

Perception Simplex: Verifiable Collision Avoidance in Autonomous Vehicles Amidst Obstacle Detection Faults

Ayoosh Bansal¹ | Hunmin Kim² | Simon Yu¹ | Bo Li¹ | Naira Hovakimyan¹ | Marco Caccamo³ | Lui Sha¹

¹University of Illinois Urbana-Champaign, Illinois, USA

²Mercer University, Georgia, USA

³Technical University of Munich, Munich, Germany

Correspondence

Ayoosh Bansal

Email: ayooshb2@illinois.edu

Funding Information

The material presented in this paper is based upon work supported by the National Science Foundation (NSF) under grant no. CNS 1932529, ECCS 2020289, the Air Force Office of Scientific Research (AFOSR) under grant no. #FA9550-21-1-0411, the National Aeronautics and Space Administration (NASA) under grant no. 80NSSC22M0070, 80NSSC20M0229, AWD-000577-G1, and University of Illinois Urbana-Champaign under grant no. STII-21-06. Marco Caccamo was supported by an Alexander von Humboldt Professorship endowed by the German Federal Ministry of Education and Research. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and don't necessarily reflect the views of the sponsors.

Abstract

Advances in deep learning have revolutionized cyber-physical applications, including the development of Autonomous Vehicles. However, real-world collisions involving autonomous control of vehicles have raised significant safety concerns regarding the use of Deep Neural Networks (DNN) in safety-critical tasks, particularly Perception. The inherent unverifiability of DNNs poses a key challenge in ensuring their safe and reliable operation.

In this work, we propose Perception Simplex (\mathcal{PS}), a fault-tolerant application architecture designed for obstacle detection and collision avoidance. We analyze an existing LiDAR-based classical obstacle detection algorithm to establish strict bounds on its capabilities and limitations. Such analysis and verification have not been possible for deep learning-based perception systems yet. By employing verifiable obstacle detection algorithms, \mathcal{PS} identifies obstacle existence detection faults in the output of unverifiable DNN-based object detectors. When faults with potential collision risks are detected, appropriate corrective actions are initiated. Through extensive analysis and software-in-the-loop simulations, we demonstrate that \mathcal{PS} provides predictable and deterministic fault tolerance against obstacle existence detection faults, establishing a robust safety guarantee.

KEYWORDS:

autonomous vehicles, software reliability, fault tolerance, cyber-physical systems, obstacle detection

1 | INTRODUCTION

Autonomous Vehicles (AV) hold great promise in enhancing road safety and potentially saving lives. However, despite the rapid progress and widespread deployment of AV prototypes, the road to achieving collision-free AV remains uncertain [2]. Instances of fatal collisions involving autonomous control of vehicles have raised significant concerns [3–22]. These incidents underscore the continued need for research and development efforts, aimed at enhancing the safety and reliability of AV.

Abbreviations: \mathcal{PS} , perception simplex; AV, autonomous vehicles; DNN, deep neural networks; CPS, cyber-physical systems.

Extension from Prior Work: This work is a substantially revised and extended version of a paper presented at ISSRE 2022 [1]. The main difference is that while the prior work presented the *detectability* model for the verifiable obstacle detection algorithm, this work develops the system architecture to utilize the verifiable capabilities and provide deterministic fault tolerance. The major differences are (i) the design, analysis and implementation for \mathcal{PS} (§7); (ii) evaluation of \mathcal{PS} using software-in-the-loop simulation based on industrial simulators (§9); (iii) alleviate a constraint of the *detectability* model by accommodating for LiDAR beam energy attenuation in presence of air impediments (§5.2.1); (iv) expanded related work to address the recent state of the art and the new contributions (§3); (v) other editorial changes to most sections, especially the abstract, introduction, scope and motivation, and conclusion.

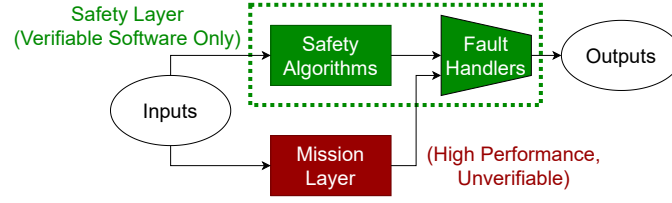


Figure 1 Perception Simplex design.

AV depend on deep learning techniques to perform various tasks, including perceiving the vehicle’s environment. However, Deep Neural Networks (DNN) operate by extrapolating correlations learned from training data, which inherently renders them unverifiable. As a result, their behavior becomes unpredictable and unreliable, rendering them unsuitable for deployment in safety-critical tasks [23–27]. Safety-critical software necessitates logical analysis and verification, requirements that current DNN solutions are not yet equipped to fulfill. The challenges are further compounded by the design of autonomous driving systems, where mission-critical tasks (navigation) and safety-critical tasks (collision avoidance) often overlap without clear delineation. This lack of separation leads to ill-formed optimization objectives for the system, and the safety problem becomes entangled and convoluted.

Hence, while DNN-based solutions are essential for fulfilling the mission of an AV, they also pose significant safety challenges. This research proposes a solution to this apparent dilemma by introducing **Perception Simplex (\mathcal{PS})**, which enables deterministic fault tolerance against perception faults in AV. To achieve this, we adopt a decoupled approach, similar to previous solutions in control systems [28, 29]. The \mathcal{PS} architecture consists of a safety layer and a mission layer. The safety layer incorporates logically analyzable and verifiable algorithms that focus on detecting faults and ensuring minimal functionality for safety-critical requirements. On the other hand, the mission layer houses the sophisticated self-driving software, *e.g.*, Apollo [30] or Autoware [31], which is responsible for fulfilling all aspects of the AV’s mission while trying to maintain safety. The proposed procedure within this architecture involves: (i) identifying the safety-critical requirements, (ii) utilizing verifiable algorithms to detect faults in meeting these requirements, and (iii) continuously monitoring for faults at runtime and initiating corrective actions when necessary. An overview of the \mathcal{PS} architecture is presented in Figure 1.

This work presents a brief survey of real-world fatal collisions associated with autonomous control of vehicles. The survey reveals that the most prevalent type of fault involved in these collisions is related to the detection of obstacle existence. To address this critical issue, we apply the \mathcal{PS} framework, which offers deterministic fault tolerance specifically targeting these obstacle existence detection faults. Additionally, we design \mathcal{PS} with a flexible and modular structure, enabling future expansions to address other perception faults. To address obstacle existence faults with \mathcal{PS} , we first carefully reduce the requirements for obstacle existence detection for collision avoidance. By focusing on only the essential elements necessary for this specific task, we eliminate several features of mission-critical object detection that are deemed unnecessary. This streamlined approach allows us to fulfill the task of obstacle existence detection while minimizing complexity. Furthermore, we conduct a detailed analysis of an existing LiDAR-based geometric algorithm called Depth Clustering [32, 33]. Through this analysis, we establish a comprehensive model, named the *detectability* model, which outlines the algorithm’s capabilities and limitations. The Depth Clustering algorithm is then integrated into the \mathcal{PS} pipeline, where the *detectability* model serves as the foundation for ensuring safety guarantees in collision avoidance scenarios. The key **contributions** of this work are:

- Reduce the complexity of safety-critical obstacle existence detection by carefully eliminating unnecessary object detection features, including object type classification (§4).
- Establish a *detectability* model for an existing LiDAR-based geometric obstacle detection algorithm, providing human perceptible bounds on capabilities and limitations (§6).
- Design, analyses, and implementation of the Perception Simplex (\mathcal{PS}) framework to achieve deterministic fault tolerance specifically targeting the perception fault of obstacle existence detection (§7).
- Evaluation of the verifiable algorithm using real sensor data [34] (§8) and software-in-the-loop simulation [30, 35, 36] demonstrating \mathcal{PS} ’s response in obstacle existence detection fault scenario (§9).¹

¹https://github.com/CPS-IL/perception_simplex

Table 1 Survey of fatal traffic collisions involving autonomous control of vehicles.

Reference	Autonomy Response [†]
[4]	A truck in the path of the vehicle was not detected and no evasive actions were initiated.
[5]	Changing the classification of a pedestrian led to the perception system ignoring their existence.
[6]	A crash attenuator in the path of the vehicle was not detected.
[8]	A white semi-trailer in the path of the vehicle was not detected.

[†]The comments brief the driver assist system's response during the incident and are not necessarily the causal faults for the incident.

2 | SCOPE AND MOTIVATION

Obstacle existence detection fault, *i.e.*, False Negative (FN) in perception is a grave safety concern [37]. A survey of fatal collisions involving AV (Table 1) points to recurring FN errors. Other fatalities involving AV [3, 7, 9–22], excluded from Table 1 due to the unavailability of public investigation reports, seem to follow this pattern. These underlying safety concerns are the primary challenge in adopting complete autonomy [38–41]. Therefore this work focuses on providing deterministic fault tolerance against obstacle existence detection faults. Causes, mechanisms, and adversarial inducement of such faults is an active area of research [42, 43]. However, these causes are out of the scope of this work, therefore, to evaluate \mathcal{PS} the faults are injected via instrumentation, as described in Section 9.1.

Analyzability and verifiability are the crucial components of the certification process of safety-critical systems [44, 45]. Verifiable algorithms, where the *causality* between the input parameters and the algorithm result can be established, are inherently suitable for safety-critical applications. This is in contrast to object detection DNN, trained using supervised learning, which effectively captures *correlations* between training input and labels. Consider the following simple example where y is obstacle height, x is obstacle distance from the AV, a and b are constant parameters based on LiDAR properties:

$$y \geq ax + b \quad (1)$$

If we want to establish that (1) is the *detectability* model, *i.e.*, when the condition in (1) is met, obstacles are always detected by the AV, the following must be defined:

Requirements: A definition of minimal requirements for what it means to *detect* an obstacle (§4).

Constraints: A set of well-defined constraints that must be met for the model to be applicable (§5).

Verification: A deterministic analysis verifying the *detectability* model (§6).

Let's assume that an AV safety standard requires that the AV be able to detect all obstacles of a minimum height of 10 *cm*. Further, let's assume all vehicles and structures on the road are also mandated to be taller than this minimum height by a road safety rule. Using LiDAR and AV parameters to determine a and b , the *detectability* model (1) can be used to determine the minimum distance at which such obstacles can be detected. This minimum distance, in conjunction with the braking capability of the AV, can then be used to determine the max speed at which the AV can safely travel. This example, admittedly simple, shows how a verified *detectability* model can bring together road safety rules, AV parameters, and AV safety policies to provide deterministic collision safety. We show in this work that such velocity limitations can be derived, as in Theorem 4, guaranteeing collision avoidance when complex mission-critical object detectors suffer obstacle existence detection faults.

It is assumed that all communication in the \mathcal{PS} system is real-time and lossless. Data shared between layers is assumed to be always fresh and valid for its use. Temporal safety *i.e.*, real-time execution and communication are important to \mathcal{PS} design, though have been the focus of prior works [46–48]. This work presents the functional design and safety guarantees of \mathcal{PS} .

While we limit this work to obstacle existence detection faults due to their importance and prevalence in real-world fatal collisions, \mathcal{PS} is designed to be extensible, to address other perception faults by utilizing verifiable capabilities to fulfill other safety-critical perception tasks, as exemplified by the handling of obstacle existence detection faults in this work. Furthermore, while the \mathcal{PS} is applied to autonomous ground vehicles in this work, the procedure for verifiable mitigation of faults in complex DNN, can be applied to other safety-critical systems reliant on DNN solutions.

3 | RELATED WORK

Redundancy is a crucial concept for designing reliable safety-critical cyber-physical systems (CPS). Full system replication enables a system to tolerate and correct random or systemic faults that an individual system cannot [49–54]. However, computational redundancy does not address faults shared by all replicas *i.e.* algorithmic or implementation faults [55, 56]. Functional or analytical redundancy for fault-tolerant control [57], addresses this by using different systems with different implementations, inputs, capabilities, or algorithms to fulfill the same function *e.g.*, GPS and inertial navigation.

Simplex architecture [28, 29, 58] and its evolutionary variations [59–61], including *PS*, are a form of asymmetric functional redundancy. This architecture safeguards systems that are controlled by a *high-performance* but potentially unreliable controller, switching to a reliable, *high-assurance* controller when needed, in order to ensure the safety of the system. Although solutions have been proposed to tackle control [62] and planning [63] in autonomous vehicles and to account for uncertainty in position and velocity measurements [64], these approaches do not directly address the critical issue of gross lapses in perception, *e.g.*, obstacle existence detection faults. *PS* is the first to propose a deterministic simplex implementation for fault tolerance against obstacle existence detection faults, by leveraging verifiable solutions for obstacle detection. *PS* is a modular solution designed to be open to future integration of other solutions for measurement uncertainty, planning, and control.

DNN verification, even for simple properties, is an NP-complete problem [65]. Gharib *et al.* [66] describe the need and current lack of verification methods for machine learning components used in safety-critical applications. Albarghouthi [67] described the challenges for neural network verification, including scale, complexity, and dynamism of the environment, all applicable for their use in Autonomous Driving. Liu *et al.* [68] survey existing verification methods and classify the verification methods into three categories: reachability, optimization, and search, identifying tradeoffs between scale and completeness of the verification methods. Even when the deep networks can be verified, it is done so for small input sets only [69]. Improving the input coverage by exploring around available inputs by adding adversarial or context-specific distortions [70, 71] is insufficient in the face of the vastness of the input sets for real-world problems, like perception.

Classical Obstacle Detection approaches have been developed using various sensors. Long-range obstacle detection based on stereo cameras was proposed by Pinggera *et al.* [72]. LiDAR sensors have shown incredible promise for their use in autonomous driving [73]. This has, in turn, accelerated the improvements in LiDAR sensor technology, increasing the range, scanning frequency, and resolution of the sensor [74]. LiDAR sensors suffer a linear error growth with distance as compared to quadratic for stereo camera [75]. These factors made the LiDAR sensor the natural choice for this work. Various LiDAR-based geometric algorithms were considered as part of this work [76–79]. Each algorithm has a similar flow; identifying points on the ground vs. obstacles, followed by clustering, segmentation, and/or classification. Depth Clustering [32, 33] is chosen as the primary example due to its deterministic explainable behavior, flexible parameters to optimize tradeoffs, and public availability of an efficient implementation with extremely low runtime of 40 *ms* on an embedded Jetson Xavier platform, using a single CPU core only, for a point cloud with 169,600 points [80].

Sensor Fusion is used in AV [81] to alleviate limitations of individual sensor types. However, the intermingling of safety and performance optimizations, limitations of the algorithms used to interpret sensor inputs, and unforeseen interactions between features unnecessary for safety and optimizations for performance have resulted in perception lapses and collisions [4–6, 8]. To address these issues, the *PS* safety layer uses algorithms that have deterministic capabilities, constraints, and limits. However verifiable algorithms using different sensor data types would help alleviate the constraints of individual algorithms. The use of sensor fusion in *PS* is left for future work.

Safe Autonomous Driving has been an active area of research with varying approaches. System architectures have been proposed that make parts of the AV software have better security and execution time predictability [48, 82–84]. While this addresses the response latency variance for the AV, it does not address the faults born out of complex algorithms in the software pipelines. Cooperative approaches [85] do not address the mixed autonomous and manually driven vehicle conditions. Techniques to improve driving policies [86, 87], path planning [88–90], risk assessment [91–95], steering or braking control [96–100] are still vulnerable to perception failures. *PS* addresses the perception failures first and is flexible to the inclusion of advancements in the mission layer.

Summary. *PS* integrates the design philosophies of various prior safety architectures for AV while providing targeted safety guarantees and the flexibility of incremental development making it practical and easy to include in existing AV software stacks. Existing approaches for safe control can be housed within this architecture to expand the capabilities and fault handling of the safety layer. This work provides a path to verifiable fulfillment of safety properties in AV, starting from Perception.

4 | MINIMAL REQUIREMENTS

In this section, we define the minimally sufficient, though not necessary and sufficient, requirements for safety-critical obstacle detection and collision avoidance. These requirements define a minimized set of features of an obstacle that the AV perception system should perceive to establish the existence of the obstacle. While some of the following observations are already well established, this work is the first to form a minimally sufficient set for safety-critical collision avoidance. Further, some requirements like object classification are typically considered crucial for all aspects of perception in AV, however, we show that while it is valuable for mission-critical requirements, it is not for safety-critical collision avoidance.

4.1 | Classification

We will use the mathematical problem formulation of reach-avoid control to justify why we do not need to classify obstacles for collision avoidance purposes. The current section is motivated by [101, 102] but we simplify the notations and the problem in a deterministic setting while the references consider stochastic problem formulations. Let $x_k \in X$ be the 2D-position of the vehicle at discrete time instance $k \geq 0$, where $X \subseteq \mathbb{R}^2$ is the set of the state space. The set $D \subset X$ is the destination, and the sets $O_i(k) \subset X$ for $i \in M_O \triangleq \{1, 2, \dots, M\}$ are obstacles, where $O_i(k)$ represents the i^{th} obstacle and M_O is the index set for the obstacles. The set $O_i(k)$ could be time-varying but it is assumed that the obstacles do not overlap with the destination, *i.e.*, $\cup_{i \in M_O} (D \cap O_i(k)) = \emptyset \forall k$. For brevity, further mentions of $O_i(k)$ are reduced to O_i . Now, the reach-avoid control problem is to drive the vehicle to the destination D while avoiding obstacles O_i for $i \in M_O$ within finite time horizon N , given initial condition x_0 . The success of this mission can be characterized by the following index κ :

$$\begin{aligned} \kappa &\triangleq \sum_{j=0}^N (\prod_{i=1}^M \prod_{t=0}^{j-1} \mathbf{1}_{X \setminus O_i}(x_t)) \mathbf{1}_D(x_j) \\ &= \begin{cases} 1, & \text{if } \exists j \in \{0, 1, \dots, N\} : x_j \in D \wedge \forall t \in \{0, 1, \dots, j-1\} : x_t \in \cap_{i=1}^M X \setminus O_i \\ 0, & \text{otherwise,} \end{cases} \end{aligned} \quad (2)$$

where $\mathbf{1}_S(\cdot) : X \rightarrow \{0, 1\}$ is the indicator function for a set S , and \wedge is the logical AND. In short, $\kappa = 1$ if and only if the objective is achieved. The index κ in (2) can be used as a cost function of the optimal control problem as in [101, 102]. Therefore, one is required to evaluate the indicator functions in (2), which means that it is required to know the set D and $X \setminus O_i$ for $i \in M_O$.

However, index κ in (2) can be equivalently formulated as

$$\kappa = \sum_{j=0}^N (\prod_{t=0}^{j-1} \mathbf{1}_{X \setminus (\cup_{i \in M_O} O_i)}(x_t)) \mathbf{1}_D(x_j) \quad (3)$$

and this expression can be used instead for the optimal reach-avoid control problem. The formulation (3) indicates that one could also address the control problem only with knowing $\cup_{i \in M_O} O_i$, but not individual obstacle sets O_i , *i.e.*, one does not need to classify/distinguish individual obstacles for the reach-avoid control purpose.

It should be noted that classification adds valuable information that supports advanced features like predictive tracking, and maneuver planning and improves AV performance. The argument here is only that object classification is not a necessity for obstacle avoidance and therefore not a part of the minimal requirements for safety-critical obstacle detection.

4.2 | Collision Risk

Collision avoidance involves many factors, from perception to vehicular control. While the dynamics and ethics of collision avoidance are complex [103], the safety-critical obstacle detection system is required to detect all obstacles that can potentially collide with the AV. We utilize a physical model for collision risk from our prior work [104], where the potential risk of collision is determined by the overlap of the existence regions [105] of obstacles and AV within the AV's time to stop.

4.3 | Height

The height of an obstacle is only useful in making a binary determination for collision avoidance, *i.e.*, whether or not the obstacle is completely clear above the height of the AV. For example, in Figure 2 (a) the overhead road sign ① is completely above the AV, its exact height has no implication for collision avoidance. The box ② contains valuable information that is required to identify the obstacle within the box to be a vehicle, however as discussed in Section 4.1, such a recognition of the obstacle class is not a requirement for collision avoidance. Thus for safety-critical collision avoidance ③ contains as much relevant

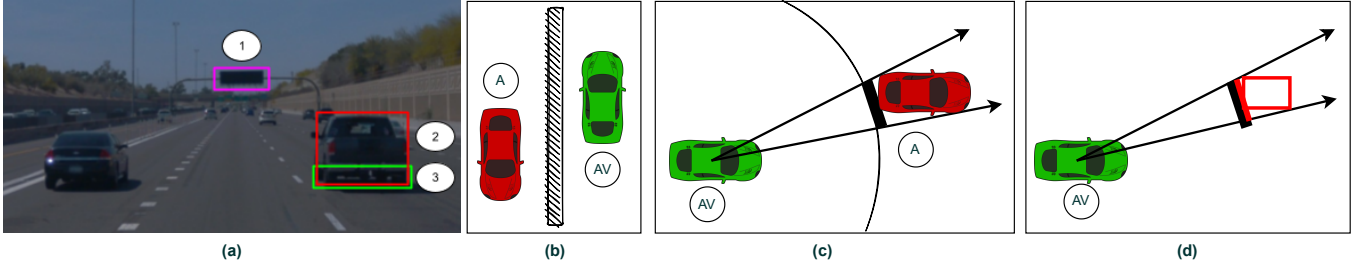


Figure 2 Example scenes. (a) Role of obstacle height. (b) Distance and Occlusion of Obstacles. (c) Projection of Obstacles.

information as ②. Therefore as long as an obstacle's height is not erroneously detected to be above and clear of the AV, we can simply consider the top or bird's eye view of the AV's surroundings. While such a view of obstacles is not traditionally used in perception systems, however, path planning in AV, an inherently 2D problem, regularly uses this representation [106].

4.4 | Distance

The distance to an obstacle must be accurately detected for the collision-free operation of the AV. For collision avoidance, this distance is the minimum distance between the perimeters of the obstacle and the AV. Many safety parameters like safe following distance, time to collision, and time to stop, are a function of the distance between the AV and obstacles [96, 99], and therein lies the difference in underestimation and overestimation error in distance detection.

4.4.1 | Underestimation Error

When obstacle distances are underestimated, *i.e.*, obstacles are detected to be closer than they are, the performance of AV suffers, however, there is no negative impact on direct collision avoidance. An implicit effect of acceptance of all detection with a lower distance than the ground truth obstacle is that occluded obstacles become unnecessary to detect as long as the occluding obstacle is detected. An example of this is Figure 2 (a) and (b). A small wall exists to the left side of the AV in both images, separating traffic moving in the opposite direction to the AV. If the wall is detected as an obstacle, the AV would not need to detect the vehicles across the wall to avoid colliding with them. Avoiding collisions with the occluding obstacle (wall here) means avoiding collisions with occluded obstacles. Note that this only applies to completely occluded obstacles. Partial occlusion is discussed in Section 4.5.

4.4.2 | Overestimation Error

Overestimation of distance is a grave safety concern. Obstacles detected to be further than they are, invalidate any safety decisions that would be based on this information. Since it is unrealistic that an obstacle detection system would always have no distance overestimation error, a strict bound on overestimation distance detection error is required. With an established max error bound, any distance-based safety guarantee can assume this error is always present, maintaining the safety guarantee in the worst case.

4.5 | Projection

We determine a projection signifying a line on the road that the AV cannot cross without colliding with the obstacle for each obstacle. Figure 2 (c) shows an example of this. A circle is drawn with its center at the AV's sensor hub and radius equal to the closest point on the obstacle from AV. The projection of the obstacle is then found as a line segment, tangent to the point where the above circle touches the obstacle. This line segment is shown as a thick line segment in Figure 2 (c). This line segment is a representation of the obstacle. Note that as new sensor inputs come in over time, the AV and obstacle move relative to each other, and the projection moves accordingly. So the line segment only represents the obstacle in the current frame to make safety-critical decisions until the next sensor frame is received and processed. With this minimal representation of the ground truth obstacles, containing only the minimal information about the obstacles as required for safety-critical collision avoidance, we can now determine when detected obstacles are sufficiently detected to avoid collisions.

4.6 | Coverage

Each detected obstacle that meets the distance criteria (§4.4) and falls within the direction of an obstacle can now be used to determine if a ground truth obstacle is sufficiently covered by detected obstacles to avoid collisions with the ground truth obstacle. Detected obstacles are projected on the projection of the ground truth obstacle to determine what parts of the ground truth projection are covered by the detected obstacles, as shown in Figure 2 (d). The projected detection must provide enough information to avoid collisions for an obstacle to be considered as detected. As with distance (§4.4) this coverage may not be perfect but should have bounded errors. A limited proportion of the ground truth projection must be covered. This is equivalent to the traditionally used intersection over union (IOU) limits. It should be noted that the error margin of coverage is less important than that of the distance. The minimum distance always tracks the distance between the closest points between the AV and the obstacle, changing with each sensor input frame.

4.7 | Summary

In this section, we have discussed various features of obstacles a perception system may detect and detailed their relevance for safety-critical collision avoidance. Many properties of obstacles that are considered critical parts of perception in AV (e.g., class, 3D dimensions, road sign information), while indeed required for achieving the mission of autonomous driving, are not necessary for safety-critical collision avoidance. In brief, an obstacle is considered detected for safety-critical obstacle avoidance if the following are accurately determined (a) the distance of the obstacle from AV, within bounded error margins. (b) a line on the road that the obstacle makes unsafe for the AV to cross. An obstacle detection system that reliably meets the above requirements can be used to detect safety-critical faults in perception by more complex perception systems, thus providing fault detection and collision avoidance system. The fusion of verifiable algorithms with DNN and the control decisions based on the verifiable algorithms have additional challenges that will be addressed in future works.

5 | PARAMETERS AND CONSTRAINTS

5.1 | LiDAR Parameters

A rotating LiDAR contains several lasers stacked vertically. Each laser is given an index $r \in \{1, 2, \dots, N\}$, where N is the number of lasers. Each laser has an inclination angle ξ_i from the horizontal, positive below the horizontal, the set of which is represented by ξ . The lowest laser below the horizontal is assigned index $r = 1$ and $r = N$ index represents the highest laser, usually inclined above the horizontal. The LiDAR is mounted on the vehicle at a height H_L above the wheelbase. For rotating LiDAR, the horizontal step angle Ψ is uniform and can be calculated as $360^\circ / \text{SamplesPerRotation}$. The sensor is considered the origin point for the coordinate system. At each rotation, a new column of N samples is recorded and indexed with c . The sensor returns a range image, a 2D depth image of range values $R_{r,c} \forall r \in \{1, 2, \dots, N\}, c \in \{1, 2, \dots, \Psi\}$. While we assume the more prolific rotating LiDAR, a solid state LiDAR [74] would have similar parameters. Table 2 and Figure 3, respectively, summarize and represent some of these parameters, using example values from real-world dataset [34].

5.2 | Constraints

We assume the following constraints for the validity of the *detectability* model in Section 6:

C1: All LiDAR beams encountering solid obstacles, including ground, within the max range of the LiDAR, return with strong enough intensity from the first obstacle they encounter so that the return is recorded. This assumption holds in real-world except when; (a) there are physical impediments in the air like dust, smoke, fog, or rain; or, (b) adversarial objects with extremely reflective, absorbent, or transparent surfaces. This constraint is partially alleviated by accommodating for LiDAR beam energy attenuation resulting in reduced effective LiDAR range, due to air impediments, as described in Section 5.2.1.

C2: The obstacle's width, projected on a plane perpendicular to the LiDAR beams falling on the obstacle, must be enough to have LiDAR points returned from it. For example, a thin sheet aligned parallel to the LiDAR beam's direction can never be detected. According to the formula for circle arc length, a conservative bound can be established:

$$\text{Width Projected} \geq \frac{\Psi D \pi}{180^\circ}. \quad (4)$$

Table 2 Symbols Summary

Symbols	Description	Example Values [‡]
N	Count of lasers in a vertical array	64 [34], 32 [35]
ξ	LiDAR Beam Angles set	$\{-2.4^\circ, \dots, 17.6^\circ\}$ [34], $\{-10.7^\circ, \dots, 30.7^\circ\}$ [35]
H_L	Height of LiDAR sensor	2.184 m [34], 2.312 m [35]
Ψ	LiDAR horizontal angle step	0.136° [34], 1° [35]
D_{min}^{LiDAR}	Distance to first LiDAR return	6.7 m [34], 3.9 m [35]
$R_{max,clear}^{LiDAR}$	Distance range of the sensor in clear conditions	75 m [34], 100 m [35]
R_{max}	Distance range of the sensor in current conditions	
α_o	Inclination of obstacle from ground	$\{0^\circ, \dots, 8^\circ\}$ [107]
h_o	Height of obstacle	Variable
α_{th}	Threshold angle for ground removal	10°

[‡]Where different, values are provided for both evaluation setups (§8, §9).

Since Ψ is small, the arc length is approximately equal, though always greater than the chord length for the same points. Being a small value, this width is not a prohibitive constraint, *e.g.*, using values from Table 2 for Waymo Open Dataset [34], the width constraint at max LiDAR range of 75 m is just 17 cm.

C3: At least one LiDAR point exists on the ground before the obstacle. This is an algorithmic constraint [33]. While we focus on the primary Top LiDAR only in the rest of this work, additional low height limited field of view (FOV) LiDAR sensors can be used for obstacles closer than D_{min} , *e.g.*, Front, Rear, Side-Left and, Side-Right LiDAR in Waymo Open Dataset [34].

Furthermore, the following are assumed initially (§6.1) and relaxed in later sections:

A1: Obstacle touches the ground at 90° , *i.e.*, $\alpha_o = 90^\circ$. Relaxed in Sections 6.2 and 6.3.

A2: There is no ground inclination change between the vehicle and the obstacle. Relaxed in Section 6.4.

A3: No noise in detected range. Relaxed in Section 6.5.

5.2.1 | LiDAR Beam Energy Attenuation

Particulate impediments in air, *e.g.*, smoke, dust, or fog, reflect and dissipate LiDAR beam energy cause false early or lost LiDAR beam returns [108–111]. The detection of these conditions is an orthogonal problem with significant prior solutions [112–116]. Note that we only consider benign faults, not LiDAR spoofing attacks [117–120], which can also violate constraint C1. We now consider the reduction in the effective range of LiDAR, from excessive attenuation of beam energy due to particulates in the air. The power of LiDAR attenuates per Beers-Lambert Law [121]:

$$\tau(R) = \frac{P(R)}{P(0)} = e^{-\sigma R} \quad (5)$$

where $P(R)$, $P(0)$ refer LiDAR power at distance R and 0, respectively. The value σ is the attenuation coefficient that characterizes how strongly a medium absorbs the wave signals. The attenuation coefficient is calculated by the Kruse model [122]

$$\sigma(\lambda) = \frac{17.35}{V} \left(\frac{\lambda}{0.55} \right)^{-q} \quad (6)$$

where V is the visibility which depends on air quality and conditions, such as fog and haze. The parameter λ is the wavelength of the signals in the micrometer. A typical LiDAR wavelength is either 905 nm or 1550 nm. The parameter q depends on visibility [123] as

$$q = \begin{cases} 1.6 & V > 50 \text{ km} \\ 1.3 & 6 \text{ km} < V < 50 \text{ km} \\ 0.16V + 0.34 & 1 \text{ km} < V < 6 \text{ km} \\ V - 0.5 & 0.5 \text{ km} < V < 1 \text{ km} \\ 0 & V < 0.5 \text{ km} \end{cases} \quad (7)$$

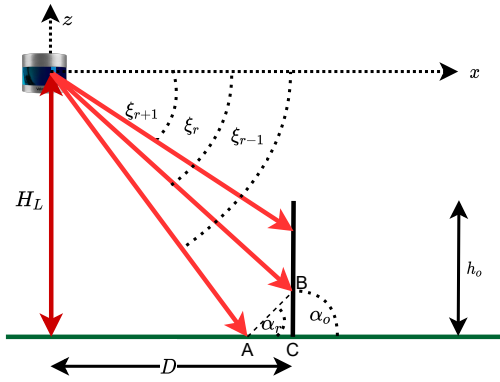


Figure 3 A representation of point segmentation to ground vs obstacle based on α thresholding when two points return from the obstacle.

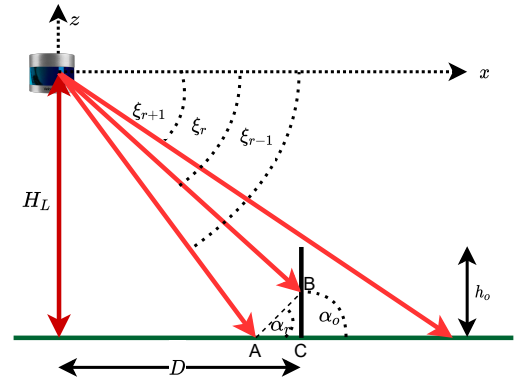


Figure 4 Representation for α thresholding with only one point on the obstacle.

Knowing the LiDAR's visibility under the current weather condition is critical because the short-sighted perception will delay the vehicle's response and limit the ability to respond to obstacles. We emphasize this limitation but leave the accurate calculation of the measurable range as future work because this is out of the current paper's scope and has been studied in prior works [124]. Instead, we provide a rule of thumb.

Regardless of the weather condition, minimum observable signal power remains unchanged for a LiDAR device, *i.e.*, there exists τ_{min} in (5). Therefore, we have

$$\tau_{min} = e^{-\sigma_{clear} R_{max,clear}^{LiDAR}} = e^{-\sigma_{current} R_{max,current}^{LiDAR}} \quad (8)$$

for clear weather and current weather. By comparing the exponent, we have the maximum measurable distance of LiDAR under the current weather as follows:

$$R_{max}^{LiDAR} = \frac{\sigma_{clear}}{\sigma_{current}} R_{max,clear}^{LiDAR}. \quad (9)$$

By definition, $\sigma_{clear} \leq \sigma_{current}$. Therefore, with the above observation, for current weather conditions, we get

$$R_{max}^{LiDAR} \leq R_{max,clear}^{LiDAR}. \quad (10)$$

Given the typical attenuation coefficients 0.1 for clear air, 1 for haze, and 10 for fog [123], the effective range reduces 10 \times in haze and 100 \times in dense fog. Going forward R_{max}^{LiDAR} , shortened to R_{max} , is used to represent the effective range of the LiDAR.

6 | DETECTABILITY MODEL

A *detectability* model describes, given an obstacle's properties, sensor parameters, and environment parameters, whether a given algorithm can detect the obstacle with a given set of algorithm parameters. The *detectability* model allows the conversion of safety standards into an algorithm and sensor parameter requirements as described with the example in Section 2. We now develop the *detectability* model for obstacle detection using a LiDAR. We use Depth Clustering [32, 33] as the example algorithm, which uses depth discontinuity to segment LiDAR points into ground vs obstacles and then determine bounding boxes for obstacles. Table 2 contains a summary of the symbols used in this section with example values from Waymo Open Dataset [34] or chosen defaults.

6.1 | Simple Model

We start with a simple system model, as shown in Figures 3 and 4, with Assumptions A1, A2 and A3.

Ground Removal: The primary part of the Depth Clustering algorithm that we consider in this work is ground removal, *i.e.*, determining which range image points are on obstacles vs ground. FN errors can occur if points on obstacles are mistakenly considered part of the ground. Ground removal is based on vertical depth discontinuity using the inclination angles α . Assuming

the first point ($r = 1$) to always be on the ground, a sharp change in α above a threshold α_{th} indicates that a point is on an obstacle. The first point must lie on the ground for the subsequent comparisons to be meaningful (C3). The inclination angle is calculated from the range image, as designed by Bogoslavskyi and Stachniss [33], shown in Figure 3, as

$$\alpha_{r,c} = \text{atan2}(|BC|, |AC|) = \text{atan2}(\Delta z, \Delta x) = \text{atan2}(|R_{r-1,c} \sin \xi_{r-1} - R_{r,c} \sin \xi_r|, |R_{r-1,c} \cos \xi_{r-1} - R_{r,c} \cos \xi_r|) \quad (11)$$

$$\Delta \alpha_{r,c} = \begin{cases} 0^\circ, & \text{if } r = 1 \\ |\alpha_{r,c} - \alpha_{r-1,c}|, & \text{otherwise} \end{cases} \quad (12)$$

The point r, c is considered to not be on the ground if, $R_{r-1,c}$ is on the ground, $\alpha_{th} < 45^\circ$, and

$$\Delta \alpha_{r,c} > \alpha_{th}, \quad (13)$$

If $R_{r-1,c}$ is not on the ground then $R_{r,c}$ is also considered to not be on the ground. The lowest row of points $R_{0,c}$ are assumed to be on the ground (§5.2 C3). The column index corresponds to the rotational position of the sensor. For brevity, the column index is omitted for points in the same column only.

At horizontal distance D from the sensor, the height of a LiDAR beam r can be calculated as:

$$H_r(D, \xi_r, H_L) = H_L - D * \tan(\xi_r). \quad (14)$$

For brevity, $H_r(D, \xi_r, H_L)$ is referred to as $H_r(D)$ from this point as ξ_r and H_L are constant once a LiDAR sensor is chosen and placed on an AV. Let us define $r = \min\{i | H_i(D) > 0^\circ\}$, i.e., R_r is the first point on the obstacle.

With this background, we can determine the minimum height of an obstacle that can be detected when the obstacle is at a distance of D from the sensor (Figure 3). It is required to check three consecutive LiDAR points to detect an obstacle, according to the ground removal algorithm presented in (11) through (13).

Theorem 1. The obstacle is detected at a distance D , if and only if one of the following conditions is true:

1. $H_r(D) \leq h_o < H_{r+1}(D)$ AND $\alpha_{th} < \text{atan2}(H_r(D), |D - \frac{H_L}{\tan(\xi_{r-1})}|)$;
2. $h_o \geq H_{r+1}(D)$.

Proof. Sufficiency: Consider case 2), i.e., $h_o \geq H_{r+1}(D)$. Then the points R_r and R_{r+1} are on the obstacle, as shown in Figure 3, and thus it holds that $\alpha_{r+1} = 90^\circ$. Since we have

$$\Delta \alpha_r = \alpha_r - 0^\circ, \quad \Delta \alpha_{r+1} = |90^\circ - \alpha_r|,$$

where $\alpha_{r-1} = 0^\circ$ by the definition of index r , it follows that

$$\max\{\Delta \alpha_r, \Delta \alpha_{r+1}\} \geq 45^\circ > \alpha_{th},$$

which guarantees that the obstacle is detected at a distance D .

Now consider case 1), i.e., $H_r(D) \leq h_o < H_{r+1}(D)$ as shown in Figure 4. By the assumption, there exists an index $r - 1$ that touches the ground. Since $h_o \geq H_r(D)$, the angle α_{r-1} can be found by (11) as follows:

$$\begin{aligned} \alpha_r &= \text{atan2}(H_r(D), |R_{r-1} \cos(\xi_{r-1}) - R_r \cos(\xi_r)|) \\ &= \text{atan2}(H_r(D), |D - \frac{H_L}{\tan(\xi_{r-1})}|), \end{aligned} \quad (15)$$

where

$$R_{r-1} \sin(\xi_{r-1}) = H_L, \quad R_r \cos(\xi_r) = D.$$

Since $\alpha_{r-1} = 0^\circ$, the above equation and the angle condition in case 1) imply $\alpha_{th} < \Delta \alpha_r$, i.e., the obstacle is detected at distance D .

Necessity: If the obstacle is detected at a distance D , then at least one point on the obstacle is at distance D , i.e., R_r is on the obstacle. This implies $h_o \geq H_r(D)$. Furthermore, if the obstacle is detected, then

$$\Delta \alpha_r = \alpha_r - \alpha_{r-1} = \alpha_r > \alpha_{th}. \quad (16)$$

There are two cases: $\alpha_{r+1} = 90^\circ$ and $\alpha_{r+1} \neq 90^\circ$. If $\alpha_{r+1} = 90^\circ$, then R_{r+1} is on the obstacle, and thus we have $h_o \geq H_{r+1}(D)$. This implies case 2). On the other hand, if $\alpha_{r+1} \neq 90^\circ$, then the inequality in (16) must hold, which renders the condition $\alpha_{th} \leq \text{atan2}(H_r(D), |D - \frac{H_L}{\tan(\xi_{r-1})}|)$. This implies case 1). This completes the proof. \square

1. $H_{r_g}(D) \leq h_o + h_g < H_{r_g+1}(D)$ AND $\alpha_{th} < \text{atan2}(H_{r_g}(D), |D - \frac{H_L}{\tan(\xi_{r_g-1})}|)$;
2. $h_o + h_g \geq H_{r_g+1}(D, \xi, H_L)$.

The minimum detectable height for this case is found by

$$H_{r_g}(D) = h_o + h_g$$

$$\alpha_{th} \leq \text{atan2}(h_o + h_g, |D - \frac{H_L}{\tan(\xi_{r_g-1})}|).$$

6.4 | Inclined Ground

Assumptions: We maintain all assumptions in Section 6.1 except A2, i.e., Ground inclination α_g could be non-zero (Figure 6).

Ground inclination affects the *detectability* model only when relative inclination changes between the obstacle and the AV. This section assumes that ground inclination starts before the obstacle position, i.e., $d < D$. There are two cases.

Case I: The inclination starts after the beam R_{r-1} , i.e.,

$$\frac{H_L}{\tan(\xi_{r-1})} \leq d \leq \frac{H_L}{\tan(\xi_r)}. \quad (18)$$

We can extend Corollary 1 to find the *detectability* condition. The first beam touching the obstacle must be higher than the ground at a distance of D . Let us define

$$r_f = \min\{i | H_i(D) > (D - d)\tan(\alpha_g)\}.$$

If R_{r_f+1} is at the end of the obstacle, then

$$H_{r_f+1}(D) = h_o \cos(\alpha_g) - h_o \sin(\alpha_g) \tan(\xi_{r_f+1}) + (D - d) \tan(\alpha_g),$$

which is the threshold to determine whether R_{r_f+1} is on the obstacle. Furthermore, α_{r_f} is found by increased height $\frac{H_{r_f}(D) - (D - d)\tan(\alpha_g)}{1 - \tan(\xi_{r_f})\tan(\alpha_g)} + (D - d)\tan(\alpha_g)$ and decreased width $D + \frac{H_{r_f}(D) - (D - d)\tan(\alpha_g)}{\tan(\xi_{r_f}) - \cot(\alpha_g)} - \frac{H_L}{\tan(\xi_{r_f-1})}$. This observation induces the following:

Corollary 3. The obstacle is detected at a distance D , if and only if one of the following conditions is true:

1. $H_{r_f}(D) \leq h_o \cos(\alpha_g) - h_o \sin(\alpha_g) \tan(\xi_{r_f+1}) + (D - d) \tan(\alpha_g) < H_{r_f+1}(D)$ AND $\alpha_{th} < \text{atan2}(\frac{H_{r_f}(D) - (D - d)\tan(\alpha_g)}{1 - \tan(\xi_{r_f})\tan(\alpha_g)} + (D - d)\tan(\alpha_g), |D + \frac{H_{r_f}(D) - (D - d)\tan(\alpha_g)}{\tan(\xi_{r_f}) - \cot(\alpha_g)} - \frac{H_L}{\tan(\xi_{r_f-1})}|)$;
2. $h_o \cos(\alpha_g) - h_o \sin(\alpha_g) \tan(\xi_{r_f+1}) + (D - d) \tan(\alpha_g) \geq H_{r_f+1}(D)$.

Case II: The inclination starts before the beam R_{r-1} , i.e.,

$$d < \frac{H_L}{\tan(\xi_{r-1})}.$$

In this case, the beam point of R_{r-1} land on a different point from that of case II. This decreases relative height $H_L - R_{r-1} \sin(\xi_{r-1})$ and increases relative width $\frac{H_L}{\tan(\xi_{r-1})} - R_{r-1} \cos(\xi_{r-1})$ between R_{r-1} and R_{r_f} . This changes α_{r_f} found in Corollary 3. It is also worth noticing that if

$$d < \frac{H_L}{\tan(\xi_{r-2})},$$

then $\alpha_{r_f-1} = \alpha_g$, and thus we have $\Delta\alpha_r = \alpha_r - \alpha_g$. Using these facts, one can find an extension of Corollary 3.

6.5 | Noise

Assumptions: We maintain all assumptions in Section 6.1 except A3, i.e., there exists a depth detection error. Minor inaccuracies in the depth detection can be a problem when finding $\Delta x, \Delta z$. In particular, the i^{th} beam returns noisy measurement $y_i = R_i \epsilon_i$ [75]

instead of its ground truth distance R_i from the sensor to the obstacle where $\epsilon_i \in [1 - \epsilon, 1 + \epsilon]$ is unknown noise with a known sensor error bound ϵ , e.g., Velodyne HDL-64E S3 LiDAR has a range detection inaccuracy of ± 2 cm.²

Under mild assumptions, Theorems 2 and 3 provide a bound of the first angle α_r on the obstacle, and a bound of the angle α_i between two consecutive points on the obstacle.

Theorem 2. Assume $y_r \cos(\xi_r) \geq y_{r-1} \cos(\xi_{r-1})$ and $y_r \sin(\xi_r) \leq y_{r-1} \sin(\xi_{r-1})$. Then, the angle α_r is lower bounded by

$$\alpha_r \geq \text{atan2}(R_{r-1}(1 - \epsilon) \sin(\xi_{r-1}) - R_r(1 + \epsilon) \sin(\xi_r), R_r(1 + \epsilon) \cos(\xi_r) - R_{r-1}(1 - \epsilon) \cos(\xi_{r-1}))$$

and upper bounded by

$$\alpha_r \leq \text{atan2}(R_{r-1}(1 + \epsilon) \sin(\xi_{r-1}) - R_r(1 - \epsilon) \sin(\xi_r), R_r(1 - \epsilon) \cos(\xi_r) - R_{r-1}(1 + \epsilon) \cos(\xi_{r-1})).$$

Proof. Angle α_r is found by (11). The assumptions $y_r \cos(\xi_r) \geq y_{r-1} \cos(\xi_{r-1})$ and $y_r \sin(\xi_r) \leq y_{r-1} \sin(\xi_{r-1})$ imply that

$$\begin{aligned} \alpha_r &= \text{atan2}(|y_r \sin(\xi_r) - y_{r-1} \sin(\xi_{r-1})|, |y_r \cos(\xi_r) - y_{r-1} \cos(\xi_{r-1})|) \\ &= \text{atan2}(y_{r-1} \sin(\xi_{r-1}) - y_r \sin(\xi_r), y_r \cos(\xi_r) - y_{r-1} \cos(\xi_{r-1})). \end{aligned}$$

Considering the fact that \tan is a strictly increasing function in the domain $[0^\circ, 90^\circ)$, we can find the lower bound and upper bounds presented in the theorem statement. \square

Theorem 3. Assume that two consecutive points R_{i-1} and R_i land on the obstacle, and that $y_i \sin(\xi_i) \leq y_{i-1} \sin(\xi_{i-1})$. Then, the angle α_i is upper bounded by 90° and lower bounded by

$$\alpha_i \geq \text{atan2}(R_i(1 + \epsilon) \sin(\xi_i) - R_{i-1}(1 - \epsilon) \sin(\xi_{i-1}), R_i(1 + \epsilon) \cos(\xi_i) - R_{i-1}(1 - \epsilon) \cos(\xi_{i-1})).$$

Proof. The function atan2 is upper bounded by 90° in all domains. Similarly, the lower bound can be found with the proof of Theorem 2. \square

The assumptions $y_r \cos(\xi_r) \geq y_{r-1} \cos(\xi_{r-1})$ and $y_r \sin(\xi_r) \leq y_{r-1} \sin(\xi_{r-1})$ are mild because $R_r \cos(\xi_r) \geq R_{r-1} \cos(\xi_{r-1})$ and $R_r \sin(\xi_r) \leq R_{r-1} \sin(\xi_{r-1})$ hold for most of the cases, and ϵ_i is around 1. For the same reason, the assumption $y_i \sin(\xi_i) \leq y_{i-1} \sin(\xi_{i-1})$ is mild as well.

6.6 | Summary

Using ξ and H_L from the Waymo Open Dataset [34], (17) yields Figure 7a for various α_{th} and D (X-Axis). Figure 7b shows the same when $\alpha_o = 60^\circ$ as in Section 6.2. Figure 7c is based on Section 6.3, assuming the obstacle is 0.271 m above the ground, based on the ground clearance of a sedan. The varying nature of the colored plot is due to the discrete nature of LiDAR beams. To get a usable bound, like (1), we determine a linear max curve fit for each case, where the dashed line is always greater than or equal to the original plot. The shaded region is where the obstacle is always detected.

In this section, we have developed a *detectability* model for the ground removal of the Depth Clustering algorithm. We find that not only is such analysis possible, but it also yields human perceptible bounds. This work provides bounds on the capabilities of the algorithm. This work stands as the first step in establishing the verifiability of classical obstacle detection algorithms and their practical use in AV for safety-critical obstacle avoidance. A limitation of the presented model is that while it describes the behavior of the core ground vs obstacle differentiation functionality, it does not cover features that reduce false positives or draw boundaries on individual obstacles. We now show present the \mathcal{PS} design that integrates this algorithm into the perception pipeline to provide guarantees for collision avoidance.

²<https://velodynelidar.com/products/hdl-64e/>

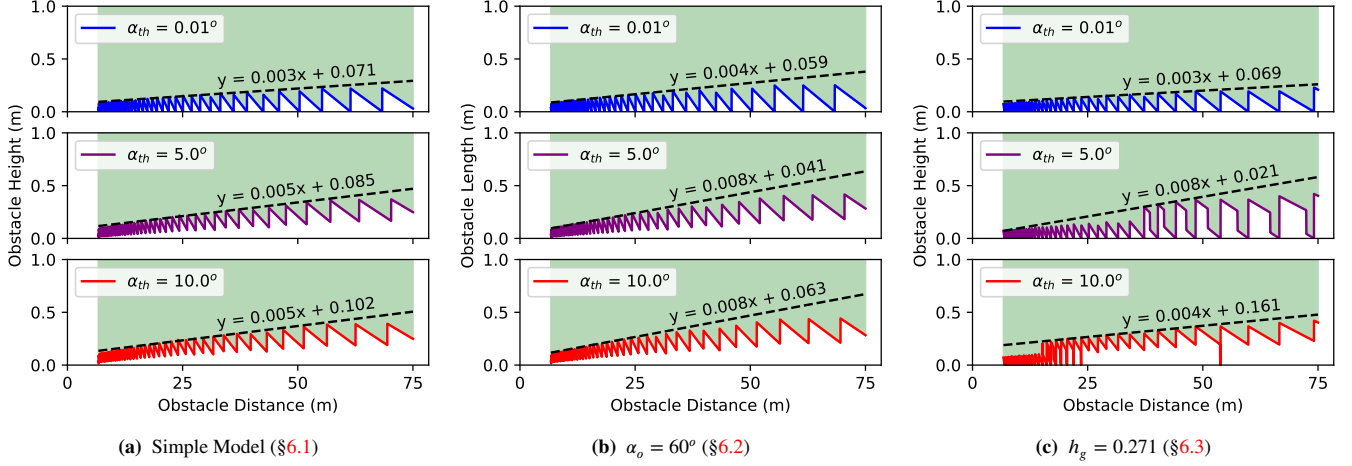


Figure 7 Minimum detectable obstacle height (Y-Axis) at varying distances from the vehicle (X-Axis). Black dashed lines are a linear max fit for the plots, with corresponding equations. The shaded region signifies where the obstacle is always detected.

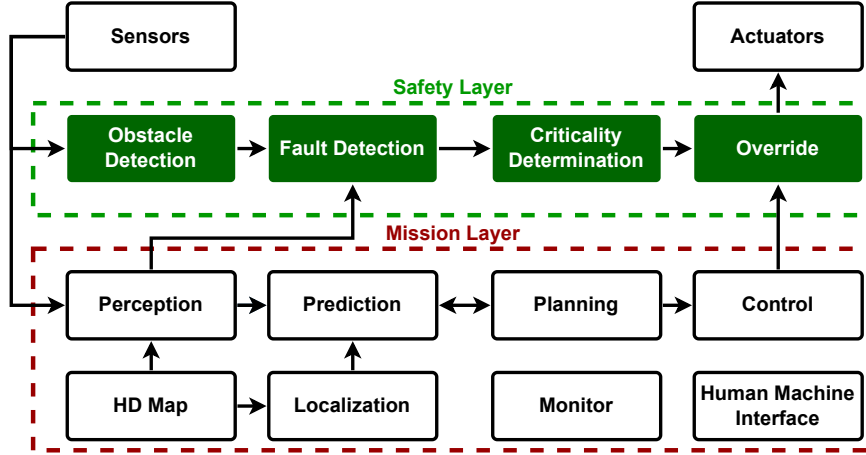


Figure 8 Control and Data Pipeline of an Autonomous Driving system within Perception Simplex (\mathcal{PS}) architecture. The mission layer uses Baidu’s Apollo 5.0 [30]. The safety layer uses Depth Clustering [33] for obstacle detection.

7 | Perception Simplex

Perception Simplex (\mathcal{PS}) presents a system architecture that can utilize the verifiable obstacle detection capabilities described in Section 6. Acknowledging the limitations on capabilities of verifiable algorithms and unverifiability of DNN, the goal of \mathcal{PS} is to provide specific safety guarantees amidst perception faults, rather than complete replacement as in traditional Simplex architecture [28, 29]. As noted before in Section 2, this work focuses on obstacle existence detection faults. Therefore \mathcal{PS} preserves the full functionality of the high-capability mission-critical software when no faults are detected. When a fault is detected, and such a fault poses a risk of collision, the safety layer intervenes, abandoning the mission of the AV, to ensure safety. Figure 8 shows the AV pipeline under \mathcal{PS} .

7.1 | Mission Layer

The mission layer is an autonomous driving agent capable of fulfilling the mission of the AV, *i.e.*, driving from one point to another while trying to maintain safety. However, if the safety layer detects a reason to override the mission layer, the mission

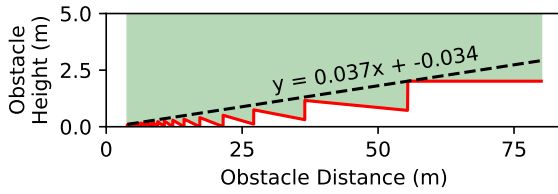


Figure 9 *Detectability* model for the software-in-the-loop simulation setup (§9). The red line plot uses Corollary 2 and parameters from the simulation setup. The black line is the linear max fit, same as Figure 7.

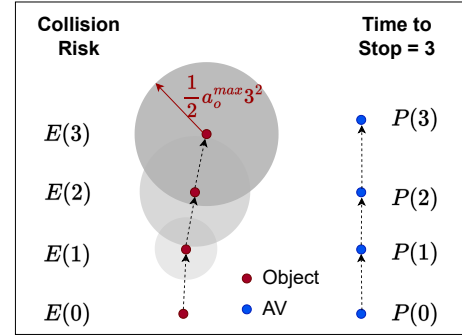


Figure 10 Collision risk within Time to Stop of the AV, adapted from prior work [104]. $E(t)$ denotes the existence region [105] of the obstacle at time t , *i.e.*, the region it can exist based on current position and velocity measurement, and maximum acceleration capability of a_o^{max} . $P(t)$ denotes the position of the AV on the local planned path at time t .

of the AV is abandoned till the mission layer faults are mitigated. The clear separation of the mission layer from the safety layer makes the \mathcal{PS} architecture forward compatible with all improvements in the autonomous driving agents. The only explicit design requirement for the mission layer in \mathcal{PS} is to explore its intermediate results and control commands to the safety layer. This is easily achieved in the publisher-subscriber model, common to the middleware of open-source AV implementations, *i.e.*, CyberRT [125] for Baidu’s Apollo 5.0 [30] or ROS [126] for Autoware [31]. In this work, Baidu’s Apollo 5.0 [30] is used as the mission layer.

7.2 | Safety Layer - Obstacle Detection

Obstacle detection is performed using the Depth Clustering [32,33] algorithm. Figure 9 shows the *detectability* model based on Corollary 2 and the parameters from evaluation setup (§8). It should be noted that while the *detectability* model covers the core ground removal, as in (11), (12), and (13), Depth Clustering algorithm has broader features, *e.g.*, points clustering to determine obstacle boundaries or noise filtering, that are used for the safety layer obstacle detection.

7.3 | Safety Layer - Fault Detection

Faults in mission layer object detection output are determined by comparing it against the safety layer’s obstacle detection output. For the comparison, we use the minimized sufficient obstacle detection requirements (§4). Intuitively, the comparison checks whether mission layer detections cover the hazards detected by the safety layer.

7.4 | Safety Layer - Criticality Determination

To determine the appropriate reaction to a fault, first, its severity must be established. For any obstacle existence detection fault, it must be ascertained if this obstacle poses a risk of collision with the AV. We use a physical model for collision risk [104], wherein an obstacle is determined to pose a risk of collision if their existence regions [105] overlap within the AV’s time to stop. Figure 10 describes the collision risk determination.

7.5 | Safety Layer - Override

Two overrides are imposed on the mission layer control output. Velocity is limited to maintain the safety guarantee, as derived later in Theorem 4 and (19). An emergency brake is applied to avoid the potential collision if a safety-critical obstacle existence detection fault is determined and evaluated using Algorithm 1.

Algorithm 1: Decision Logic.**Result:** No Override or Brake**Input:** O_m : Mission layer's current set of objects, O_s : Safety layer's current set of obstacles

```

1 if  $O_m.update() \parallel O_s.update()$  then           // Run when either detection set changes
2   foreach  $o_s$  in  $O_s$  do
3     if  $isDetected(o_s, O_m)$  then
4       continue
5     if  $isCollisionRisk(o_s)$  then
6       return (Brake)
7 return (No Override)

8 Function  $isDetected(o_s, O_m)$ :
9   return Safety-critical minimized obstacle detection requirements (§4, §7.3).

10 Function  $isCollisionRisk(o_s)$ :
11   return Potential for Collision [104] using Existence Regions [105] (§7.4, Figure 10).
```

7.6 | Safety Guarantee

Algorithm 1 describes the safety layer decision logic. We can now derive \mathcal{PS} 's safety guarantee.

Theorem 4. Let's assume an AV must detect all obstacles with height $\geq H^O$, defined by safety regulations. When *detectability* model constraints are met, collisions with stationary obstacles of height $\geq H^O$, and that the mission layer fails to detect, are always avoided, when the AV's operating velocity is always below a predetermined limit:

$$v_{max}^{safe} = \sqrt{(a_{max}^{av} L_{max})^2 + 2a_{max}^{av} D_{max}^{stop}} - a_{max}^{av} L_{max} \quad (19)$$

where a_{max}^{av} , L_{max} are the maximum safe deceleration of the AV, and worst-case computational latency, respectively. Furthermore, maximum stopping distance D_{max}^{stop} depends on the *detectability* model and the current maximum detection range of LiDAR.

Proof. We will first show that v_{max}^{safe} in (19) is the maximum safe velocity for collision avoidance by the safety layer. Let's denote the maximum range of the LiDAR sensor in clear weather, a sensor specification, by $R_{max,clear}^{LiDAR}$. This is a sensor specification, e.g., 75 m for top LiDAR in Waymo Open Dataset [34] and 100 m for the LiDAR used in simulation [127]. As described in Section 5.2.1, this range reduces in low visibility conditions, e.g., due to fog or haze. The maximum detection range of LiDAR R_{max}^{LiDAR} is as in (9). Using the *detectability* model (§6) the maximum range at which an object of height H^O can be detected can be determined. Let's denote this range as R_{max}^O . The maximum range at which obstacles can be reliably detected (R_{max}) is then

$$R_{max} = \min(R_{max}^O, R_{max}^{LiDAR}). \quad (20)$$

(20) establishes the maximum detection range. However, as a safety policy and to accommodate for any error allowances within the system, a safety margin can be added to this range, denoted as D^{Safety} . Combining this with (20) we get the maximum stopping distance for the AV and can now derive v_{max}^{safe} by solving:

$$D_{max}^{stop} = R_{max} - D^{Safety} = v_{max}^{safe}(t + L_{max}) - \frac{1}{2}a_{max}^{av}t^2 \quad (21)$$

where t is deceleration time. Substituting t in (21) with $t = \frac{v_{max}^{safe}}{a_{max}^{av}}$, we get (19).

Now we prove the rest of the statement. Since the obstacle height is $\geq H^O$, the obstacle's existence is detected by the safety layer at all distances $\leq R_{max}$ in (20) i.e., it is part of the set of obstacles detected by safety layer (O_s). When obstacle o_s is not detected by the mission layer, i.e., it is not part of the set of mission layer detections (O_m) i.e., $o_s \notin O_m$, but $o_s \in O_s$, Algorithm 1 returns "Brake" if this obstacle poses collision risk. If the current velocity is less than v_{max}^{safe} in (19), we can guarantee collision avoidance for the obstacle o_s . \square

Table 3 Results for False Negative (FN) evaluation using real-world dataset.

Count	Category
98224	Total obstacles (Vehicles, Pedestrian, Cyclist, Unknown)
93030	Total without obstacles closer than D_{min} (C3 §5.2)
7565	Obstacles that pose collision risk (§4.2) [104]
7402	Detected meeting minimally sufficient requirements (§4)
153	Ground Truth larger than obstacle
10	Oversegmentation
0	Remaining FN count at 75% coverage

It should be noted that while Theorem 4 considers stationary obstacles, trivially, any obstacle moving away from the AV’s position is also avoided. Even when AV velocity is higher than v_{max}^{safe} , braking will reduce collision velocity resulting in a reduced impact of any collisions. Further, since all parameters here can be determined statically or based on AV’s operating conditions (e.g., a_{max}^{av}), v_{max}^{safe} can be calculated to provide a guarantee against collisions with static obstacles. Only H^O is based on obstacle properties, which we assume are lower bound by regulations and policies (§2). There needs to be a policy or safety standard defining the minimum height of obstacles that must be detected and avoided by AV. Given such a policy the operational design domain of the presented \mathcal{PS} design is determined by (a) adherence to the constraints of the *detectability* model (§5); (b) v_{max}^{safe} , adjustable by selecting sensor, AV and algorithm parameters.

8 | DATASET EVALUATION

In this section, we evaluate the obstacle detection algorithm using real-world sensor data from Waymo Open Dataset [34]. We first determine if the algorithm meets the safety-critical requirements proposed in this work (§4). For obstacle existence fault or FN evaluation, we randomly select 16 clear weather scenes from the 202 scenes in the validation dataset. The random down-sampling was necessitated by the manual effort involved in analyzing FN candidates. The scenes included various scenarios, including heavy to low traffic, residential with pedestrians, city, and highway driving.

Table 2 summarizes the parameters used. α_{th} was set at 10° . Distance overestimation error was bounded to $10\text{ cm} + 5\%$ of actual distance. The constant was chosen as a small value to keep the error bound low at close distances. However, since the depth perception accuracy reduces with distance an additional percentage-based threshold allows accommodation of sensor limits while having a low impact. A limited coverage threshold of 75% was used as a threshold for True Positive (TP) detection, chosen to parallel the strict metric for IOU from COCO Dataset [128]. All FN candidates were analyzed manually using a visualizer provided by Bogoslavskyi and Stachniss [33]. We enhanced the visualizer with selective Ground Truth (GT) annotations and point cloud coloring, to aid the manual inspection. Table 3 summarizes the results.

8.0.1 | GT Counts

We first determine the total number of GT in the included scenes. All classes, excluding road signs, were counted to yield a total of **98224**. Obstacles closer than the first beam ($R_{1,c}$) on the ground, i.e., closer than D_{min} are removed as per C3 in Section 5.2, reducing the count of GT to **93030**. Obstacles that do not pose a risk of collision are also removed (§4.2), leaving **7565** obstacles.

8.0.2 | Automated Evaluation

We run the detection algorithm and evaluate results based on requirements described in Section 4. **7402** True Positives and **163** FN candidates were found. The FN candidates were then manually inspected.

8.0.3 | GT Dimension Error

The most common reason for erroneous FN indication was inaccurate GT labels. This inaccuracy of GT was determined based on a visual inspection of point clouds. The GT was visibly larger or offset from the actual obstacle. Figures 11a and 11b show

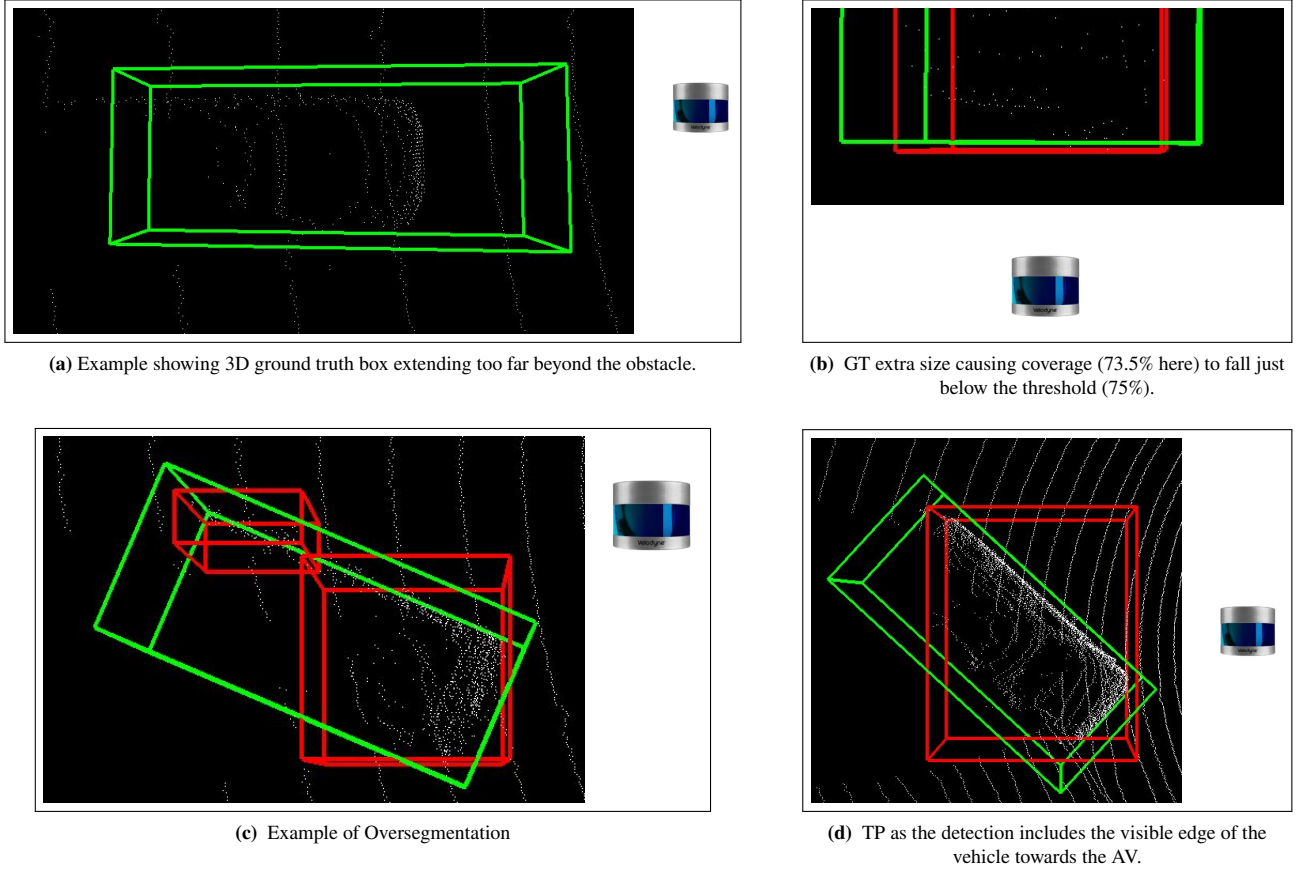


Figure 11 Examples from manual inspections of FN candidates. The white points are in 3D space as returned by LiDAR *i.e.*, the point cloud. Ground truth 3D labels are drawn as green bounding boxes, while detection bounding boxes are red. The point cloud is shown from a top view, and the LiDAR direction is indicated with the sensor image [129].

such examples. Drawn as per the provided GT, the green box is either larger or offset from the contained obstacle. Similarly, as shown in Figure 11b, the GT inaccuracy is enough to bring the coverage below the 75% threshold. In some cases, the error is small, *e.g.*, Figure 11d. However, when close to the AV, the small GT error can still be larger than the distance overestimation error allowed. We consider these detections as TP after ascertaining that the detection bounding box meets the requirements. A total of **153** FN candidates were found to fall in this category.

8.0.4 | Oversegmentation

Figure 11c shows a case where the obstacle was adequately detected but segmented into more than one bounding box. Since the second bounding box did not meet the distance threshold, the automated analysis ignored it. However, given the presence of both bounding boxes, we argue that this detection should be considered True Positive. The points on the obstacle were not erroneously considered to be drivable ground. **10** instances of this scenario on the same vehicle were found in consecutive frames.

8.0.5 | Obstacle Existence Fault

No FN, *i.e.*, obstacle existence faults were found. This is not surprising given the low minimum height bounds determined in Section 6 and Figure 7. The max curve fits in Figures 7a, 7b and 7c are conservative linear approximations. Ignoring the max curve fits, the actual bounds in Figures 7a, 7b and 7c were less than 0.5 m in all cases within the sensor range of 75 m.

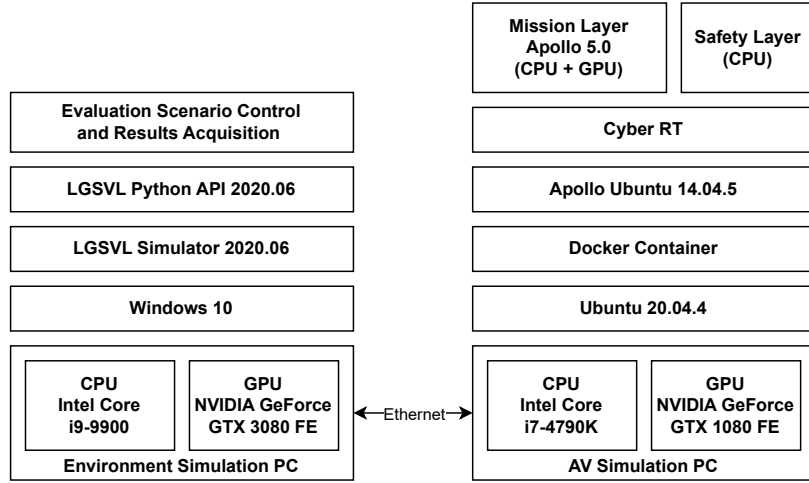


Figure 12 Software-in-the-loop simulation setup for the evaluation. LGSVL [35] simulates the driving environment, controlled via its Python API interface. The mission and safety layers components are shown in Figure 8.

9 | SOFTWARE-IN-THE-LOOP SIMULATION

We evaluate the efficacy of \mathcal{PS} using software-in-the-loop simulation. Figure 12 describes the setup, based on Apollo 5.0 [30] and LGSVL Simulator [35].

9.1 | Setup

To evaluate the efficacy of \mathcal{PS} against obstacle existence faults, the scenario as shown in Figure 13 is used. The scenario starts with the AV moving with initial velocity $v(0) \in \{5, 10, \dots, 40\} \text{ m/s}$ with a vehicle in its path at distance, $d(0) \in \{10, 20, \dots, 100\} \text{ m}$. The AV's mission is to pass this obstacle. The safety requirement during this mission is to avoid any collisions. Each simulation run results in one of the following outcomes, summarized in Table 4.

- ✗ *Collision*: AV collides with the obstacle, therefore the safety requirement is violated and the mission cannot be fulfilled.
- *Safe Stop*: AV safely stops without colliding with the obstacle, however, it is unable to continue and pass the obstacle.
- + *Safe Pass*: AV passes the obstacle without colliding with it.

After the simulation start ($t > 0$), \mathcal{PS} controls the AV in one of the following three configurations of the \mathcal{PS} pipeline (Figure 8):

- *Mission Only* (\mathcal{MO}): Mission Layer directly controls the AV. The safety layer is disabled. While this scenario provides the baseline for the mission capabilities of the system, however, the mission layer has considerations beyond collision avoidance, e.g., mission completion and rider comfort, which makes a direct comparison less useful.
- *Mission Crash* (\mathcal{MC}): Mission layer becomes unavailable at $t = 0$, therefore brakes are applied $\forall t > 0$ and hence is the synthetic ideal case for safety.
- *Fault Injected* (\mathcal{FI}): All components of \mathcal{PS} are functional and available. However, obstacle existence detection faults are synthetically injected into the object detection output communicated from the mission layer to the safety layer. Safe handling of obstacle existence detection faults is the primary focus of this work (§2), therefore, this configuration represents the evaluation target.

\mathcal{MO} setup provides a baseline for the mission behavior, while \mathcal{MC} shows the best-case braking response, given the dynamics of the scenario. \mathcal{FI} is the evaluation target. Using simulation parameters, $R_{max,clear}^{LiDAR} = 100 \text{ m}$, $D^{Safety} = 0.1 \text{ m}$, $a_{max}^{av} = 7.5 \text{ m/s}^2$, $L_{max} = 0.01 \text{ s}$, and $H_O = 0.75 \text{ m}$ for sedan's back with (19) we get $v_{max}^{safe} = 17.71 \text{ m/s}$. For the evaluation, the velocity limit is not enforced by the safety layer, rather it is used here to compare the \mathcal{PS} behavior when AV is below this limit vs not.

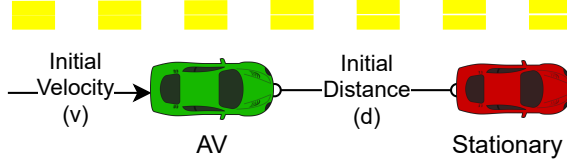


Figure 13 Evaluation scenario.

Table 4 Outcomes

	Safety	Mission
Collision	Violated ✗	Violated ✗
Safe Stop	Met ✓	Violated ✗
Safe Pass	Met ✓	Met ✓

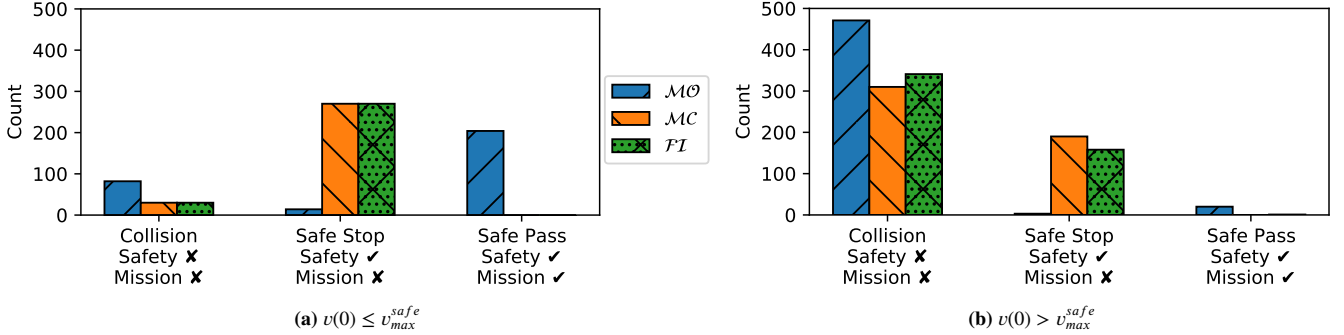


Figure 14 Comparison of scenario outcomes (X-Axis). Collision, Safe Stop, and Safe Pass are the possible outcomes (Table 4). The Y-Axis counts how often an outcome is observed among all the simulation instances for each configuration (§9.1).

9.2 | Safety - Overview

For each $v(0)$, $d(0)$ and PS configuration combination, 10 simulation repetitions are run *i.e.*, a total of 2400 instances.³ We now compare the prevalence of different outcomes. Figure 14a shows the count of outcomes when initial velocity is below the safe limit, *i.e.*, $v(0) \leq v_{max}^{safe}$ and Figure 14b shows the same when initial velocity is above the safe limit, *i.e.*, $v(0) > v_{max}^{safe}$.

Discussion. When $v(0) \leq v_{max}^{safe}$ the target FI results are identical to the best case configuration of MC . This supports the primary claim of this work, *i.e.*, deterministic collision avoidance in the presence of obstacle existence detection faults when operating within predictable safety limits.

When $v(0) > v_{max}^{safe}$ the max range at which the obstacle is detected by the safety layer obstacle detection algorithm (R_{max}^O) is not guaranteed to be a sufficient distance for the AV to brake to stop. This results in more collision outcomes for the target result in FI as compared to the ideal case of MC in Figure 14b.

Despite the absence of a deterministic guarantee, Safe stops still happen when $v(0) > v_{max}^{safe}$, because (i) although not guaranteed, as shown in Figure 9, the discrete nature of LiDAR beams leads to regions below the linear bound where obstacles may be detected when beyond R_{max}^O ; (ii) false positives cause unnecessary early braking which in this case creates additional slack time for collision avoidance; (iii) when $d(0)$ is large, the mission layer's actuation commands reduce AV speed to match its planning and control goals, helping avoid collisions.

It should be noted that the safety layer in PS does not share the mission layer's goal of completing the AV's mission. Hence, when the mission layer has safety-critical faults, MC and FI , the mission of the AV, *i.e.*, *Safe Pass*, is abandoned.

v_{max}^{safe} is derived from vehicle, algorithm, and sensor parameters, its value can be increased by changing these parameters, *e.g.*, larger LiDAR sensor array.

9.3 | Safety - Worst Case

The simulation results are now presented in higher detail, emphasizing the worst case for each $d(0)$, $v(0)$ and PS configuration combination, as shown in Figure 15. Among the 10 repetitions for each combination, if any instance results in a collision, the

³Out of any 10 repetitions, at minimum 8 instances yielded valid outcomes. Rarely (0.3%), when initial velocity is high (35 m/s or 40 m/s), AV becomes unstable and goes out of the road boundaries. Such instances are not included in the results.

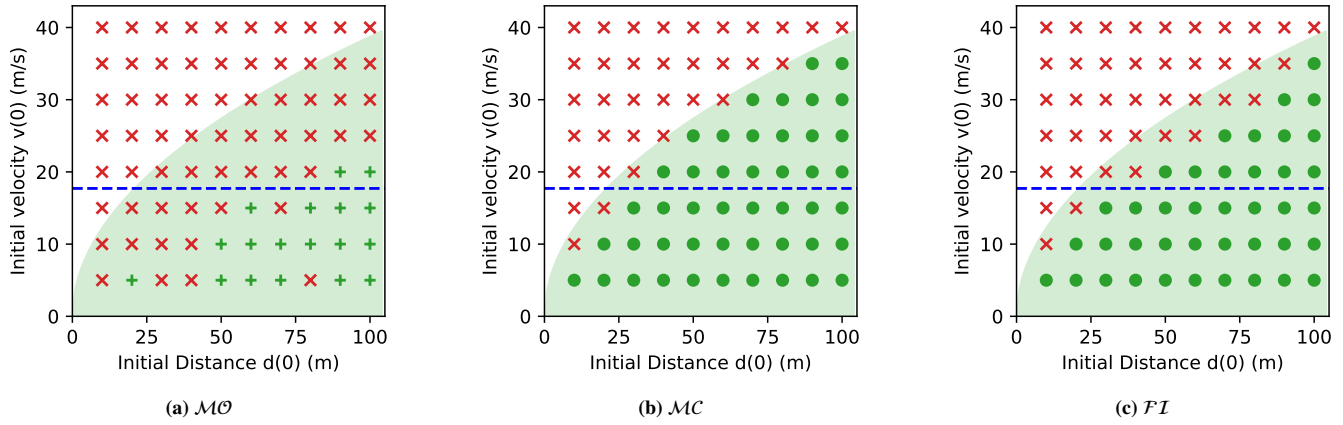


Figure 15 Each data point represents the worst-case outcome among the 10 repetitions conducted for each combination of $d(0)$ (X-Axis), $v(0)$ (Y-Axis), and PS configuration. The shaded green area, defined by $v(0) \leq \sqrt{2 * 7.5 * d(0)}$, represents the distance and velocity combinations at which it is physically possible for the AV to avoid a collision, assuming constant deceleration of 7.5 m/s^2 and zero response latency. \times represents collision, \bullet represents safe stop and $+$ represents safe pass (§9.1). The blue dashed line notes $v_{max}^{safe} = 17.71 \text{ m/s}$.

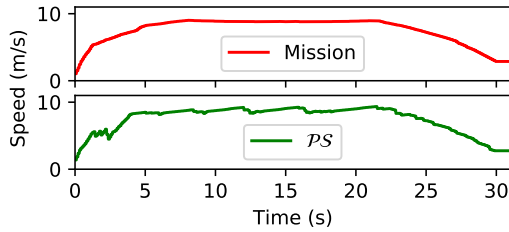


Figure 16 Performance comparison. $d(0) = 50\text{m}$, $v(0) = 1\text{m/s}$. At $t = 2$ safety layer first detects the obstacle.

Table 5 Computational Latency

Layer	Module	Average	Worst Case
Mission	Segmentation	69362 μs	128634 μs
Mission	Fusion	922 μs	3665 μs
Safety	Obstacle Detection	7273 μs	38391 μs
Safety	Fault Handlers	387 μs	9801 μs

worst-case result notes it as *Collision* (\times). Only when all repetitions safely pass the obstacle the noted result is a *Safe Pass* ($+$). When no collisions occur and at least one instance does not lead to a safe pass, the result is noted to be *Safe Stop* (\bullet).

$v_{max}^{safe} = 17.71 \text{ m/s}$ is indicated by blue horizontal lines in each plot. The shaded area ($v(0) \leq \sqrt{2 * 7.5 * d(0)}$) notes the best possible braking response, assuming $a_{max}^{av} = 7.5 \text{ m/s}^2$. This ignores any computation latency or dynamics of the simulation. Figure 15b is the synthetic best case for safety where the AV applies maximum brake at the start of the simulation run.

Discussion. Figure 15c captures the PS behavior under obstacle existence detection faults, the focus of this work (§2). This result substantiates the $v_{max}^{safe} = 17.71 \text{ m/s}$ safe operating limit established earlier, as below this limit, the worst case result of the ideal scenario (Figure 15b) is same as Figure 15c, the PS behavior under obstacle existence detection faults.

9.4 | Performance

Figure 16 shows the comparison of AV speed over time when controlled by the Mission layer only ($M\Theta$) vs. PS without any instrumented faults. The evaluation scenario is similar to Figure 13, except that the obstacle is not in the same lane as the AV, rather it is in an adjacent lane. Initial AV state is fixed at $d(0) = 50\text{m}$, $v(0) = 1\text{m/s}$. MC and FI are not evaluated here as they would lead to a safe stop but the mission of the AV will be abandoned due to the instrumented faults. Compared to the baseline of Mission, PS shows some additional braking, however erroneous overrides due to transient false positive detections from the safety layer algorithm are mitigated with minimal impact on AV speed, as enforced by Algorithm 1. The only significant difference near the start ($t \approx 2$) is when the safety layer detects the obstacle whereas the mission layer does not yet.

9.5 | Computational Latency

The final result is a comparison of the execution latency of safety and mission layer components. The results are presented in Table 5. Safety layer Fault Handler time includes fault detection, criticality determination, and override. The values only include the computational latency *i.e.*, any time spent waiting for inputs is excluded. Safety layer components have significantly lower latency than those in the mission layer, a direct result of the simplified safety requirements and minimalist design of the safety layer components. The low latency supports a faster response to external stimuli.

10 | DISCUSSION

Incremental Design. The PS architecture presented in this work is designed to focus specifically on obstacle existence detection faults due to their prevalence in real-world fatal collisions (§2). PS can be expanded to address other perception faults if verifiable algorithms exist that can fulfill the corresponding perception task. In the absence of such capabilities, safety overrides must be applied precisely to only handled fault types.

Consider the scenario of a deer suddenly jumping in front of an AV using the presented PS . If the mission layer is unable to detect the existence of the deer, the safety layer overrides the control and applies brakes. This has multiple beneficial effects, such as completely avoiding the collision or reducing the impact velocity if a collision is unavoidable. Additionally, the time before the collision is increased, giving the mission layer more time to react. When the mission layer does detect the existence of the deer, the safety layer removes its override, even if doing so leads to a collision with the deer. Within the presented design, the safety layer does not have the capability to determine the best possible path for the AV. A collision with the deer can be the best response, avoiding more serious collisions with other vehicles. The safety layer, as presented in this work, does not have the capability to make this distinction and therefore does not intervene. Therefore, PS provides guarantees against specific faults only, which can be expanded with the inclusion of other verifiable capabilities. Rather than replacing other prior solutions, the modular design of PS solves the most critical issue first while remaining open to future expansion and integration.

Trade-offs. The improved safe operating capabilities of the AV come at the cost of degradation of rider comfort and performance, *e.g.*, due to reduced speed and increased braking (Figure 16). PS minimizes the performance loss by responding to specific faults and only when faults pose a risk of collision. An advantage of using verifiable and logically analyzable software is that for each hardware and software parameter, the impact of different values can be determined analytically.

Limitations and Future Works. There are two main limitations of this work which are also the topic for future works. First, the presented verification model addresses the main ground vs obstacle differentiation step, however, the complete depth clustering algorithm goes beyond this. These additional steps reduce the occurrence of false positive detections. Additionally, as detailed in Section 5.2 constraints on obstacle width (C2) and first point of ground (C3) are minimally restrictive and can be easily mitigated. However, the inherent limitation of LiDAR sensor (C1) cannot be mitigated directly. Therefore one direction for future work is to create an expanded *detectability* model unifying different sensor types, all using verifiable algorithms, *i.e.*, sensor fusion, and including the additional algorithmic steps.

Second, while PS presents safety guarantees in the presence of faults without compromising the higher mission layer capabilities when faults are not detected, the response to faults is limited to braking and does not leverage the mission layer capabilities. Therefore the main focus of our future work is to utilize mission layer capabilities in fault tolerance and override actions, while maintaining the verifiable safety guarantees presented in this work.

11 | CONCLUSION

This work presents Perception Simplex (PS), a system architecture for fault tolerance against perception faults. In this work, PS is applied to the critical issue of obstacle existence detection faults. The modular design of PS makes it extensible to cover other types of perception faults, while the verifiable nature of its safety layer ensures that critical safety requirements are met without relying on uncertain machine learning models. PS provides a promising direction for the design of fault-tolerant AV.

References

- [1] Ayoosh Bansal, Hunmin Kim, Simon Yu, Bo Li, Naira Hovakimyan, Marco Caccamo, and Lui Sha. Verifiable Obstacle Detection. In *2022 IEEE 33rd International Symposium on Software Reliability Engineering (ISSRE)*, pages 61–72. IEEE, 2022. doi: [10.1109/ISSRE55969.2022.00017](https://doi.org/10.1109/ISSRE55969.2022.00017).
- [2] Philip Koopman. Autonomous vehicle myths: The dirty dozen. <https://www.eetimes.com/autonomous-vehicle-myths-the-dirty-dozen/>, 2021. Accessed: Jun 6, 2023.
- [3] Dong Ngo. Dashcam shows fatal Tesla Model S crash in China. <https://www.cnet.com/news/dash-cam-showed-fatal-tesla-crash-in-china>, Sep 2016. Accessed: Jun 6, 2023.
- [4] National Transportation Safety Board. Collision Between a Car Operating With Automated Vehicle Control Systems and a Tractor-Semitrailer Truck Near Williston, Florida, May 7, 2016. Accident Report NTSB/HAR-17/02 PB2017-102600, September 2017. URL: <https://www.nts.gov/investigations/accidentreports/reports/har1702.pdf>.
- [5] National Transportation Safety Board. Collision Between Vehicle Controlled by Developmental Automated Driving System and Pedestrian Tempe, Arizona, March 18, 2018. Accident Report NTSB/HAR-19/03 PB2019-101402, 2019. URL: <https://www.nts.gov/investigations/AccidentReports/Reports/HAR1903.pdf>.
- [6] National Transportation Safety Board. Collision Between a Sport Utility Vehicle Operating With Partial Driving Automation and a Crash Attenuator Mountain View, California, March 23, 2018. Accident Report NTSB/HAR-20/01 PB2020-100112, February 2020. URL: <https://www.nts.gov/investigations/AccidentReports/Reports/HAR2001.pdf>.
- [7] Edward C. Chen and Joel Greer. Umeda v. Tesla, Inc., United States District Court for the Northern District of California. Case No. 5:20-cv-2926: Docket No. 1, Complaint for damages, April 2020. Accessed: Jun 6, 2023. URL: https://regmedia.co.uk/2020/04/30/tesla_complaint.pdf.
- [8] National Transportation Safety Board. Highway accident brief HWY19FH008. Highway Accident Brief, Accident Number: HWY19FH008, March 2019. URL: <https://www.nts.gov/investigations/AccidentReports/Reports/HAB2001.pdf>.
- [9] Neil Boudette. It happened so fast: Inside a fatal tesla autopilot accident. <https://www.nytimes.com/2021/08/17/business/tesla-autopilot-accident.html>, 2021. Accessed: Jun 6, 2023.
- [10] Dan Kerman. Lawsuit: Family blames tesla’s autopilot for deadly crash. <https://www.kron4.com/news/bay-area/lawsuit-family-blames-teslas-autopilot-for-deadly-crash/>, 2021. Accessed: Jun 6, 2023.
- [11] Jared Jernagan. Fatal tesla crash into fire truck was one year ago today. <https://www.bannergraphic.com/story/2856765.html>, 2020. Accessed: Jun 6, 2023.
- [12] AP News. Feds will investigate deadly tesla crash in california. <https://apnews.com/article/technology-business-los-angeles-us-news-california-6eae3986e7d9c1d00db7d52146cddf23>, 2020. Accessed: Jun 6, 2023.
- [13] GEIR RØED. Tesla on auto-steering when man was cut down [translated]. <https://www.motor.no/autopilot-nyheter-tesla/tesla-pa-auto-styring-da-mann-ble-meid-ned/188623>, 2020. Accessed: Jun 6, 2023.
- [14] John Antczak and Tom Krisher. Crash, arrest draw more scrutiny of tesla autopilot system. <https://apnews.com/article/ca-state-wire-california-arrests-business-technology-388ae97b2c5c8a7ca7ae816323e7b829>, 2021. Accessed: Jun 6, 2023.
- [15] NBC New York. 52-year-old fixing flat on long island expressway hit, killed by tesla driver: Cops. <https://www.nbcnewyork.com/news/local/52-year-old-fixing-flat-on-long-island-expressway-hit-killed-by-tesla-driver-cops/3175414/>, 2021. Accessed: Jun 6, 2023.
- [16] T Kebang. Believing too much in the words “automatic driving”, the ceo of a chinese startup company crashed into a construction vehicle in self-driving mode and died, 2021. Accessed: Jun 6, 2023. URL: <https://www.newsdirectory3.com/believing-too-much-in-the-words-automatic-driving-the-ceo-of-a-chinese-startup-company-crashed-into-a-construction-vehicle-in-self-driving-mode-and-died-t%E5%AE%A2%E9%82%A6/>.

- [17] Tony Kovalski. Fatal tesla crash in evergreen now part of federal investigation. <https://www.denver7.com/news/investigations/fatal-tesla-crash-in-evergreen-now-part-of-federal-investigation>, 2022. Accessed: Jun 6, 2023.
- [18] Miranda Ceja. Pedestrian killed in fatal dui crash tuesday in mission viejo. <https://patch.com/california/missionviejo/pedestrian-killed-fatal-dui-crash-tuesday-mission-viejo>, 2022. Accessed: Jun 6, 2023.
- [19] Tom Krisher. Nhtsa to investigate florida tesla crash that killed 2. <https://apnews.com/article/california-florida-gainesville-lompoc-1ec22080a57d4077f517a31ae021965a>, 2022. Accessed: Jun 6, 2023.
- [20] Andrew J. Hawkins. Tesla driver using autopilot kills motorcyclist, prompting another nhtsa investigation. <https://www.theverge.com/2022/7/27/23280461/tesla-autopilot-crash-motorcyclist-fatal-utah-nhtsa>, 2022. Accessed: Jun 6, 2023.
- [21] Olivia Lloyd. Motorcyclist dies after being rear-ended by high-speed tesla in boca raton. <https://www.sun-sentinel.com/local/palm-beach/boca-raton/fl-ne-motorcyclist-dies-in-rear-end-crash-with-tesla-20220826-buh3wz4m2vcvdhenhwd3npkszu-story.html>, 2022. Accessed: Jun 6, 2023.
- [22] ABC7 News. Tesla that hit fire truck in deadly i-680 crash in walnut creek was on autopilot, company says, 2023. Accessed: Jun 6, 2023. URL: <https://abc7news.com/tesla-autopilot-crash-driver-assist-crashes-into-fire-truck-walnut-creek-fatal/13144903/>.
- [23] Douglas Heaven et al. Why deep-learning ais are so easy to fool. *Nature*, 574(7777):163–166, 2019. doi:10.1038/d41586-019-03013-5.
- [24] Xiaowei Huang, Daniel Kroening, Wenjie Ruan, James Sharp, Youcheng Sun, Emese Thamo, Min Wu, and Xinpeng Yi. A survey of safety and trustworthiness of deep neural networks: Verification, testing, adversarial attack and defence, and interpretability. *Computer Science Review*, 37:100270, 2020. doi:10.1016/j.cosrev.2020.100270.
- [25] Ana Pereira and Carsten Thomas. Challenges of machine learning applied to safety-critical cyber-physical systems. *Machine Learning and Knowledge Extraction*, 2(4):579–602, 2020. doi:10.3390/make2040031.
- [26] Oliver Willers, Sebastian Sudholt, Shervin Raafatnia, and Stephanie Abrecht. Safety concerns and mitigation approaches regarding the use of deep learning in safety-critical perception tasks. In *International Conference on Computer Safety, Reliability, and Security*, pages 336–350. Springer, 2020. doi:10.1007/978-3-030-55583-2_25.
- [27] Florian Tambon, Gabriel Laberge, Le An, Amin Nikanjam, Paulina Stevia Nouwou Mindom, Yann Pequignot, Foutse Khomh, Giulio Antoniol, Ettore Merlo, and François Laviolette. How to certify machine learning based safety-critical systems? a systematic literature review. *Automated Software Engineering*, 29(2):38, 2022. doi:10.1007/s10515-022-00337-x.
- [28] Lui Sha et al. Using simplicity to control complexity. *IEEE Software*, 18(4), 2001. doi:10.1109/MS.2001.936213.
- [29] Stanley Bak, Deepti K Chivukula, Olugbemiga Adekunle, Mu Sun, Marco Caccamo, and Lui Sha. The system-level simplex architecture for improved real-time embedded system safety. In *2009 15th IEEE Real-Time and Embedded Technology and Applications Symposium*, pages 99–107. IEEE, 2009. doi:10.1109/RTAS.2009.20.
- [30] Baidu. Apollo. <http://apollo.auto/>, 2023. Accessed: Jun 6, 2023.
- [31] Shinpei Kato, Shota Tokunaga, Yuya Maruyama, Seiya Maeda, Manato Hirabayashi, Yuki Kitsukawa, Abraham Monroy, Tomohito Ando, Yusuke Fujii, and Takuya Azumi. Autoware on board: Enabling autonomous vehicles with embedded systems. In *2018 ACM/IEEE 9th International Conference on Cyber-Physical Systems (ICCPS)*, pages 287–296. IEEE, 2018. doi:10.1109/ICCPS.2018.00035.
- [32] Igor Bogoslavskyi and Cyrill Stachniss. Fast range image-based segmentation of sparse 3d laser scans for online operation. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 163–169. IEEE, 2016. doi:10.1109/IROS.2016.7759050.

- [33] Igor Bogoslavskyi and Cyrill Stachniss. Efficient online segmentation for sparse 3d laser scans. *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 85(1):41–52, 2017. doi:10.1007/s41064-016-0003-y.
- [34] Pei Sun, Henrik Kretzschmar, Xerxes Dotiwalla, Aurelien Chouard, Vijaysai Patnaik, Paul Tsui, James Guo, Yin Zhou, Yuning Chai, Benjamin Caine, et al. Scalability in perception for autonomous driving: Waymo open dataset. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2446–2454, 2020. doi:10.48550/arXiv.1912.04838.
- [35] LG Electronics America R&D Lab. LGSVL Simulator. <https://www.svl simulator.com/>, 2020. Accessed: Jun 6, 2023.
- [36] LG Electronics America R&D Lab. LG Silicon Valley Lab Apollo 5.0 Fork. <https://github.com/lgsvl/apollo-5.0>, 2019. Accessed: Jun 6, 2023.
- [37] Qinghua Yang, Hui Chen, Zhe Chen, and Junzhe Su. Introspective false negative prediction for black-box object detectors in autonomous driving. *Sensors*, 21(8):2819, 2021. doi:10.3390/s21082819.
- [38] Steven Posada, Office of Defects Investigation, NHTSA, USDOT. ODI RESUME. Investigation PE 21-020, 2021 [Online]. URL: <https://static.nhtsa.gov/odi/inv/2021/INOA-PE21020-1893.PDF>.
- [39] Praveena Penmetsa, Pezhman Sheinidashtegol, Aibek Musaev, Emmanuel Kofi Adanu, and Matthew Hudnall. Effects of the autonomous vehicle crashes on public perception of the technology. *IATSS Research*, 2021. doi:10.1016/j.iatssr.2021.04.003.
- [40] Wei Li, Yi Huang, Shichao Wang, and Xuecai Xu. Safety criticism and ethical dilemma of autonomous vehicles. *AI and Ethics*, 2(4):869–874, 2022. doi:10.1007/s43681-021-00128-2.
- [41] Peng Jing, Yunhao Cai, Baihui Wang, Bichen Wang, Jiahui Huang, Chengxi Jiang, and Chenglu Yang. Listen to social media users: Mining chinese public perception of automated vehicles after crashes. *Transportation Research Part F: Traffic Psychology and Behaviour*, 93:248–265, 2023. URL: <https://www.sciencedirect.com/science/article/pii/S1369847823000207>, doi:10.1016/j.trf.2023.01.018.
- [42] James Tu, Mengye Ren, Sivabalan Manivasagam, Ming Liang, Bin Yang, Richard Du, Frank Cheng, and Raquel Urtasun. Physically realizable adversarial examples for lidar object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 13716–13725, 2020. doi:10.48550/arXiv.2004.00543.
- [43] Dimity Miller, Peyman Moghadam, Mark Cox, Matt Wildie, and Raja Jurdak. What’s in the black box? the false negative mechanisms inside object detectors. *IEEE Robotics and Automation Letters*, 7(3):8510–8517, 2022. doi:10.1109/LRA.2022.3187831.
- [44] Peter Feiler, John Goodenough, Arie Gurfinkel, Charles Weinstock, and Lutz Wrage. Four pillars for improving the quality of safety-critical software-reliant systems. Technical report, Carnegie-Mellon Univ Pittsburg PA Software Engineering Inst, 2013.
- [45] Mats Heimdahl, Nancy Leveson, Julie Redler, Melanie Felton, and Grady Lee. Software assurance approaches, considerations, and limitations: Final report. Technical report, Technical Report DOT/FAA/TC-15/57. Federal Aviation Administration, US DOT, 2016.
- [46] Christian Menard, Andrés Goens, Marten Lohstroh, and Jeronimo Castrillon. Achieving determinism in adaptive AUTOSAR. In *2020 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pages 822–827. IEEE, 2020. doi:10.23919/DATE48585.2020.9116430.
- [47] Wojciech Baron, Anna Arestova, Christoph Sippl, Kai-Steffen Hielscher, and Reinhard German. Lett: An execution model for distributed real-time systems. In *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*, pages 1–7. IEEE, 2021. doi:10.1109/VTC2021-Fall152928.2021.9625343.

- [48] Reza Miroslou, Mohamed Hassan, and Rodolfo Pellizzoni. Duetto: Latency guarantees at minimal performance cost. In 2021 Design, Automation & Test in Europe Conference & Exhibition (DATE), pages 1136–1141. IEEE, 2021. doi:[10.23919/DATE51398.2021.9474062](https://doi.org/10.23919/DATE51398.2021.9474062).
- [49] Jaynarayan H Lala and Richard E Harper. Architectural principles for safety-critical real-time applications. Proceedings of the IEEE, 82(1):25–40, 1994. doi:[10.1109/5.259424](https://doi.org/10.1109/5.259424).
- [50] R. Hammett. Design by extrapolation: an evaluation of fault-tolerant avionics. In 20th DASC. 20th Digital Avionics Systems Conference (Cat. No.01CH37219), volume 1, pages 1C5/1–1C5/12 vol.1, Oct 2001. doi:[10.1109/DASC.2001.963314](https://doi.org/10.1109/DASC.2001.963314).
- [51] Rolf Isermann, Ralf Schwarz, and Stefan Stolz. Fault-tolerant drive-by-wire systems. IEEE Control Systems Magazine, 22(5):64–81, 2002. doi:[10.1109/MCS.2002.1035218](https://doi.org/10.1109/MCS.2002.1035218).
- [52] Massimo Baleani, Alberto Ferrari, Leonardo Mangeruca, Alberto Sangiovanni-Vincentelli, Maurizio Peri, and Saverio Pezzini. Fault-tolerant platforms for automotive safety-critical applications. In Proceedings of the 2003 international conference on Compilers, architecture and synthesis for embedded systems, pages 170–177. ACM, 2003. doi:[10.1145/951710.951734](https://doi.org/10.1145/951710.951734).
- [53] Hyunki Kim, Hyeuntae Lee, and Keyseo Lee. The design and analysis of avtmr (all voting triple modular redundancy) and dual-duplex system. Reliability Engineering & System Safety, 88(3):291–300, 2005. doi:[10.1016/j.ress.2004.08.012](https://doi.org/10.1016/j.ress.2004.08.012).
- [54] Oliver Rooks, Michael Armbruster, Armin Sulzmann, Gernot Spiegelberg, and Uwe Kiencke. Duo duplex drive-by-wire computer system. Reliability Engineering & System Safety, 89(1):71–80, 2005. doi:[10.1016/j.ress.2004.08.001](https://doi.org/10.1016/j.ress.2004.08.001).
- [55] J. Wei, J. M. Snider, J. Kim, J. M. Dolan, R. Rajkumar, and B. Litkouhi. Towards a viable autonomous driving research platform. In 2013 IEEE Intelligent Vehicles Symposium (IV), pages 763–770, June 2013. doi:[10.1109/IVS.2013.6629559](https://doi.org/10.1109/IVS.2013.6629559).
- [56] P. Koopman and M. Wagner. Autonomous vehicle safety: An interdisciplinary challenge. IEEE Intelligent Transportation Systems Magazine, 9(1):90–96, Spring 2017. doi:[10.1109/MITS.2016.2583491](https://doi.org/10.1109/MITS.2016.2583491).
- [57] Mogens Blanke, Michel Kinnaert, Jan Lunze, Marcel Staroswiecki, and Jochen Schröder. Diagnosis and fault-tolerant control, volume 2. Springer, 2006.
- [58] D Seto, Bruce H Krogh, Lui Sha, and A Chutinan. Dynamic control system upgrade using the simplex architecture. IEEE Control Systems Magazine, 18(4):72–80, 1998. doi:[10.1109/37.710880](https://doi.org/10.1109/37.710880).
- [59] Seong Kyung Kwon, Ji Hwan Seo, Jin-Woo Lee, and Kyoung-Dae Kim. An approach for reliable end-to-end autonomous driving based on the simplex architecture. In 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV), pages 1851–1856. IEEE, 2018. doi:[10.1109/ICARCV.2018.8581113](https://doi.org/10.1109/ICARCV.2018.8581113).
- [60] Dung T Phan, Radu Grosu, Nils Jansen, Nicola Paoletti, Scott A Smolka, and Scott D Stoller. Neural simplex architecture. In NASA Formal Methods: 12th International Symposium, NFM 2020, Moffett Field, CA, USA, May 11–15, 2020, Proceedings 12, pages 97–114. Springer, 2020. doi:[10.1007/978-3-030-55754-6_6](https://doi.org/10.1007/978-3-030-55754-6_6).
- [61] Xiaofeng Wang, Naira Hovakimyan, and Lui Sha. Rsimplex: A robust control architecture for cyber and physical failures. ACM Transactions on Cyber-Physical Systems, 2(4):1–26, 2018. doi:[10.1145/3121428](https://doi.org/10.1145/3121428).
- [62] Yanbing Mao, Yuliang Gu, Naira Hovakimyan, Lui Sha, and Petros Voulgaris. S11-simplex: Safe velocity regulation of self-driving vehicles in dynamic and unforeseen environments. ACM Transactions on Cyber-Physical Systems, 7(1):1–24, 2023. doi:[10.1145/3564273](https://doi.org/10.1145/3564273).
- [63] Patrick Musau, Nathaniel Hamilton, Diego Manzananas Lopez, Preston Robinette, and Taylor T Johnson. On using real-time reachability for the safety assurance of machine learning controllers. In 2022 IEEE International Conference on Assured Autonomy (ICAA), pages 1–10. IEEE, 2022. doi:[10.1109/ICAA52185.2022.00010](https://doi.org/10.1109/ICAA52185.2022.00010).

- [64] Julian Bernhard, Patrick Hart, Amit Sahu, Christoph Schöller, and Michell Guzman Cancimance. Risk-based safety envelopes for autonomous vehicles under perception uncertainty. In 2022 IEEE Intelligent Vehicles Symposium (IV), pages 104–111. IEEE, 2022. doi:[10.1109/IV51971.2022.9827199](https://doi.org/10.1109/IV51971.2022.9827199).
- [65] Guy Katz, Clark Barrett, David L Dill, Kyle Julian, and Mykel J Kochenderfer. Reluplex: An efficient smt solver for verifying deep neural networks. In International Conference on Computer Aided Verification, pages 97–117. Springer, 2017. doi:[10.1007/978-3-319-63387-9_5](https://doi.org/10.1007/978-3-319-63387-9_5).
- [66] Mohamad Gharib, Paolo Lollini, Marco Botta, Elvio Amparore, Susanna Donatelli, and Andrea Bondavalli. On the safety of automotive systems incorporating machine learning based components: a position paper. In 2018 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W), pages 271–274. IEEE, 2018. doi:[10.1109/DSN-W.2018.00074](https://doi.org/10.1109/DSN-W.2018.00074).
- [67] Aws Albarghouthi et al. Introduction to neural network verification. Foundations and Trends® in Programming Languages, 7(1–2):1–157, 2021.
- [68] Changliu Liu, Tomer Arnon, Christopher Lazarus, Christopher Strong, Clark Barrett, Mykel J Kochenderfer, et al. Algorithms for verifying deep neural networks. Foundations and Trends® in Optimization, 4(3-4):244–404, 2021. doi:[10.1561/24000000035](https://doi.org/10.1561/24000000035).
- [69] Hoang-Dung Tran, Stanley Bak, Weiming Xiang, and Taylor T Johnson. Verification of deep convolutional neural networks using imagestars. In International Conference on Computer Aided Verification, pages 18–42. Springer, 2020. doi:[10.1007/978-3-030-53288-8_2](https://doi.org/10.1007/978-3-030-53288-8_2).
- [70] Xiaowei Huang, Marta Kwiatkowska, Sen Wang, and Min Wu. Safety verification of deep neural networks. In International conference on computer aided verification, pages 3–29. Springer, 2017. doi:[10.1007/978-3-319-63387-9_1](https://doi.org/10.1007/978-3-319-63387-9_1).
- [71] Yuchi Tian, Kexin Pei, Suman Jana, and Baishakhi Ray. Deeptest: Automated testing of deep-neural-network-driven autonomous cars. In Proceedings of the 40th international conference on software engineering, pages 303–314, 2018. doi:[10.1145/3180155.3180220](https://doi.org/10.1145/3180155.3180220).
- [72] Peter Pinggera, Uwe Franke, and Rudolf Mester. High-performance long range obstacle detection using stereo vision. In 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 1308–1313. IEEE, 2015. doi:[10.1109/IROS.2015.7353537](https://doi.org/10.1109/IROS.2015.7353537).
- [73] You Li and Javier Ibanez-Guzman. Lidar for autonomous driving: The principles, challenges, and trends for automotive lidar and perception systems. IEEE Signal Processing Magazine, 37(4):50–61, 2020. doi:[10.1109/MSP.2020.2973615](https://doi.org/10.1109/MSP.2020.2973615).
- [74] Ricardo Roriz, Jorge Cabral, and Tiago Gomes. Automotive LiDAR technology: A survey. IEEE Transactions on Intelligent Transportation Systems, 2021. doi:[10.1109/TITS.2021.3086804](https://doi.org/10.1109/TITS.2021.3086804).
- [75] Yan Wang, Wei-Lun Chao, Divyansh Garg, Bharath Hariharan, Mark Campbell, and Kilian Q Weinberger. Pseudo-lidar from visual depth estimation: Bridging the gap in 3d object detection for autonomous driving. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 8445–8453, 2019. doi:[10.1109/CVPR.2019.00864](https://doi.org/10.1109/CVPR.2019.00864).
- [76] Michael Himmelsbach, Felix V Hundelshausen, and H-J Wuensche. Fast segmentation of 3d point clouds for ground vehicles. In 2010 IEEE Intelligent Vehicles Symposium, pages 560–565. IEEE, 2010. doi:[10.1109/IVS.2010.5548059](https://doi.org/10.1109/IVS.2010.5548059).
- [77] Dmitry Korchev, Shinko Cheng, Yuri Owechko, et al. On real-time lidar data segmentation and classification. In Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV), page 1. The Steering Committee of The World Congress in Computer Science, Computer ..., 2013.
- [78] Alireza Asvadi, Paulo Peixoto, and Urbano Nunes. Detection and tracking of moving objects using 2.5 d motion grids. In 2015 IEEE 18th International Conference on Intelligent Transportation Systems, pages 788–793. IEEE, 2015. doi:[10.1109/ITSC.2015.133](https://doi.org/10.1109/ITSC.2015.133).

- [79] Dimitris Zermas, Izzat Izzat, and Nikolaos Papanikolopoulos. Fast segmentation of 3d point clouds: A paradigm on lidar data for autonomous vehicle applications. In 2017 IEEE International Conference on Robotics and Automation (ICRA), pages 5067–5073. IEEE, 2017. doi:10.1109/ICRA.2017.7989591.
- [80] Jiyang Chen, Simon Yu, Rohan Tabish, Ayoosh Bansal, Shengzhong Liu, Tarek Abdelzaher, and Lui Sha. Lidar cluster first and camera inference later: A new perspective towards autonomous driving. arXiv preprint arXiv:2111.09799, 2021. doi:10.48550/arXiv.2111.09799.
- [81] Miguel Realpe, Boris X Vintimilla, and Ljubo Vlacic. Multi-sensor fusion module in a fault tolerant perception system for autonomous vehicles. In 2nd International Conference on Robotics and Artificial Intelligence, Los Angeles, USA, 2016.
- [82] Paolo Burgio, Marko Bertogna, Nicola Capodieci, Roberto Cavicchioli, Michal Sojka, Přemysl Houdek, Andrea Marongiu, Paolo Gai, Claudio Scordino, and Bruno Morelli. A software stack for next-generation automotive systems on many-core heterogeneous platforms. Microprocessors and Microsystems, 52:299–311, 2017. doi:10.1016/j.micpro.2017.06.016.
- [83] Nvidia Corporation. Nvidia drive os. <https://developer.nvidia.com/drive/drive-software>driveav, 2020. Accessed: Jun 6, 2023.
- [84] Long Chen, Yunqing Zhang, Bin Tian, Yunfeng Ai, Dongpu Cao, and Fei-Yue Wang. Parallel driving os: A ubiquitous operating system for autonomous driving in cpss. IEEE Transactions on Intelligent Vehicles, 7(4):886–895, 2022. doi:10.1109/TIV.2022.3223728.
- [85] Jing Chen, Cong Zhao, Shengchuan Jiang, Xinyuan Zhang, Zhongxin Li, and Yuchuan Du. Safe, efficient, and comfortable autonomous driving based on cooperative vehicle infrastructure system. International journal of environmental research and public health, 20(1):893, 2023. doi:10.3390/ijerph20010893.
- [86] Yali Yuan, Robert Tasik, Sripriya Srikant Adhatarao, Yachao Yuan, Zheli Liu, and Xiaoming Fu. Race: reinforced cooperative autonomous vehicle collision avoidance. IEEE Transactions on Vehicular Technology, 2020. doi:10.1109/TVT.2020.2974133.
- [87] Zhong Cao, Shaobing Xu, Xinyu Jiao, Huei Peng, and Diange Yang. Trustworthy safety improvement for autonomous driving using reinforcement learning. Transportation research part C: emerging technologies, 138:103656, 2022. doi:10.1016/j.trc.2022.103656.
- [88] Kibeom Lee and Dongsuk Kum. Collision avoidance/mitigation system: Motion planning of autonomous vehicle via predictive occupancy map. IEEE Access, 7:52846–52857, 2019. doi:10.1109/ACCESS.2019.2912067.
- [89] Hong Wang, Yanjun Huang, Amir Khajepour, Yubiao Zhang, Yadollah Rasekhipour, and Dongpu Cao. Crash mitigation in motion planning for autonomous vehicles. IEEE transactions on intelligent transportation systems, 20(9):3313–3323, 2019. doi:10.1109/TITS.2018.2873921.
- [90] Haseeb Tahir, Mujahid N Syed, and Uthman Baroudi. Heuristic approach for real-time multi-agent trajectory planning under uncertainty. IEEE Access, 8:3812–3826, 2019. doi:10.1109/ACCESS.2019.2962785.
- [91] Samyeul Noh and Kyoungwan An. Decision-making framework for automated driving in highway environments. IEEE Transactions on Intelligent Transportation Systems, 19(1):58–71, 2017. doi:10.1109/TITS.2017.2691346.
- [92] Samyeul Noh. Decision-making framework for autonomous driving at road intersections: Safeguarding against collision, overly conservative behavior, and violation vehicles. IEEE Transactions on Industrial Electronics, 66(4):3275–3286, 2018. doi:10.1109/TIE.2018.2840530.
- [93] Donghoon Shin, Beomjun Kim, Kyongsu Yi, Ashwin Carvalho, and Francesco Borrelli. Human-centered risk assessment of an automated vehicle using vehicular wireless communication. IEEE Transactions on Intelligent Transportation Systems, 20(2):667–681, 2018. doi:10.1109/TITS.2018.2823744.

- [94] Ming-Yuan Yu, Ram Vasudevan, and Matthew Johnson-Roberson. Occlusion-aware risk assessment for autonomous driving in urban environments. IEEE Robotics and Automation Letters, 4(2):2235–2241, 2019. doi:10.1109/LRA.2019.2900453.
- [95] Guofa Li, Yifan Yang, Tingru Zhang, Xingda Qu, Dongpu Cao, Bo Cheng, and Keqiang Li. Risk assessment based collision avoidance decision-making for autonomous vehicles in multi-scenarios. Transportation research part C: emerging technologies, 122:102820, 2021. doi:10.1016/j.trc.2020.102820.
- [96] Peter Seiler, Bongsob Song, and J Karl Hedrick. Development of a collision avoidance system. SAE transactions, pages 1334–1340, 1998. URL: <https://www.jstor.org/stable/44741070>.
- [97] Joseph Funke, Matthew Brown, Stephen M Erlien, and J Christian Gerdes. Collision avoidance and stabilization for autonomous vehicles in emergency scenarios. IEEE Transactions on Control Systems Technology, 25(4):1204–1216, 2016. doi:10.1109/TCST.2016.2599783.
- [98] Xiangkun He, Yulong Liu, Chen Lv, Xuewu Ji, and Yahui Liu. Emergency steering control of autonomous vehicle for collision avoidance and stabilisation. Vehicle system dynamics, 57(8):1163–1187, 2019. doi:10.1080/00423114.2018.1537494.
- [99] Shuo Cheng, Liang Li, Hong-Qiang Guo, Zhen-Guo Chen, and Peng Song. Longitudinal collision avoidance and lateral stability adaptive control system based on mpc of autonomous vehicles. IEEE Transactions on Intelligent Transportation Systems, 21(6):2376–2385, 2019. doi:10.1109/TITS.2019.2918176.
- [100] Reza Hajiloo, Mehdi Abroshan, Amir Khajepour, Alireza Kasaiezadeh, and Shih-Ken Chen. Integrated steering and differential braking for emergency collision avoidance in autonomous vehicles. IEEE Transactions on Intelligent Transportation Systems, 2020. doi:10.1109/TITS.2020.2984210.
- [101] Sean Summers and John Lygeros. Verification of discrete time stochastic hybrid systems: A stochastic reach-avoid decision problem. Automatica, 46(12):1951–1961, 2010. doi:10.1016/j.automatica.2010.08.006.
- [102] Sean Summers, Maryam Kamgarpour, John Lygeros, and Claire Tomlin. A stochastic reach-avoid problem with random obstacles. In Proceedings of the 14th international conference on Hybrid systems: computation and control, pages 251–260, 2011. doi:10.1145/1967701.1967738.
- [103] Noah J Goodall. Ethical decision making during automated vehicle crashes. Transportation Research Record, 2424(1):58–65, 2014. doi:10.3141/2424-07.
- [104] Ayoosh Bansal, Jayati Singh, Micaela Verucchi, Marco Caccamo, and Lui Sha. Risk ranked recall: Collision safety metric for object detection systems in autonomous vehicles. In 2021 10th Mediterranean Conference on Embedded Computing (MECO), pages 1–4, 2021. doi:10.1109/MECO52532.2021.9460196.
- [105] Christian Schmidt, Fred Oechsle, and Wolfgang Branz. Research on trajectory planning in emergency situations with multiple objects. In 2006 IEEE Intelligent Transportation Systems Conference, pages 988–992. IEEE, 2006. doi:10.1109/ITSC.2006.1707153.
- [106] Yu Ming, Yanqiang Li, Zihui Zhang, and Weiqi Yan. A survey of path planning algorithms for autonomous vehicles. SAE international journal of commercial vehicles, 14(1):02–14, 2021. doi:10.4271/02-14-01-0007.
- [107] American Association of State Highway and Transportation Officials. Policy on geometric design of highways and streets. American Association of State Highway and Transportation Officials, Washington, DC, 1(990):158, 2001.
- [108] Mario Bijelic, Tobias Gruber, and Werner Ritter. A benchmark for lidar sensors in fog: Is detection breaking down? In 2018 IEEE Intelligent Vehicles Symposium (IV), pages 760–767. IEEE, 2018. doi:10.1109/IVS.2018.8500543.
- [109] Robin Heinzler, Philipp Schindler, Jürgen Seekircher, Werner Ritter, and Wilhelm Stork. Weather influence and classification with automotive lidar sensors. In 2019 IEEE Intelligent Vehicles Symposium (IV), pages 1527–1534. IEEE, 2019. doi:10.1109/IVS.2019.8814205.

- [110] Andrew M Wallace, Abderrahim Halimi, and Gerald S Buller. Full waveform lidar for adverse weather conditions. *IEEE transactions on vehicular technology*, 69(7):7064–7077, 2020. doi:10.1109/TVT.2020.2989148.
- [111] You Li, Pierre Duthon, Michèle Colomb, and Javier Ibanez-Guzman. What happens for a tof lidar in fog? *IEEE Transactions on Intelligent Transportation Systems*, 22(11):6670–6681, 2020. doi:10.1109/TITS.2020.2998077.
- [112] Mario Pavlić, Heidrun Belzner, Gerhard Rigoll, and Slobodan Ilić. Image based fog detection in vehicles. In *2012 IEEE Intelligent Vehicles Symposium*, pages 1132–1137. IEEE, 2012. doi:10.1109/IVS.2012.6232256.
- [113] Philip Sallis, Clemens Dannheim, Christian Icking, and Markus Maeder. Air pollution and fog detection through vehicular sensors. In *2014 8th Asia Modelling Symposium*, pages 181–186. IEEE, 2014. doi:10.1109/AMS.2014.43.
- [114] Derek J Straub et al. Detecting the presence of fog using low-cost proximity sensors. *Aerosol and Air Quality Research*, 20(5):981–990, 2020. doi:10.4209/aaqr.2019.11.0612.
- [115] Kim Larson. Dust detection and warning system tracks its first season. <https://azdot.gov/adot-blog/dust-detection-and-warning-system-tracks-its-first-season>, 2020. Accessed: Jun 6, 2023.
- [116] Răzvan-Cătălin Miclea, Vlad-Ilie Ungureanu, Florin-Daniel Sandru, and Ioan Silea. Visibility enhancement and fog detection: Solutions presented in recent scientific papers with potential for application to mobile systems. *Sensors*, 21(10):3370, 2021. doi:10.3390/s21103370.
- [117] Jonathan Petit, Bas Stottelaar, Michael Feiri, and Frank Kargl. Remote attacks on automated vehicles sensors: Experiments on camera and lidar. *Black Hat Europe*, 11(2015):995, 2015.
- [118] Hocheol Shin, Dohyun Kim, Yujin Kwon, and Yongdae Kim. Illusion and dazzle: Adversarial optical channel exploits against lidars for automotive applications. In *International Conference on Cryptographic Hardware and Embedded Systems*, pages 445–467. Springer, 2017. doi:10.1007/978-3-319-66787-4_22.
- [119] Yulong Cao, Chaowei Xiao, Dawei Yang, Jing Fang, Ruigang Yang, Mingyan Liu, and Bo Li. Adversarial objects against lidar-based autonomous driving systems. *arXiv preprint arXiv:1907.05418*, 2019. doi:10.48550/arXiv.1907.05418.
- [120] Jiachen Sun, Yulong Cao, Qi Alfred Chen, and Z Morley Mao. Towards robust {LiDAR-based} perception in autonomous driving: General black-box adversarial sensor attack and countermeasures. In *29th USENIX Security Symposium (USENIX Security 20)*, pages 877–894, 2020. doi:10.48550/arXiv.2006.16974.
- [121] Hugo Weichel. *Laser beam propagation in the atmosphere*, volume 10319. SPIE press, 1990.
- [122] Sanna Kaasalainen, Hannu Hyypä, Antero Kukko, Paula Litkey, Eero Ahokas, Juha Hyypä, Hubert Lehner, Anttoni Jaakkola, Juha Suomalainen, Altti Akujärvi, et al. Radiometric calibration of lidar intensity with commercially available reference targets. *IEEE Transactions on Geoscience and Remote Sensing*, 47(2):588–598, 2008. doi:10.1109/TGRS.2008.2003351.
- [123] Isaac I Kim, Bruce McArthur, and Eric J Korevaar. Comparison of laser beam propagation at 785 nm and 1550 nm in fog and haze for optical wireless communications. In *Optical wireless communications III*, volume 4214, pages 26–37, 2001. doi:10.1117/12.417512.
- [124] Răzvan-Cătălin Miclea, Ciprian Dughir, Florin Alexa, Florin Sandru, and Ioan Silea. Laser and lidar in a system for visibility distance estimation in fog conditions. *Sensors*, 20(21):6322, 2020. doi:10.3390/s20216322.
- [125] Baidu. Apollo Cyber RT. <https://cyber-rt.readthedocs.io/en/latest/>, 2022. Accessed: Jun 6, 2023.
- [126] The ROS Community. Robot Operating System. <https://ros.org/>, 2022. Accessed: Jun 6, 2023.
- [127] LG Electronics America R&D Lab. Lidar Sensor Plugin. <https://www.svl simulator.com/docs/archive/2020.06/lidar-plugin/>, 2020. Accessed: Jun 6, 2023.
- [128] COCO Consortium. Detection evaluation. <https://cocodataset.org/detection-eval>, 2021. Accessed: Jun 6, 2023.
- [129] Velodyne Lidar. Puck. <https://velodynelidar.com/products/puck/>, 2016. Accessed: Jun 6, 2023.