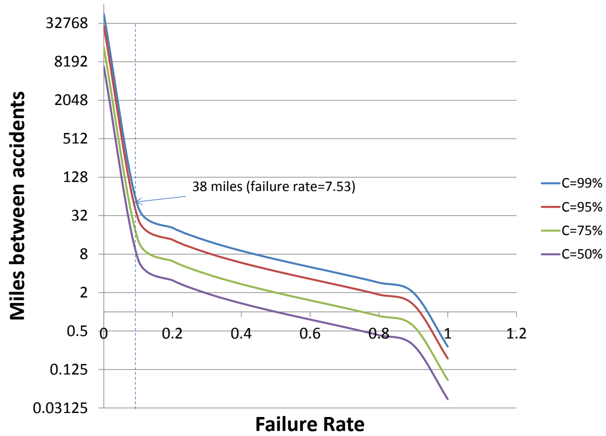

Figure 3: Failure rates and corresponding distance between accidents in autonomous vehicles

we are interested in the number of miles that an autonomous vehicle is driven before a failure occurs for a given confidence and reliability. We can rewrite the above expression as

$$n \;=\; \frac{ln(1-C)}{ln(R)}$$

To understand the behavior of failure rate in autonomous vehicles, we first plot various failure rates and the resulting number of accidents. To operationalize this, we use the expression for n given above. We draw graphs for a few different C values. While $C = 95\%$ is more commonly used, the other values are also used sometimes [9]. The miles between accidents is highest for C value of 99%, which is followed respectively by those for 95%, 75%, and 50%. Overall, the average number miles that pass between any two successive accidents decreases with increase in failure rate for all C values. We see from Figure 3 that for a system failure rate of 7.53, which is the number of failures per frame that we determined in Section III-B, on average there would be 38.27 miles between failures for $C = 95\%$. The corresponding actual (human-driven vehicle) figure is 402.98 $\left(= \frac{2{,}600{,}000{,}000 \text{ miles}}{6{,}452{,}000 \text{ accidents}}\right)$ miles between accidents which is associated with system failure rate of 0.741. The stark difference between miles between failure for autonomous and human-driven vehicle shows that although autonomous vehicles are devoid of negative human characteristics such as being distracted while driving, human drivers are much better at situational awareness in being able to react to other objects in their path and successfully navigate a good number of accident-prone situations.

D. Easy, moderate, and hard cases

Based on the bounding box height and occlusion/truncation level, the benchmark data used in Wang et al. [5] has three cases that include easy, moderate, and hard. The easy case corresponds to objects that are within 30 meters of the ego vehicle and hard case includes objects that are farther from the ego vehicle and appear smaller, which renders their detection rather difficult. In Section III-C, we considered the easy case, which is the best case scenario. We now consider the moderate and hard cases. For the moderate case, the corresponding values include λ_L $(= 1 - 84 = 16)$, λ_P $(= 1 - 9.1 = 90.9)$, λ_B $(= 1 - 24 = 76)$, and λ_O $(= 1 - 47.2 = 52.8)$. From equation (1), the failure rate for the entire moderate system is 10.66499 failures per frame. For the hard case, the corresponding values include λ_L $(= 1 - 75.3 = 24.7)$, λ_P $(= 1 - 9.1 = 90.9)$, λ_B $(= 1 - 21.9 = 78.1)$, and λ_O $(= 1 - 40.3 = 59.7)$. From equation 1, we determine the failure rate for the entire hard system to be 16.45852 failures per frame.

In reality, any given autonomous vehicle may experience a mix of easy, moderate, and hard scenarios that depend on location, time of day, characteristics of the path, among others. For example, highway driving in light traffic is generally less complex and much easier than city driving in heavy traffic. Highways are also generally devoid of objects that move in unpredictable ways and cause much uncertainty such as pedestrians and bicycles. As compared to city driving, highway driving is often at a relatively high speed, which has its own dynamics in terms of reaction time to sudden unexpected events such as a vehicle moving into the ego vehicle's path with no notice, emergency situations that develop either inside or outside the vehicle, and extreme or unexpected behavior from other objects that violate law and order. When hard braking is necessary under such a circumstance, high speed renders it rather difficult to accomplish within a short stopping distance. Although city driving occurs at relatively low speeds, uncertainties introduced by bicycles, pedestrians, the large number of intersections, among others impose additional layers of complexities.

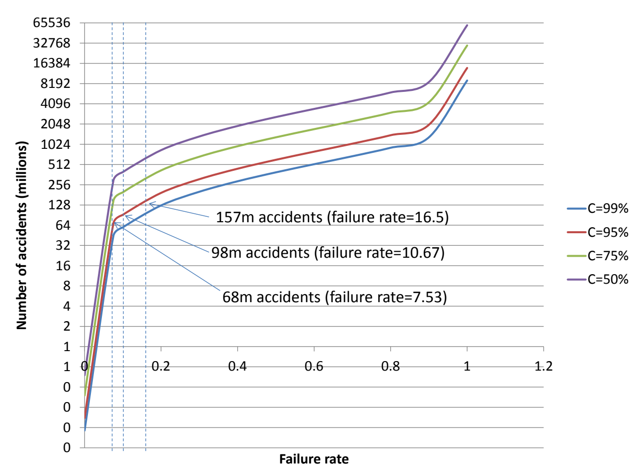

Figure 4: Failure rates and number of accidents in autonomous vehicles [note: fractions not displayed in the vertical axis]

As in Figure 3, we consider the four C values of 99%, 95%, 75%, and 50% and plot the number of accidents in millions vs. the failure rate. As can be seen in Figure 4, the lowest number of accidents for any given failure rate corresponds to the lowest C value. As is most common, we now focus only on C value of 95%. For the three cases (hard, moderate, and easy) with failure rates of 16.5, 10.67, and 7.53, the respective number of accidents are 157 million, 98 million, and 68 million. In reality, for human-driven vehicles, the number of accidents in the United States was around 6.452 million in 2017 which corresponds to a failure rate of 0.7407. These real-life accident count includes all cases that had fatalities, injuries, and only property damage. We considered the number of accidents for autonomous vehicles at the aggregate level and are unable to separate the autonomous vehicle statistic into fatalities, injuries, and only property damage cases simply because we do not yet have necessary available data for the aggregate, let alone for individual categories [28]. For example, the total number of miles driven by all autonomous vehicles across all years over all countries in the world so far eclipses when considered against the number of human-driven miles in one year (2017) of 2.6 trillion miles just in the United States. Although the number of property damages that have been caused by autonomous vehicles is probably in the low hundreds to date, the number of fatalities caused by autonomous vehicles are in the low dozens at best. Similarly, the number of injuries that have been caused by autonomous vehicles are quite low when compared against those of human-driven vehicles. This is to be expected since the number of autonomous vehicles is a rather small fraction of the total number of vehicles.

E. Active objects in reachable space

In a good number of extant autonomous driving studies, the models developed take a snapshot of the surroundings with particular emphasis placed on the reachable space [29] of both the ego vehicle and all potential objects in its path. The ego vehicle's trajectory is then determined (i.e., motion planning problem) based on the expected location of all such objects as they move along their planned path. Of particular interest are cases where the reachable space of the ego vehicle and that of the other objects in its path have the potential to (partially or fully) overlap. Care is taken to individually consider each such object and its dynamic reachable space over time as long as potential for overlap remains.

Fixed and mobile objects trigger different responses from the ego vehicle, which has to respectively take into consideration fixed and dynamic reachable spaces of other objects in its path. The reachable space for a fixed object is fixed. For a mobile object, the reachable space of interest to the ego vehicle depends on various parameters that include its relative speed (vs. ego vehicle), direction, and footprint.

When the ego vehicle is moving in the same direction as the other moving objects in its potential path, the size and relative speed of these objects with respect to that of the ego vehicle play a significant role due to various reasons (e.g., imminent possibility of contact, implications of Solomon curve). Moreover, since multiple objects share the same environment, it is necessary to address the most imminent threat first while simultaneously accounting for the potential threats that immediately follow since the next threats may be closely spaced in time and plan to address those threats need to be simultaneously generated when dealing with the first threat.

While a good first step is to assume that the other objects will continue in their current paths at the same speed, etc., when deciding on the ego vehicle's trajectory, it is worth noting that each of the other objects may also simultaneously consider the objects in their paths to make necessary modifications to their potential trajectory. What was considered reachable space overlap at an earlier time point may not remain so as the other party makes necessary adjustments to avoid a collision situation and vice versa. Rogue and/or adversarial vehicles add yet another level of complexity to this environment [30] [31].

Reachability space determination has a tremendous effect on vehicle safety. Methods that use reachability cannot guarantee safety when these spaces are modeled with improper or invalid assumptions, and an overly conservative model results in large delays and lower throughput. There is a trade-off, and a good feasible solution lies between these extremes. Almost all studies that consider reachable space for trajectory generation do not consider pedestrians [32] or bicycles that are in their path as the movement of these objects are unpredictable. This is a complex problem to solve even without consideration of related concepts such as risk homeostasis or the trolley problem or the idea that the ego car driver's life is somehow more important [33].

Real-time trajectory generation must be accomplished with the simultaneous incorporation of the actions of other participants on the reachable space of the ego vehicle. The ideal active solution would be for the ego vehicle to communicate with every object in its trajectory and then determine which object(s) yield to which other objects. However, this solution is infeasible at this point in time because (a) the infrastructure to accomplish this does not yet exist and (b) even if bicycles and pedestrians can communicate with the ego vehicle, some of the objects in its path (e.g., a disabled vehicle, a fallen object on the road) may not be able to communicate in order to resolve any conflicts with respect to the ego vehicle's reachable space. A reasonably feasible active alternative is for the ego vehicle to communicate with as many other objects, which include other vehicles, that are in its reachable space to coordinate their relative movements and then generate the reachable space with the incorporation of estimated movement and location of the other objects.

A passive means is for the ego vehicle to make a reasonable guess about the other vehicles before it generates its reachable space. For example, the ego vehicle can track each of the other vehicles that have the potential to be in its path and then see if their acceleration or deceleration and location indicate that they are indeed aware of the ego vehicle's path and are generating their own response to avoid overlaps in reachable space at appropriate points in time. Such estimates on the action by other vehicles can be made through analysis of relevant patterns in historical data from this domain. For example, data used for such analyses could include attributes of the other vehicles such as their acceleration or deceleration rates, their average and relative (vs. the ego vehicle) speeds, their relative location with respect to that of the ego vehicle, the existence of other vehicles in their path and whether their reactions if any are made in response to accommodate these other vehicles, the number of vehicles in their path traveling in the same direction, the number of vehicles that are in the same general area but are traveling in the opposite direction, road visibility conditions (e.g., due to weather, time of day), presence of stationary object(s) in its path, among others. Once the patterns are generated through analysis of such data, these patterns can be used to develop ego vehicle reachable space that is less conservative and more accurate in general. While such a process accounts for other vehicles that have the ability to modify their speed in response to the presence of the ego vehicle, among others, it is difficult to accomplish the same by pedestrians and bicycles on the road since they are the weaker objects and they generally tend to accommodate and yield to motorized vehicles. From the ego vehicle's perspective, the movement of these objects (pedestrians, bicycles) are unpredictable and bicycles are rather difficult to identify since their silhouettes vary depending on what else is hanging (e.g., flag(s)) from the bicycle, what is worn by the bicycle rider, and the number of objects (e.g., people, pets, bags) that are on the bicycle. Overall, when an ego vehicle takes into consideration in real-time the response from all other vehicles on the road, both trajectory generation and generated reachable space will be less conservative and more efficient and effective for the entire set of vehicles on the road.

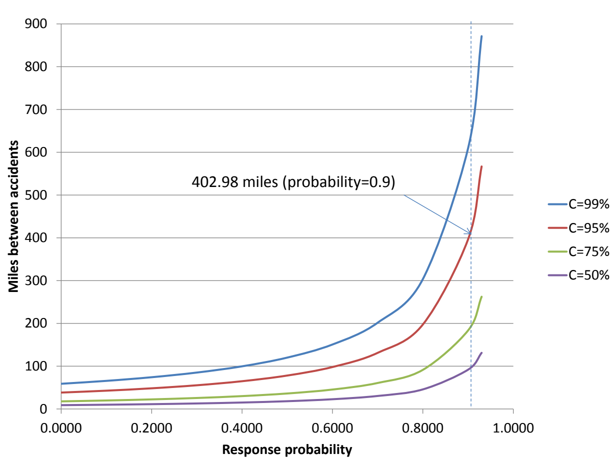

Figure 5: Other objects' response probability and miles between accidents

In some sense, an accident could be avoided when the ego vehicle actively takes necessary steps to prevent such an incident or the objects in its path somehow all act proactively to avoid colliding with the ego vehicle. To this end, we model the aggregate response of the objects (animals, bicyclists, other vehicles, and pedestrians) on the path of the ego vehicle as a single response probability. Response probability of zero models the scenario in which none of the objects on the path of the ego vehicle modifies its intended path in response to the ego vehicle's presence. Response probability value of one models the scenario in which all of the objects on the path of the ego vehicle successfully alter their intended path to avoid collision with the ego vehicle. As can be seen in Figure 5, response probability value of 0.9 corresponds to a similar distance between accidents as human driven vehicles. While this raises some hope, 0.9 is a very high probability and objects in an autonomous vehicle's path cannot be expected to appropriately respond with such high probability. Ideally, the design of an autonomous vehicle must be such that it is able to safely navigate without expecting anything of other objects in its path. Note that the act of avoiding collision with one another by itself could inadvertently lead to collisions when the response time is insufficient to avoid such an occurrence.

F. Appropriate and timely response

We considered the ability of vision systems in autonomous vehicles to detect objects in its path. What we have not considered is the ability of such a vehicle to make an appropriate response on time to avert any negative incident. Clearly, detection without an appropriate response is useless. Moreover, an appropriate response that is delayed or comes in too late for the intended purpose is not useful. For example, if an appropriate response to a detected object in a scene (e.g., an object in the ego vehicle's path) is coming to a stop, the ego vehicle cannot accomplish this act if detection came too late, the object is too close, and/or the ego vehicle is traveling at a speed that does not allow it to stop on time. The vision system should ideally be able to locate and identify the object(s) in its path with enough time and distance to spare for it to facilitate collision avoidance. Stopping distance therefore plays a critical role under such circumstances.

Stopping distance is the sum of reaction distance (i.e., distance traveled from the time the object was identified and the time a response was instantiated) and braking distance. While reaction distance is generally shorter in autonomous vehicles when compared to human-driven vehicles, it depends on several factors. For example, the camera frame rate is an important factor in this scenario wherein a higher frame rate translates to earlier identification of the object and therefore the ability to react sooner. It has also been shown [34] that higher frame rates could allow for simpler and efficient tracking algorithms such as Correlation Filter (CF) based methods with a tradeoff between significantly lower computational processing load per frame and need to process more frames.

To avoid an accident, the ego vehicle must (a) identify accident prone situations, (b) determine the most appropriate response for a given situation, and (c) instantiate a timely response. For successful execution of (c), both (a) and (b) need to occur with enough advance notice. Therefore, although correct identification of an impending accident situation is a necessary step, it alone is not sufficient to avoid an accident.

Take for example Luminar's Hydra LiDAR, which was unveiled at the Consumer Electronics Show in Las Vegas on 7 January 2020. Hydra is specifically targeted for use in highways and purports to track the road out to 80 meters, lanes out to 150 meters and objects to 250 meters. However, the braking distance of a typical vehicle that travels at 70mph is about 75 meters. Even with minimal response time, it is easy to see that the stopping distance could easily go beyond 80 meters (the road tracking distance of Hydra).

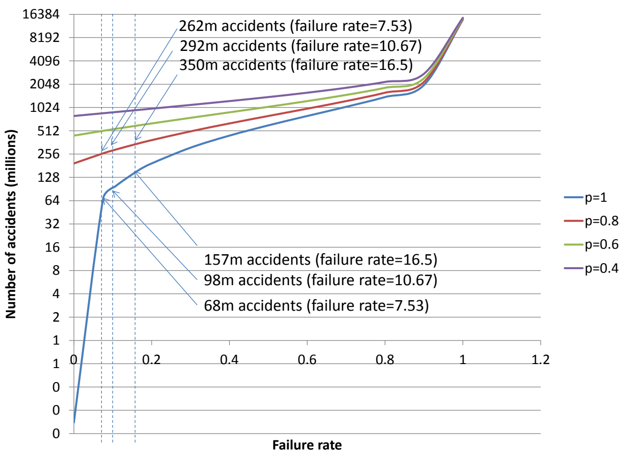

Figure 6: Failure rates and number of accidents in autonomous vehicles due to insufficient warning lead-time [C=95%]

From the above discussion, it is clear that knowledge lead-time plays a significant role in preventing accidents. It is not always the case that accidents are prevented upon correct accident prone situation identification. To this end, we consider a few different probabilities (p=1, 0.8, 0.6, 0.4) under which the system correctly identifies a situation and successfully gets out of an accident situation. The number of accidents corresponding to the three failure rates are marked only for p=1 and p=0.8 in Figure 6. As can be seen in this figure, even going from p=1 to p=0.8 more than doubles the number of accidents. Although in reality, the probability is context-specific and will vary across different incidents, p=0.8 is a conservative estimate. The 2017 statistic of 6.452 million accidents by human-driven vehicles can be matched only by a p value of 0.9927, which is high given the unpredictability of traffic conditions. Higher p values leave no chance to match human driving performance on average. In other words, autonomous vehicles have to identify an object correctly and respond with enough time and distance to spare about 99.27% of the times to match human performance.

IV. Discussion

Various studies over the years have documented the negative aspects of human vehicle drivers that portray them as distracted, drunk, lethargic, and irrational at times [35] [36]. Autonomous automobiles have been touted as a solution that does not entail the drawbacks that are generally associated with human drivers. With the advent and availability of a plethora of facilitating technologies, autonomous driving appears to be within reach, contingent upon resolution of identified deficits.

We summarize the results from our analysis in Table I, which shows the extent of deficit that must be addressed in current vision-based systems. Note that for human-driven vehicles, the performance value includes all possible weather and traffic conditions, whereas for the autonomous vehicle we only consider fair weather conditions. To match human-driven vehicle performance in terms of average number of miles between pairs of accidents, autonomous vehicles have to lower their system failure rate to 0.741 from 7.53 for the easy case. In terms of the number of accidents, it was 6.452 million for human-driven vehicles which corresponds to a system failure rate of 0.7407. For autonomous vehicles, it is 68 million (failure rate of 7.53), 98 million (failure rate of 10.67), and 157 million (failure rate of 16.5) respectively for the easy, moderate and hard cases. To avoid accidents with the ego autonomous vehicle, other objects have to react correctly 90% of the time for the ego vehicle to match human-driven vehicle performance. From the ego autonomous vehicle perspective, it has to accurately respond on time to allow for enough distance and time in order to avoid accidents with a probability of 0.9927. These results show that the performance of vision systems in autonomous vehicles is below par when compared against that of human-driven vehicles.

Table I presents the *potential* for accidents in autonomous driving vehicles with current state-of-the-art vision technology and machine learning algorithms. Results on the miles between accidents and number of accidents illustrate the 10-fold difference between human-driven and autonomous vehicles for the easy cases. For the hard cases, the difference is 20-folds in order. As the ego vehicle has minimal influence on what the other objects on its path do, it needs to rely on its own ability to avoid accidents. The 90% response probability expected of the other objects in the ego vehicle's path is therefore unrealistic. The ego vehicle must be able to accurately identify objects in its path (99.27%) *on time* for it to avoid an accident in parity with human-driven vehicles. These results clearly point to the need to focus on the ego vehicle's ability to simultaneously (1) recognize accurately an imminent accident situation, (2) actuate appropriate response, and (3) do these in a timely manner to avoid that accident. Among these, (1) involves vision and machine learning components, (2) involves the vehicle's mechanical components, and (3) mostly depends on the ability of the machine learning component to learn to quickly recognize so (2) can be executed in a timely manner. An accident points to deficits in at least one of these three steps. These steps are interconnected as (1) requires knowledge on the strengths and limitations that correspond to (2) for actionable intelligence that is instantiated on time. Our results show the extent of deficit in current vision-based systems in autonomous driving vehicles that must be addressed to reach parity with human-driven vehicles for accident proneness.

We did not consider the number of accidents that solely had fatalities, injuries, and only property damage since data

Performance criterion	Human-driven vehicle	Autonomous
Miles between accidents	402.98 (0.741)	38 (7.53)
Number of accidents [easy]	6.452 million overall	68 million (7.53)
Number of accidents [moderate]	6.452 million overall	98 million (10.67)
Number of accidents [hard]	6.452 million overall	157 million (16.5)
Other object response probability for parity	n/a	0.9
Ego vehicle response probability to avoid accident	n/a	0.9927

TABLE I. Summary of results for $C = 95\%$ [failure rates in parentheses]

for such events are not available. As we are interested in the overall reliability of autonomously driven vehicles for this study, separation of accidents into those that had fatalities, injuries, and only property damage is not paramount.

We set out to answer questions on whether current vision sensors and machine learning algorithms are sufficient for autonomous vehicle accident rates to be below that of human-driven vehicles and what (vision sensors+algorithm) performance is necessary for parity. The data and algorithms used for learning to identify objects in the path of the vehicle play a significant role in accident avoidance. There have been initiatives to develop data sets that help with this process. For example, Barnes et al. [37] develop a dataset with "over 240000 scans from a Navtech CTS350-X radar and 2.4 million scans from two Velodyne HDL-32E 3D LiDARs; along with six cameras, two 2D LiDARs, and a GPS/INS receiver" with specific emphasis on the Frequency-Modulated Continuous-Wave (FMCW) class of radar that "provides a 360°-view of the scene and is capable of detecting targets at ranges far exceeding those of automotive 3D LiDAR." The algorithms used in these applications are generally some variant of CNN. While lower classification accuracy might be inefficient, ineffective, or just inconvenient in a significant number of applications, data and algorithms that synergistically work together to achieve close to perfect classification accuracy is necessary in automated driving systems as anything less could precipitate in human fatalities. There are also initiatives for more complicated annotations such as segmentation tracking.

One such is Caesar et al. [38] who develop an annotated data set comprising 1000 scenes from a fully autonomous vehicle sensor suite with 6 cameras, 5 radars and 1 LiDAR. Their average precision (AP) results based on 2D center distance on the ground plane are 85% for cars, 82% for pedestrians, and 37% for bicycles. Among the sensors used, their AP results for LiDAR data processed through PointPillars and camera data processed through image detector respectively were 68.4% and 47.8% for cars, 59.7% and 37.0% for pedestrians, 1.1% and 24.5% for bicycles, and 30.8% and 48.7% for traffic cones. Based on these results, we see that the overall identification accuracy is lacking even with annotated data. There is a need to improve identification of pedestrians and bicycles as these are the most vulnerable objects on the street. Similarly, identification under inclement weather conditions also need to be improved with appropriate sensors and learning algorithms.

We considered the commonly used vision system that includes camera, LiDAR, and RADAR for autonomous driving applications and showed that the current state-of-the-art of these vision-related devices and related data processing algorithms may not be sufficient for safe operationalization of autonomous vehicles. We now discuss a few other sensors that could possibly be used in consort with the ones that are already in such vehicles to improve their vision sense.

It is believed that the eyes of some living beings (e.g., frogs) only detect motion, perhaps to utilize the vision facet of their sensor suite efficiently. The relatively new biologically-inspired *event cameras* operate in a similar manner to capture motion in a scene through detection of changes in pixel-level light intensity [39] [40]. These cameras do not see anything if nothing changes in a scene. When even a single pixel-level light intensity changes, these cameras are capable of capturing that event with minimal blur and power consumption. Event cameras capture variations in brightness of every pixel in a scene in an asynchronous and independent manner. These cameras show great promise in autonomous driving applications. Thermal imaging infrared cameras are another option. These cameras create images using infrared radiation and are widely used in firefighting applications since they see through dark environments. These cameras identify infrared output differentials of objects whereby objects of the same temperature appear to be of the same color. However, the images produced by these cameras have poor depth perception. When the entire object of interest is at the same temperature, the object details are not visible in the image. While the latter may not be an issue for autonomous driving applications, the lack of depth perception is a serious drawback. Perhaps once an object is identified by this camera, the other devices in the autonomous vehicle's vision suite can take over.

V. Conclusion

As the world prepares for higher levels of autonomous driving, it helps to take stock of the state-of-the-art of sensors that facilitate vision and to determine if these sensors are sufficient for a given level of autonomous vehicle. We considered the roadway accident statistics from 2017 based on human-driven vehicles and compared that against related statistics when human-driven vehicles are replaced by autonomous vehicles with current state-of-the-art vision technology. Based on our analysis, we show that we have a long way to go to achieve completely autonomously driven vehicles that are able to navigate any road condition anywhere in the world. Our results show that in addition to better situational awareness, the failure rate of human drivers is more than an order of magnitude lower than autonomous vehicles when compared across all driving environments. Human drivers are also less susceptible to adverse actions from active objects in reachable space and, for the most part, are able to correctly identify and respond in a timely manner to objects in the vehicle's path.

References

[1] X. Mosquet, T. Dauner, N. Lang, M. Ruessmann, A. Mei-Pochtler, R. Agrawal, F. Schmieg (2015) "Revolution in the Driver's Seat - The Road to Autonomous Vehicles," April, The Boston Consulting Group.

[2] S.K. Gehrig, F.J. Stein (1999) "Dead reckoning and cartography using stereo vision for an autonomous car," *Proceedings of International Conference on Intelligent Robots and Systems*, 3, 1507-1512.

[3] SAE International (2016) "Automated Driving." https://www.sae.org/standards/content/j3016_202104/preview/

[4] Autonomous Vehicle Testing Permit Holders (2021) www.dmv.ca.gov/portal/vehicle-industry-services/autonomous-vehicles/autonomous-vehicle-testing-permit-holders/

[5] Y. Wang, W.-L. Chao, D. Garg, B. Hariharan, M. Campbell, K.Q. Weinberger (2020) "Pseudo-LiDAR from Visual Depth Estimation: Bridging the Gap in 3D Object Detection for Autonomous Driving." 22 February. https://arxiv.org/pdf/1812.07179.pdf

[6] L. Gomes (2014) "Hidden Obstacles for Google's Self-Driving Cars: Impressive Progress Hides Major Limitations of Google's Quest for Automated Driving." *MIT Technology Review*, August 28.

[7] The Milwaukee Sentinel (1926) "'Phantom auto' will tour city." 8/12,4

[8] J. McCarthy (1968) "Computer Controlled Cars." http://www-formal.stanford.edu/jmc/progress/ cars/cars.html

[9] N. Kalra, S.M. Paddock (2016) "Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?" *Transportation Research - Part A*, 94, 182-193.

[10] J. Stout (2017) "How simulation turns one flashing yellow light into thousands of hours of experience," https://medium.com/waymo/simulation-how-one-flashing-yellow-light-turns-into-thousands-of-hours-of-experience-a7a1cb475565

[11] NTSB (2018) *Preliminary Report Highway HWY18MH010* https://www.ntsb.gov/investigations/AccidentReports/Reports/HWY18MH010-prelim.pdf

[12] Volvo Cars USA (2015) "US urged to establish nationwide Federal guidelines for autonomous driving." October 7. https://www.media.volvocars.com/us/en-us/media/pressreleases/167975/us-urged-to-establish-nationwide-federal-guidelines-for-autonomous-driving

[13] N. Zhou (2017) "Volvo admits its self-driving cars are confused by kangaroos." *The Guardian*, 30 June. https://www.theguardian.com/technology/2017/jul/01/volvo-admits-its-self-driving-cars-are-confused-by-kangaroos

[14] E. Arnold, O. Y. Al-Jarrah, M. Dianati, S. Fallah, D. Oxtoby, A. Mouzakitis (2019) "A Survey on 3D Object Detection Methods for Autonomous Driving Applications." *IEEE Transactions on Intelligent Transportation Systems*, 20(10), 3782-3795.

[15] S. Chadwick, W. Maddern, P. Newman (2019) "Distant Vehicle Detection Using Radar and Vision." *Proceedings of the International Conference on Robotics and Automation (ICRA)*, 8311-8317.

[16] A. Singandhupe, H.M. La (2019) "A Review of SLAM Techniques and Security in Autonomous Driving." *Proceedings of the International Conference on Robotic Computing (IRC)*, 602-607.

[17] Y. Kang, H. Yin, C. Berger (2019) "Test Your Self-Driving Algorithm: An Overview of Publicly Available Driving Datasets and Virtual Testing Environments." *IEEE Transactions on Intelligent Vehicles*, 4(2),171-185

[18] M. Yang, S. Wang, J. Bakita, T. Vu, F.D. Smith, J.H. Anderson, J.-M. Frahm (2019) "Re-Thinking CNN Frameworks for Time-Sensitive Autonomous-Driving Applications: Addressing an Industrial Challenge." *Proceedings of the IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS)*, 305-317.

[19] P. Sun, H. Kretzschmar, X. Dotiwalla, A. Chouard, V. Patnaik, P. Tsui, J. Guo, Y. Zhou, Y. Chai, B. Caine, V. Vasudevan, W. Han, J. Ngiam, H. Zhao, A. Timofeev, S. Ettinger, M. Krivokon, A. Gao, A. Joshi, Y. Zhang, J. Shlens, Z. Chen, D. Anguelov (2020) "Scalability in Perception for Autonomous Driving: Waymo Open Dataset." *Proceedings of Computer Vision and Pattern Recognition (CVPR)*, 2446-2454.

[20] H. Shin, D. Kim, Y. Kwon, Y. Kim (2017) "Illusion and Dazzle: Adversarial Optical Channel Exploits Against Lidars for Automotive Applications." *Proceedings of the Conference on Cryptographic Hardware and Embedded Systems (CHES)*, Springer LNCS 10529, 445-467.

[21] K. Eykholt, I. Evtimov, E. Fernandes, B. Li, A. Rahmati, C. Xiao, A. Prakash, T. Kohno, D. Song (2018) "Robust physical-world attacks on deep learning models," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1625-1634.

[22] I. Goodfellow, P. McDaniel, N. Papernot (2018) "Making Machine Learning Robust Against Adversarial Inputs." *Communications of the ACM*, July, 61(7), 56-66.

[23] X.Boucherat (2018) "Could LiDAR be the key to five nines reliability?" *Automotive World*, 23 May. www.automotiveworld.com/articles/lidar-key-five-nines-reliability/

[24] NHTSA (2017) *Traffic Safety Facts Annual Report Tables* https://cdan.nhtsa.gov/tsftables/tsfar.htm

[25] AAA (2019) *American Driving Survey, 2014 - 2017*. aaafoundation.org/wp-content/uploads/2019/02/18-0783_AAAFTS-ADS-Brief_r8.pdf

[26] NACTO (2019) *Blueprint for Autonomous Urbanism*, 2nd edition. https://nacto.org/publication/bau2/

[27] A. Kleyner (2014) "How Stress Variance in the Automotive Environment will Affect a 'True' Value of the Reliability Demonstrated by Accelerated Testing," *SAE International Journal*, 7(2), 552-559.

[28] M. Blanco, J. Atwood, S. Russell, T. Trimble, J. McClafferty, M. Perez (2016) "Automated Vehicle Crash Rate Comparison Using Naturalistic Data." Virginia Tech Transport Institute.

[29] K. Driggs-Campbell, R. Dong, S. Sastry, R. Bajcsy (2018) "Robust, Informative Human-in-the-Loop Predictions via Empirical Reachable Sets," *IEEE Transactions on Intelligent Vehicles*, 3(3), 300-309.

[30] T. Bjørnskau (2017) "The Zebra Crossing Game - Using Game Theory to Explain a Discrepancy between Road User Behaviour and Traffic Rules," *Safety Science*, 92, 298-301.

[31] B. Groza (2017) "Traffic models with adversarial vehicle behaviour" https://arxiv.org/abs/1701.07666

[32] A. Millard-Ball (2018) "Pedestrians, Autonomous Vehicles, and Cities," *Journal of Planning Education and Research*, 38(1), 6-12.

[33] M. Taylor (2016) "Self-Driving Mercedes-Benzes will Prioritize Occupant Safety over Pedestrians," *Car and Driver*, 7 Oct. https://www.caranddriver.com/news/self-driving-mercedes-will-prioritize-occupant-safety-over-pedestrians

[34] H.K. Galoogahi, A. Fagg, C. Huang, D. Ramanan, S. Lucey (2017) "Need for Speed: A Benchmark for Higher Frame Rate Object Tracking," *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 1125-1134.

[35] Vanderbilt, T., B. Brenner (2008) *Traffic: Why We Drive the Way We Do (and what it says about us)*, Alfred A. Knopf, New York.

[36] WHO (2018) *Global Status Report on Road Safety*. https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/

[37] D. Barnes, M. Gadd, P. Murcutt, P. Newman, I. Posner (2020) "The Oxford Radar RobotCar Dataset: A Radar Extension to the Oxford RobotCar Dataset." https://arxiv.org/abs/1909.01300

[38] H. Caesar, V. Bankiti, A.H. Lang, S. Vora, V.E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan, O. Beijbom (2020) "nuScenes: A multimodal dataset for autonomous driving." *Computer Vision and Pattern Recognition (CVPR)*.

[39] G. Gallego, T. Delbrück, G. Orchard, C. Bartolozzi, B. Taba, A. Censi, S. Leutenegger, A. Davison, J. Conradt, K. Daniilidis, D. Scaramuzza (2019) "Event-based Vision: A Survey," https://arxiv.org/pdf/1904.08405.pdf

[40] E. Yurtsever, J. Lambert, A. Carballo, K. Takeda (2020) "A Survey of Autonomous Driving: Common Practices and Emerging Technologies." *IEEE Access*. DOI: 10.1109/ACCESS.2020.2983149