NTNU
Kunnskap for en bedre verden

Department of Computer Science

TDT39 - Empirical Studies

Improving Safety of Self-Driving Vehicles in the CARLA Simulator

Author:
Aksel Lunde Aase

Title of the study	Improving Safety of Self-Driving Vehicles in the CARLA Simulator
Name of the student	Aksel Lunde Aase
Number of words	1196

9th December 2022

Purpose

Today, multiple car manufacturers are developing self-driving vehicles (SDVs) [1][2][3]. SDVs come in many variants, and are often classified using the SAE International Levels of Automation [4]. Today's cars commonly include simple driver assistants like Cruise Control [5], which corresponds to level 1. At level 2, the system is able to control both steering and acceleration simultaneously. Several manufacturers are currently working on producing level 3 capable SDVs [6, 7], which is characterized by the ability for the driver pay attention to other activities, for instance watching a movie.

But with higher levels comes greater risk, due to the reduction of human oversight. In one survey concerning SDVs, Boston Consulting Group report that "Concerns About the Safety of SDVs Are a Significant Hurdle" for adoption [8]. Another study shows that people generally demand safer-than-human driving before they trust SDVs on public infrastructure [9].

Unfortunately, these strict criteria are not yet met. Collision report data from SDVs in California reveal a crash rate that is three to five times that of human drivers [10]. As the authors note, these SDVs even have the benefit of only driving in the ideal weather conditions of sunny California. Since sensor data quality can be significantly degraded by rainy or snowy conditions [11], it is fair to assume that the relative collision rate would be even worse in these conditions.

Thus, motivated by the current lack of sufficient safety in SDVs, the purpose of this research is to evaluate and improve the safety of SDVs. We choose to focus on improving safety in harsh winter conditions to offset the imbalance of current research in clear conditions. Specifically, we intend to evaluate and compare the deep learning-based autonomous driver TransFuser [12] in both snowy and clear conditions in the CARLA simulator [13]. We modify TransFuser to recognize the current weather conditions and automatically apply weather-dependent safety rules to improve it's safety score.

With this objective in mind, we propose the following research questions:

- **RQ 1.1** How does the original TransFuser (model **A**) perform when driving in adverse weather conditions?
- **RQ 1.2** How does a re-trained TransFuser (model **B**) perform after being trained on data containing adverse weather conditions?
- **RQ 2.1** Can a modified TransFuser trained to predict weather conditions (model **C**) accurately recognize which conditions it is currently driving in?
- **RQ 2.2** Can weather-specific safety constraints improve the safety of the agent when driving in adverse weather conditions?

Contributions

This research will bring the following contributions:

1. Implementation of snowy weather in the CARLA simulator, as well as effects like slippery roads and reduced sensory range.
2. Modifications to the TransFuser agent to predict weather conditions.
3. Weather-specific rules to limit the action space of TransFuser to safer choices.
4. Quantitative comparisons of model **A**, **B** and **C** in various weather conditions, that shows how the weather and weather-specific rules impact the SDVs' safety and driving performance.

Considering the current lack of adverse weather in CARLA, these contributions are a novel addition to the simulator and related literature. The contributions will enable further research into the effects of adverse weather, and increase focus on snowy climates by implementing winter conditions in an already popular research tool.

Research Method

This research quantitatively evaluates the performance of variants of TransFuser in the CARLA simulator. As is common in the literature [12, 14, 15], we compare agents using the Driving Score (DS) and Infraction Score (IS) as defined by the CARLA Leaderboard [16]. The latter metric specifically measures the agents' safety performance. However it is meaningless to improve IS alone, since simply standing still will yield perfect IS but zero DS. We therefore aim to simultaneously improve IS and DS.

Unlike the literature however, we do not submit agents to the CARLA Leaderboard for evaluation, since this online evaluation does not contain any maps with adverse weather. Instead, we modify the Longest6 benchmark as defined by TransFuser by adding snow and ice, thereby creating the Longest6-Winter benchmark. We then evaluate model **A**, **B** and **C** on both Longest6 and Longest6-Winter. This enables us to measure impact on safety in both clear and adverse conditions, and will answer **RQ** 1.1, 1.2, and 2.2. **RQ** 2.1 will be answered by measuring the accuracy of model **C**'s predictions.

The predictions of neural networks may be significantly influenced by the random initialization of parameters during training. As shown by Reimers, training and evaluating deep-learning models once is therefore not sufficient to claim statistical significant comparisons [17]. In TransFuser, the authors account for this variance by training their model with three different random seeds, and report metrics for an ensamble of these three training runs. However, this still only produces one evaluation sample, and is therefore prone to the same statistical errors of single-run-comparisons that Reimers present. We therefore take additional care to evaluate the significance of our findings.

Specifically, we hypothesise that weather-specific rules affecting decisions such as driving speed and distance to the vehicle in front will improve the IS in winter conditions. To support this claim, we perform experiments by training multiple independent instances of each of the three variants by randomly choosing the initial seed, and use Welch's t-test [18] with $\alpha = 0.05$ to determine significance of performance differences between the three model populations. This test is justified by the fact that the DS and IS metrics are calculated as averages over 36 independent routes in the Longest6-Winter benchmark, and then again averaged across all model instances, thus the Central Limit Theorem [19] applies and fulfills the required normality condition.

A custom training dataset will be generated using TransFuser's method with added snow. This is required since no current dataset from CARLA contain adverse weather, as this is not yet implemented. The dataset will be generated and agents will be trained using the IDUN cluster at NTNU [20].

Participants and Ethical Considerations

This research is conducted by myself, together with my thesis supervisor Frank Lindseth and co-supervisor Gabriel Kiss. Data generation and processing happens purely digitally, with no external subjects involved.

However, this does not relieve the research from ethical considerations, especially not within the fields of machine learning and autonomous driving. Image classification models have previously been shown to exhibit sexist and racist behaviours [21]. SDVs have been the subject of several ethical discussions. For instance, if a fatal accident is inevitable, should the vehicle preferentially protect its passengers or pedestrians [22]? If not carefully implemented, SDVs might preferentially hurt racial minorities due to imbalances in the dataset. Answers to these problems are difficult to formulate exactly, but being aware of these issues is one step towards removing the researchers' inherent bias from transferring to the models.

Research Paradigm

In this research we adopt a positivist approach. We begin by reproducing the results from the original TransFuser agent to establish the validity of our implementation. The impact of adverse weather on driving safety is then quantitatively measured, and we measure the effect of weather-specific safety rules. The results are subjected to statistical tests to see whether the rules demonstrably affect the agents safety performance.

Final Delivery and Dissemination

This research will produce open-source code modifications to the CARLA simulator and TransFuser agent. These modifications and the results will be discussed in a written master thesis, and finally presented in a thesis defence. We hope to integrate the modifications in the CARLA Leaderboard to ensure future research is also conducted with adverse weather in mind.

Bibliography

[1] Tesla. 'Autopilot and full self-driving capability', [Online]. Available: https://www.tesla.com/support/autopilot (visited on 28th Nov. 2022).

[2] Waymo. 'Home - waymo'. (2022), [Online]. Available: https://waymo.com/ (visited on 28th Nov. 2022).

[3] Honda. 'Honda, cruise and gm take next steps toward autonomous vehicle mobility service business in japan'. (2021), [Online]. Available: https://global.honda/newsroom/news/2021/c210120eng.html (visited on 28th Nov. 2022).

[4] S. International. 'Automated driving'. (2014), [Online]. Available: https://www.sae.org/binaries/content/assets/cm/content/news/press-releases/pathway-to-autonomy/automated_driving.pdf (visited on 28th Nov. 2022).

[5] Wikipedia. 'Adaptive cruise control', [Online]. Available: https://en.wikipedia.org/wiki/Adaptive_cruise_control (visited on 29th Nov. 2022).

[6] 'Honda very carefully sets loose its level 3 autonomous car'. (2021), [Online]. Available: https://www.slashgear.com/honda-very-carefully-sets-loose-its-level-3-autonomous-car-04662291/ (visited on 29th Nov. 2022).

[7] 'Mercedes-benz wins world's first approval for level 3 autonomous cars: What's that mean?' (2021), [Online]. Available: https://www.slashgear.com/782536/mercedes-benz-wins-worlds-first-approval-for-level-3-autonomous-cars-whats-that-mean/ (visited on 29th Nov. 2022).

[8] B. C. Group. 'Self-driving vehicles, robo-taxis, and the urban mobility revolution'. (2016), [Online]. Available: https://www.bcg.com/publications/2016/automotive-public-sector-self-driving-vehicles-robo-taxis-urban-mobility-revolution (visited on 28th Nov. 2022).

[9] M. A. Nees, 'Safer than the average human driver (who is less safe than me)? examining a popular safety benchmark for self-driving cars', *Journal of Safety Research*, vol. 69, pp. 61–68, 2019, ISSN: 0022-4375. DOI: https://doi.org/10.1016/j.jsr.2019.02.002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022437518304511.

[10] R. L. McCarthy, 'Autonomous Vehicle Accident Data Analysis: California OL 316 Reports: 2015–2020', *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg*, vol. 8, no. 3, Sep. 2021, 034502, ISSN: 2332-9017. DOI: 10.1115/1.4051779. eprint: https://asmedigitalcollection.asme.org/risk/article-pdf/8/3/034502/6762299/risk_008_03_034502.pdf. [Online]. Available: https://doi.org/10.1115/1.4051779.

[11] N. Goberville *et al.*, 'Analysis of lidar and camera data in real world weather conditions for autonomous vehicle operations', *SAE Int. J. Advanc. Curr. Prac. Mobil*, vol. 2, pp. 2428–2434, 2020.

[12] K. Chitta, A. Prakash, B. Jaeger, Z. Yu, K. Renz and A. Geiger, 'Transfuser: Imitation with transformer-based sensor fusion for autonomous driving', *Pattern Analysis and Machine Intelligence (PAMI)*, 2022.

[13] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez and V. Koltun, 'CARLA: An open urban driving simulator', in *Proceedings of the 1st Annual Conference on Robot Learning*, 2017, pp. 1–16.

[14] H. Shao, L. Wang, R. Chen, H. Li and Y. Liu, 'Safety-enhanced autonomous driving using interpretable sensor fusion transformer', *arXiv preprint arXiv:2207.14024*, 2022.

[15] Papers With Code. 'Carla leaderboard benchmark (autonomous driving)'. (2022), [Online]. Available: https://paperswithcode.com/sota/autonomous-driving-on-carla-leaderboard (visited on 29th Nov. 2022).

[16] Carla. 'Carla autonomous driving leaderboard'. (2022), [Online]. Available: https://leaderboard.carla.org/leaderboard/ (visited on 28th Nov. 2022).

[17] N. Reimers and I. Gurevych, 'Why comparing single performance scores does not allow to draw conclusions about machine learning approaches', *arXiv preprint arXiv:1803.09578*, 2018.

[18] Wikipedia. ‘Welch’s t-test’. (2022), [Online]. Available: https://en.wikipedia.org/wiki/Welch%27s_t-test (visited on 29th Nov. 2022).

[19] Wikipedia. ‘Central limit theorem’. (2022), [Online]. Available: https://en.wikipedia.org/wiki/Central_limit_theorem (visited on 29th Nov. 2022).

[20] M. Själander, M. Jahre, G. Tufte and N. Reissmann, *EPIC: An energy-efficient, high-performance GPGPU computing research infrastructure*, 2019. arXiv: 1912.05848 [cs.DC].

[21] A. Yapo and J. Weiss, ‘Ethical implications of bias in machine learning’, 2018.

[22] G. Contissa, F. Lagioia and G. Sartor, ‘The ethical knob: Ethically-customisable automated vehicles and the law’, *Artificial Intelligence and Law*, vol. 25, no. 3, pp. 365–378, Sep. 2017, ISSN: 1572-8382. DOI: 10.1007/s10506-017-9211-z. [Online]. Available: https://doi.org/10.1007/s10506-017-9211-z.