Sensor phenomena

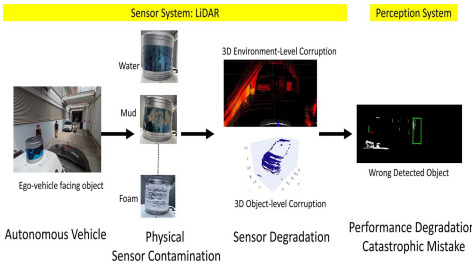
LIDAROC: Realistic LiDAR Cover Contamination Dataset for Enhancing Autonomous Vehicle Perception Reliability

Grafika Jati[1*], Martin Molan[1], Francesco Barchi[1], Andrea Bartolini[1*], Giuseppe Mercurio[2], and Andrea Acquaviva[1*]

[1]*Department of Electrical, Electronic and Information Engineering, University of Bologna, 40136 Bologna, Italy*
[2]*FEV Italia s.r.l., 40033 Bologna, Italy*
[*]*Member, IEEE*

Manuscript received 30 April 2024; revised 28 June 2024; accepted 9 July 2024. Date of publication 29 July 2024; date of current version 20 August 2024.

Abstract—LiDAR is the foundation of many autonomous vehicle perception systems, so it is essential to study and ensure the integrity and robustness of the data collected by LiDAR. To facilitate future research into robust and resilient LiDAR processing, we present a dataset containing a collection of uncontaminated and realistically contaminated LiDAR samples. We have also studied the effect of contaminants on the object detection task. The state-of-the-art object detection algorithms produce catastrophic errors in detection, such as failure to identify objects, detection of ghost objects, and wrong detection with high confidence. Based on the number of such catastrophic errors, we introduce a novel measure for the LiDAR data's contamination level. The results of the empirical evaluation of the effect of the contaminants on object detection motivate the necessity of further research into contaminant detection and contaminant-resilient data processing, which are all enabled by the dataset collected by this work.



Index Terms—Sensor phenomena, anomaly, autonomous vehicle, contamination, dataset, LiDAR corruption, object detection benchmark, perception robustness testing, sensor.

I. INTRODUCTION

To ensure the safety of all traffic participants, autonomous driving systems must be robust and resilient against various types of data corruption and environmental contamination [1], [2]. A critical task reliant on LiDAR data is 3-D object detection, which encounters limitations in scenarios with unforeseen abnormalities [3].

Utilizing real-world data is essential, as is considering additional corruption types, particularly those that may occur in an autonomous vehicle's sensors. For example, sensor cover contamination is a potential external factor, yet datasets with real contaminated LiDAR data to test sensor performance using state-of-the-art object detection methods are currently unavailable.

Existing realistic contamination datasets primarily focus on environmental factors such as weather and are limited to specific scenarios, making accurate contamination characterization challenging [4]. Our letter introduces a dataset containing multiple real physical contaminants, such as water, dust, uniform mud, drop mud, lubricant/oil, saltwater, and foam, applied in three severity levels: 1) low; 2) mid; and 3) high. This dataset is instrumental in analyzing the geometric and reflectance properties of LiDAR sensors under real-world conditions and aids in developing appropriate contamination detection methods to trigger effective cleaning procedures.

Corresponding author: Grafika Jati (e-mail: grafika.jati2@unibo.it).
Associate Editor: Francisco Falcone.
The dataset and supplement materials are available at https://gitlab.com/ecs-lab/lidaroc.
Digital Object Identifier 10.1109/LSENS.2024.3434624

II. RELATED WORK: EXISTING CORRUPTED DATA

In focusing on corruption datasets, the KITTI, nuScenes, and Waymo datasets are initially produced in pristine conditions. Dong et al. [3] then introduced synthetic corruption to create variants like KITTI-C, nuScenes-C, and Waymo-C. However, these adaptations do not fully represent real-world corruption scenarios. Another dataset, the LIBRE dataset, compares multiple 3-D LiDAR datasets using 12 different LiDAR sensors [5]. It captures data under adverse weather conditions like fog, rain, and strong light, but does not address sensor cover contamination.

Rivero et al. [6] characterized and simulated the impact of road dirt on the transmission and reflection properties of the LiDAR sensor covers. Trierweiler et al. [7] further explored how dust accumulation on sensor covers can reduce a LiDAR's maximum range by up to 75%. However, these studies focused on isolated contamination events and did not explore dynamic contamination scenarios. James et al. [8] delved into the effects of dirt, salt, and frost on LiDAR sensors, but it was limited to a single sensor position without placing target objects in front of the sensor to simulate complete cover contamination. Besides, the previous study do not present dataset, it also does not discuss about contamination affecting the detection of target object shapes.

Schlager et al. [9] experimentally measured the impact of physical damage, such as scratches, cracks, and holes on sensor covers. The findings showed a reduction in the LiDAR's range, but did not assess how much damage alters the point cloud at the object level. The most recent study involving sensor cover contamination was conducted by Schlager et al. [10], which used one static position, simple single target, and one level contamination. While still leaveing gaps in understanding how more complex and dynamic environments affect LiDAR data integrity under contamination.

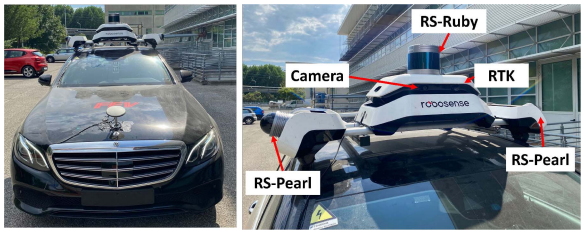


Fig. 1. Ego-vehicle with sensors.

Compared to existing datasets, our work presents the first dataset in which different sensor cover contaminants are studied in a realistic and dynamic environment and where the effect of the contaminants on critical automotive applications (autonomous driving) is quantified.

III. LIDAR COVER CONTAMINATION SETUP

a) Multisensor and Ego-vehicle: The data is collected using automotive grade RS-Ruby as the main LiDAR sensor with specification: 128-channel, minimum 0.1° Vertical Angular Resolution, Horizontal resolution 0.2°, 250 m range, Horizontal FoV 360°, Vertical FoV 40°, and frame rate 10 Hz. The LiDAR generates at least 2.3 million points/second and supports L4+ autonomous driving. The use of high-resolution LiDAR in this test represents a new analysis, as the most recent trials under varying weather conditions by Lambert et al. [11] use commercial 64-channel LiDAR. The system also has two RS-pearl LiDAR installed as a blind spot LiDAR and six cameras AR0233-5200-GMSL2 with resolution 1920 × 1080, which are installed in front right, front left, right, left, back right, and back left. The sensor acquisition system also has RTK, Millimeter Wave Radar, CAN, IPC, and mounting brackets. This sensor acquisition system mounted on our real ego-vehicle with a Robosense Reference System to generate calibrated data sensor and editable Ground Truth for 2-D/3-D object detection is shown in Fig. 1.

b) Multi and Various Object Target: The relevant targets are placed in front of the ego-vehicle at various distances and include various shapes and materials: car, motorcycle, pedestrian, and whiteboards with three different covers (original white, black sheet, and aluminum sheet). We introduce fake pedestrians made of cardboard to stress object detection, inspired by [12]. This case is also inspired by Elhafsi et al. [13], which introduced an anomalous event or edge case in terms of semantic anomalies. Imagine that a car detects an advertising board with a pedestrian figure pretending to be a real pedestrian; this causes the car to make the wrong decision. Furthermore, in this letter, we present a novel challenge for distinguishing human-shaped fake pedestrians from real pedestrians. Regarding using a cardboard pedestrian model, the material does not impact sensing performance; instead, the flat human shape poses a detection challenge. Existing object detection algorithms trained on real human forms find it more difficult to identify these models. The setting environment and target from the perspective of the ego-vehicle car and acquired LiDAR and camera data can be seen in Fig. 2.

c) Multienvironment: The experiment was carried out in two environments: A subterranean narrow hallway with the target about 5 m away, called the 5 m data, simulating a complex city driving scenario. The second environment was a spacious outdoor area with two distance variations. We consider car stopping distance range from 12 m (for 20 mph) to 23 m (for 30 mph) [14]. Considering a 20 mph city speed limit [15], targets were placed at 10 and 20 m, referred to as the 10 m and 20 m datasets. Fig. 2 shows the ego-vehicle with the sensor facing the target objects in these environments.

d) Multicontamination: We contaminated the main RS-Ruby LiDAR cover with various fluid and nonfluid contaminants, including water, dust, uniform mud, drop mud, engine lubricant/oil, salt water, and foam. Each contaminant is categorized into three levels: 1) low; 2) medium; and 3) high. For water, salt, and lubricant, we use one spray for low, three for medium, and five for high levels. Mud has two types: 1) uniform, where the sensor is evenly covered and 2) drop, where wet mud is dripped onto the sensor. For salt and mud, we create a cover on the sensor and let it dry, simulating dried mud and salt on a car. Water, dust, and lubricant are applied just before data collection. Each measurement is taken for 60 s and repeated three times.

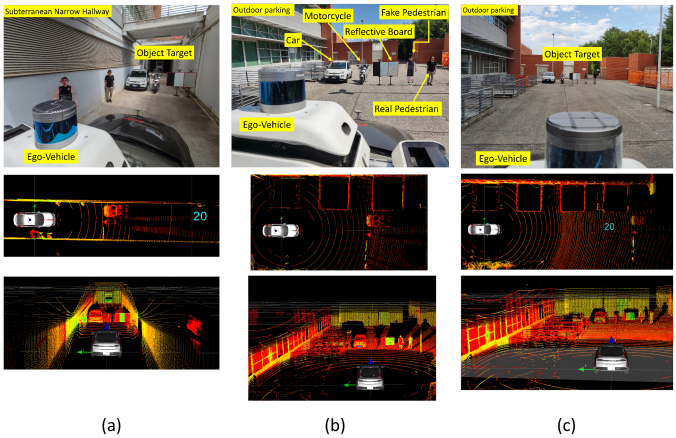


Fig. 2. Top-row: setting environment, Middle-row: Clean LiDAR point cloud from top view, and Bottom-row: Clean LiDAR point from bird's-eye view. Target object has various point intensity value depend on material and distance. The Red in RGB (255, 0, 0) indicates the lowest intensity. (a) 5 m. (b) 10 m. (c) 20 m.

Fig. 3 top-row shows the contamination applied to the LiDAR sensor. Then, Fig. 3 middle-row depicts the effect of each contamination in the LiDAR point cloud, indicating a reduction in intensity at each point even in some cases, the points disappeared. Fig. 4 shows the intensity of a target board. The intensity of contaminant is lower than the baseline (clean and cover), where *mud* is the most significant corruption. We observe that the low-level contamination has a higher intensity than the mid and high levels. Cover contamination creates an environment- and object-level corruption. Fig. 3 middle-row shows that contamination becomes an environment-level corruption, reducing the intensity of the whole point. We find new insight that contamination also creates object-level corruption. Fig. 3 bottom-row shows corruption change the 3-D shape of the car.

IV. BENCHMARK DETECTION ROBUSTNESS ON CONTAMINATED LIDAR

The main sensing task is object detection [2]. So, we empirically evaluate the state-of-the-art object detection on real contaminated LiDAR and find catastrophic mistakes indicated by misdetection with high confidence. We use the PointRCNN object detector [16] since it is the best performance in terms of the lowest RCE (relative corruption error) by measuring the performance drop over synthetic corruption [3]. We used PointRCNN trained using KITTI taken from [17]. We choose car detection as one of the relevant objects in the autonomous driving environment. Since this letter focuses on evaluating the significance of contamination effects on detection, we measure the distance between the center point (X, Y) of the detected bounding box (BB) and the ground truth (GT). In Tables 1 and 2, some measurements were not performed due to technical constraints, indicated as *no_exp*.

a) Detection accuracy: Contamination creates instability in detection, indicated by the number of frames in which nothing is detected. Therefore, detection accuracy is calculated as the total number of

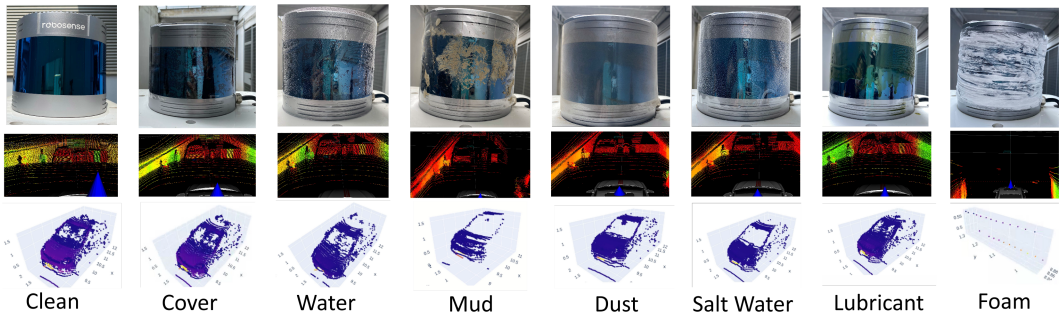


Fig. 3. Top row: Procedure contaminating LiDAR in 5 m dataset: 1. clean (baseline), 2. cover-clean (baseline) 3. water, 4. mud, 5. dust, 6. salt, 7. lubricant/engine oil, and 8. foam. Middle row: 3-D environment-level corruption. Bottom row: 3-D object-level corruption.

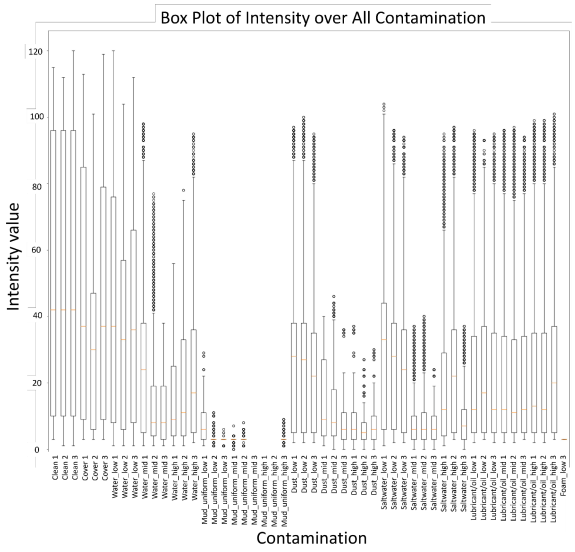


Fig. 4. Box plot of intensity value on reflective board over all contamination in 5 m data.

Table 1. Detection and Localization Accuracy of PointRCNN

No	Contaminant	5m		10m		20m	
		Detection	Localization	Detection	Localization	Detection	Localization
1	Clean (Baseline)	100%	100%	100%	100%	100%	100%
2	Cover_clean (Baseline)	100%	100%	100%	100%	100%	100%
3	Water_low	100%	100%	100%	100%	100%	100%
4	Water_mid	100%	100%	100%	100%	100%	100%
5	Water_high	100%	100%	100%	100%	100%	100%
6	Mud_drop_low	no exp	no exp	100%	100%	100%	100%
7	Mud_drop_mid	no exp	no exp	100%	100%	100%	100%
8	Mud_drop_high	no exp	no exp	100%	100%	100%	100%
9	Mud_uniform_low	100%	91%	100%	100%	100%	98%
10	Mud_uniform_mid	61%	7%	0%	0%	0%	0%
11	Mud_uniform_high	33%	33%	no exp	no exp	no exp	no exp
12	Dust_low	100%	100%	100%	100%	100%	100%
13	Dust_mid	100%	100%	100%	100%	100%	100%
14	Dust_high	100%	100%	100%	100%	100%	100%
15	Salt_low	100%	100%	100%	100%	100%	100%
16	Salt_mid	100%	100%	100%	100%	100%	100%
17	Salt_high	100%	100%	100%	100%	100%	100%
18	Lubricant_low	100%	100%	100%	100%	100%	100%
19	Lubricant_mid	100%	100%	no exp	no exp	100%	100%
20	Lubricant_high	100%	100%	no exp	no exp	100%	100%
21	Foam_low	0%	0%	no exp	no exp	no exp	no exp

frames minus the instability. The lower accuracy indicates the contamination severity on object detection. Based on Table 1, we observed that *mud_uniform* significantly reduces the detection accuracy. The *mud_uniform_high* in 5 m data, only get 33%. The most severe contamination is *foam*, resulting in the accuracy of 0%. The *foam* reduces the number and shape of the point cloud, making the object detector unable to detect objects. Fig. 3 bottom-row, *foam* remove car's point cloud. Concerning the object's distance, it is shown that *mud_uniform_mid* exacerbates detection accuracy issues at 10 m and 20 m with 0% detection. Such occurrences seriously threaten the autonomous vehicle as the inability to detect objects can lead to catastrophic consequences for passenger and environmental safety.

Table 2. Catastrophic: Inaccurate Localization and Ghost Object

No	Contaminant	5m		10m		20m	
		Innacurate Loc	Ghost Detect	Innacurate Loc	Ghost Detect	Innacurate Loc	Ghost Detect
1	Clean (Baseline)	0.67	0.00	0.00	0.00	8.67	1.00
2	Cover_clean (Baseline)	0.33	0.00	0.00	0.00	18.00	1.33
3	Water_low	0.00	0.00	0.00	0.00	23.00	6.00
4	Water_mid	2.33	0.33	0.67	0.00	38.33	6.67
5	Water_high	1.00	0.00	0.67	0.00	40.33	7.33
6	Mud_drop_low	no exp	no exp	0.00	0.00	29.50	5.00
7	Mud_drop_mid	no exp	no exp	0.33	0.00	15.00	1.00
8	Mud_drop_high	no exp	no exp	0.67	0.00	10.00	0.00
9	Mud_uniform_low	55.00	42.67	5.67	1.33	97.33	24.33
10	Mud_uniform_mid	3.33	0.00	0.00	0.00	0.00	0.00
11	Mud_uniform_high	0.00	0.00	no exp	no exp	no exp	no exp
12	Dust_low	0.00	0.00	0.00	0.00	20.00	3.00
13	Dust_mid	0.67	0.33	1.00	0.00	25.00	6.00
14	Dust_high	2.00	0.00	0.00	0.00	43.00	4.00
15	Salt_low	0.33	0.00	0.00	0.00	49.00	7.00
16	Salt_mid	3.67	0.00	0.00	0.00	36.00	6.00
17	Salt_high	1.67	0.00	0.33	0.00	29.00	5.00
18	Lubricant_low	0.33	0.00	0.00	0.00	38.00	5.00
19	Lubricant_mid	0.00	0.00	no exp	no exp	15.00	1.00
20	Lubricant_high	0.00	0.00	no exp	no exp	86.00	17.00
21	Foam_low	0.00	0.00	no exp	no exp	no exp	no exp

The value indicates average number of case, the lower the better.

b) Localization accuracy: However, the emphasis shifts towards achieving precise localization identification beyond mere detection. The priority shifts from merely detecting the object to accurately pinpointing its location within the environment. Therefore, while detection serves as an initial step, achieving accurate localization adds a layer of reliability to autonomous vehicle perception. When the detected BB and GT are less than or equal to 1.0 and 0.5 m for the X and Y axes, respectively, the detected label is *car* with model confidence at least *0.7* of *1.0*, considered correctly localized. Localization accuracy is then calculated as the number of frames localized correctly over all tested frames. Based on Table 1, localization accuracy decreases in *mud_uniform* for the 5 m, 10 m, and 20 m data.

Inaccurate localization occurs when the detector detects an object with a confidence level at least *0.7*, but the center location of the detected bounding box is far from the GT, which is more than 1.0 m and 0.5 m for the X and Y axes, respectively. Based on Table 2, comparing the 10 m and 20 m datasets, there is an increase in inaccurate detection in the 20 m dataset, where all contaminants lead to an increase in inaccurate detection, with *mud_uniform* and *lubricant* being the most influential contaminants.

c) Ghost Object: Contaminants reduce detection capabilities, cause inaccurate localization, and can pose adversarial attacks on object detectors. Table 2 shows instances of ghost/anomaly detection, where the detector makes inaccurate detection with high confidence at least *0.9*. This indicates high confidence in results despite errors. For example, $mud\_uniform\_low$ results in 42.67 ghost detections of all tested frames. Ghost detection is more severe in the 20 m dataset, with $mud\_uniform\_low$ accounting for 24 cases and $lubricant\_high$ for 17 cases. Although $lubricant\_high$ achieves 100% detection and localization (Table 1), it leads to ghost detection where no objects exist. Examples of catastrophic mistakes can be seen in Fig. 5. Fig. 5(a)

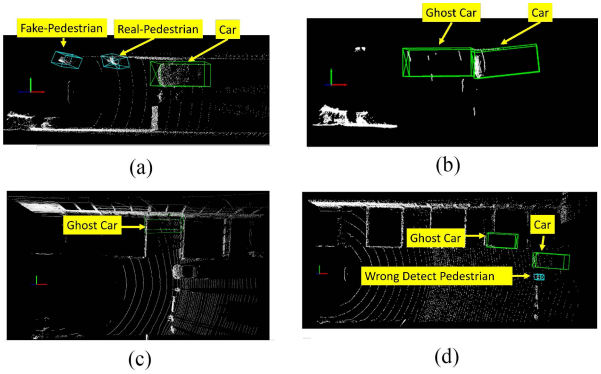


Fig. 5. Catastrophic mistake: Detect Ghost Object with high confidence. Green box as car, blue box as pedestrian. (a) Data 5 m: Clean cover (baseline). (b) Data 5 m: Catastrophic mistake: Detecting ghost car in mud_uniform_low. (c) Data 10 m: Undetected car while detecting ghost car in mud_uniform_low. (d) Data 20 m: Detecting ghost car, undetected pedestrian, detect motorcycle as pedestrian in lubricant_high.

shows the detection results on clean 5 m data as a baseline example, where the detection results for cars are indicated by green bounding boxes. In subfigure (a), the object detector can accurately identify the car. In Fig. 5(b), ghost detection occurs where the car is detected twice when only one car exists. In Fig. 5(c), the real car is missed, and a false detection occurs elsewhere. Fig. 5(d) also displays ghost objects. These examples highlight the issue of ghost detection under certain conditions.

We introduce novel cases of fake pedestrians, where the detector makes mistakes by detecting fake pedestrians as real. In Fig. 5(a), there are two blue bounding boxes where there should only be one. Then, in Fig. 5(d), there is a wrong detection in which the motorcycle is detected as a pedestrian. The analysis of pedestrian detection is not further elaborated in this letter and can be a future work of this cover-contaminated dataset.

V. CONCLUSION

Through detailed empirical research, we have uncovered significant negative impacts of contaminants on current object detection methodologies, underscoring the need for specialized solutions. As demonstrated in the empirical evaluation, models trained solely on clean data falter under contaminated conditions, making incorporating contaminated data during the training phase imperative. Furthermore, our dataset establishes a standard for evaluating perception degradation in adverse conditions, supports the development of generative AI simulations that realistically model these scenarios, and facilitates testing in sensor fusion with adaptive weighting [18], [19], and [20].

The dataset we have developed will be instrumental in assessing and crafting a sensor fusion model inspired by recent advancements in the field. In our dataset, while the main LiDAR sensor experiences corruption, the auxiliary LiDAR sensors and all cameras remain unaffected. This setup mirrors real-world challenges where the reliability of individual sensors is compromised. Consequently, our dataset is invaluable for training models to be resilient against sensor corruption. This letter code and dataset will be available upon request at https://gitlab.com/ecs-lab/lidaroc.

ACKNOWLEDGMENT

This work was supported in part by EU NRRP (PNRR) under Grant DM 352/2022, and in part by the Next Generation EU initiative, the EU Horizon project Edge AI Technologies for Optimised Performance Embedded Processing: EdgeAI under Grant 101097300. The authors would like to thank FEV Italia s.r.l. for the collaboration and data acquisition hardware and also thank Giammarco Cecchini for helping during data acquisition.

REFERENCES

[1] P. H. Chan, G. Dhadyalla, and V. Donzella, "A framework to analyze noise factors of automotive perception sensors," *IEEE Sens. Lett.*, vol. 4, no. 6, Jun. 2020, Art. no. 6001004.
[2] A. Pandharipande et al., "Sensing and machine learning for automotive perception: A review," *IEEE Sensors J.*, vol. 23, no. 11, pp. 11097–11115, Jun. 2023.
[3] Y. Dong et al., "Benchmarking robustness of 3D object detection to common corruptions," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2023, pp. 1022–1032.
[4] J. Ren, L. Pan, and Z. Liu, "Benchmarking and analyzing point cloud classification under corruptions," in *Proc. Int. Conf. Mach. Learn.*, PMLR, 2022, pp. 18559–18575.
[5] A. Carballo et al., "LIBRE: The multiple 3 D LiDAR dataset," in *Proc. 2020 IEEE Intell. Veh. Symp. (IV)*, IEEE, 2020, pp. 1094–1101.
[6] J. R. V. Rivero et al., "Characterization and simulation of the effect of road dirt on the performance of a laser scanner," in *Proc. 2017 IEEE 20th Int. Conf. Intell. Transp. Syst. (ITSC)*, 2017, pp. 1–6.
[7] M. Trierweiler, P. Caldelas, G. Gröninger, T. Peterseim, and C. Neumann, "Influence of sensor blockage on automotive LiDAR systems," in *Proc. IEEE SENSORS*, 2019, pp. 1–4.
[8] J. K. James, G. Puhlfürst, V. Golyanik, and D. Stricker, "Classification of LiDAR sensor contaminations with deep neural networks," in *Proc. Comput. Sci. Cars Symp. Future Challenges Artif. Intell. Secur. Auton. Vehicles*, Munich, Germany. ACM, New York, NY, USA, Sep. 2018, p. 8. [Online]. Available: https://doi.org/10.1145/3273946.3273952
[9] B. Schlager, T. Goelles, and D. Watzenig, "Effects of sensor cover damages on point clouds of automotive LiDAR," in *Proc. IEEE Sensors*, 2021, pp. 1–4.
[10] B. Schlager, T. Goelles, S. Muckenhuber, and D. Watzenig, "Contaminations on LiDAR sensor covers: Performance degradation including fault detection and modeling as potential applications," *IEEE Open J. Intell. Transp. Syst.*, vol. 3, pp. 738–747, 2022.
[11] J. Lambert et al., "Performance analysis of 10 models of 3D LiDARs for automated driving," *IEEE Access*, vol. 8, pp. 131699–131722, 2020.
[12] J.-A. Bolte, A. Bar, D. Lipinski, and T. Fingscheidt, "Towards corner case detection for autonomous driving," in *Proc. 2019 IEEE Intell. Veh. Symp. (IV)*, 2019, pp. 438–445.
[13] A. Elhafsi et al., "Semantic anomaly detection with large language models," *Auton. Robots*, vol. 47, no. 8, pp. 1035–1055, 2023.
[14] Brake, "Stopping distances," 2024. Accessed: Jun. 25, 2024. [Online]. Available: https://www.brake.org.uk/get-involved/take-action/mybrake/knowledge-centre/speed/stopping-distances
[15] G. F. Nightingale, "Evaluating the citywide edinburgh 20mph speed limit intervention effects on traffic speed and volume: A pre-post observational evaluation," *PLoS One*, vol. 16, no. 12, 2021, Art. no. e0261383.
[16] S. Shi, X. Wang, and H. Li, "PointRCNN: 3D object proposal generation and detection from point cloud," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 770–779.
[17] O. D. Team, "OpenPCDet: An open-source toolbox for 3D object detection from point clouds," 2020. [Online]. Available: https://github.com/open-mmlab/OpenPCDet
[18] R. Ravindran, M. J. Santora, and M. M. Jamali, "Camera, LiDAR, and radar sensor fusion based on Bayesian neural network (CLR-BNN)," *IEEE Sensors J.*, vol. 22, no. 7, pp. 6964–6974, Apr. 2022.
[19] J. Wang, Y. Lu, and H. Jiang, "FAFNs: Frequency-aware LiDAR–camera fusion networks for 3-D object detection," *IEEE Sensors J.*, vol. 23, no. 24, pp. 30847–30857, Dec. 2023.
[20] Y. Huang et al., "Improving robustness of LiDAR-camera fusion model against weather corruption from fusion strategy perspective," 2024, *arXiv:2402.02738*.

Open Access funding provided by 'Alma Mater Studiorum – Università di Bologna' within the CRUI CARE Agreement