DARWIN AI

Removing Regulatory Barriers for Vehicles with Automated Driving Systems

DarwinAI Comments
49 CFR Part 571
[Docket No. NHTSA-2019-0036]
[RIN 2127-AM00]

Table of Contents

List of Figures

A. About DarwinAI

DarwinAI is an artificial intelligence startup based in Waterloo, Ontario, Canada. The company was founded in 2017 and received funding from Obvious Ventures and iNovia Capital. The team has won numerous awards including best workshop papers at the renowned NeurIPS conference.

DarwinAI was co-founded by Dr. Alexander Wong. Dr. Wong is currently the Canada Research Chair in Artificial Intelligence and Medical Imaging, Member of the College of the Royal Society of Canada, co-director of the Vision and Image Processing Research Group, and an associate professor in the Department of Systems Design Engineering at the University of Waterloo. He has published over 500 refereed journal and conference papers, as well as 30 patents, in various fields such as computational imaging, artificial intelligence, computer vision, and pattern recognition.

Dr. Wong is an editor-in-chief, section editor, and associate editor on a number of peer-reviewed international scientific journals and on chair and program committee member on a number of international scientific conferences, and is currently a committee member on the Partnership on AI's (PAI) Annotation and Benchmarking on Understanding and Transparency of Machine Learning Lifecycles (ABOUT ML) committee for defining best practices for transparency in machine learning.

DarwinAI’s patent-pending technology enables safe autonomous vehicles by increasing the speed and explainability of the underlying neural networks that power Automated Driving Systems Dedicated Vehicles (ADS-DV) decision making. The platform is able to identify and explain the root causes and rationale of decisions made by a self-driving neural network. It also speeds up ADS-DV neural network decision making by as much as 38x. DarwinAI's partners include Nvidia and Intel.

DarwinAI is currently working with top automotive manufacturers and Tier 1 suppliers on their Autonomous Vehicle and Advanced Driver Assistance Systems. We are keen to apply our experience on the frontlines to help develop ADS-DV neural networks and provide feedback to NHTSA.

B. Question 9 Response

- For compliance testing methods that replace physical tests with non-physical requirements (simulation, documentation):

    - If the test method is used to determine compliance with a real-world test, how can NHTSA validate the accuracy of a simulation or documentation? If NHTSA must run real-world tests to validate a simulation or documentation, what is the advantage of non-physical requirements over these other compliance methods?

- **Answer:**

    - *The use of reliable simulations is imperative as testing a large number of practical scenarios under varying circumstances using "real world" data is simply not realistic. The accuracy of a simulation can be validated by cross-checking the behavior of the self-driving vehicle under simulation with its behavior under real-world tests using explainability tools. By checking the type of errors that occur during simulation as well as the type of errors that occur during real-world tests, a company can uncover biases in the simulation. Once it's been validated that the simulation exhibits similar behavior as in real-world tests, the former does not need to be done again. In sum, simulations can reflect real-world scenarios if careful thought is put into designing and implementing them.*

Figure 1 - Explainability example

- Figure 1 *illustrates the type of explainability that would be helpful in validating the accuracy of a simulation. The ADS-DV is put through a simulation with highway driving data. Parts of the source images are shaded by the explainability tool to indicate important areas regarding the decision made by the ADS-DV. The blue shaded area highlights the portions of the image that caused ADS-DV to turn right and the red shaded area highlight the portions that caused ADS-DV to turn left. This type of visualization is immensely helpful in understanding whether the ADS-DV is making the correct decisions for the appropriate reasons. If the decisions are incorrect, the ADS-DV can be supplemented with additional data to correct its behavior.*

    In this analysis, for example, the data illustrates that the self-driving vehicle ignores lane *markings when deciding to turn right or left. Determining the factors the system is using to make its decision is imperative is ensuring safe an accurate behavior in real-life scenarios.*

C. Question 10 Response

- Would non-physical requirements simply replicate the existing physical tests in a virtual world? If not, what would be the nature of the non-physical requirements (that is, what performance metrics would these requirements use, and how would NHTSA measure them)? Are there ways that NHTSA could amend the FMVSSs to remove barriers to ADS-DVs that would not require using the compliance test methods described in below?

- **Answer:**

  - *Non-physical requirements should not simply replicate existing physical tests in the virtual world, but instead extend the permutation of tested scenarios that would be too tedious or difficult to measure in the real world. For example, in the simulation in* Figure 2 *, the system is able to simulate a particular scenario across a multitude of visibility and weather conditions.*

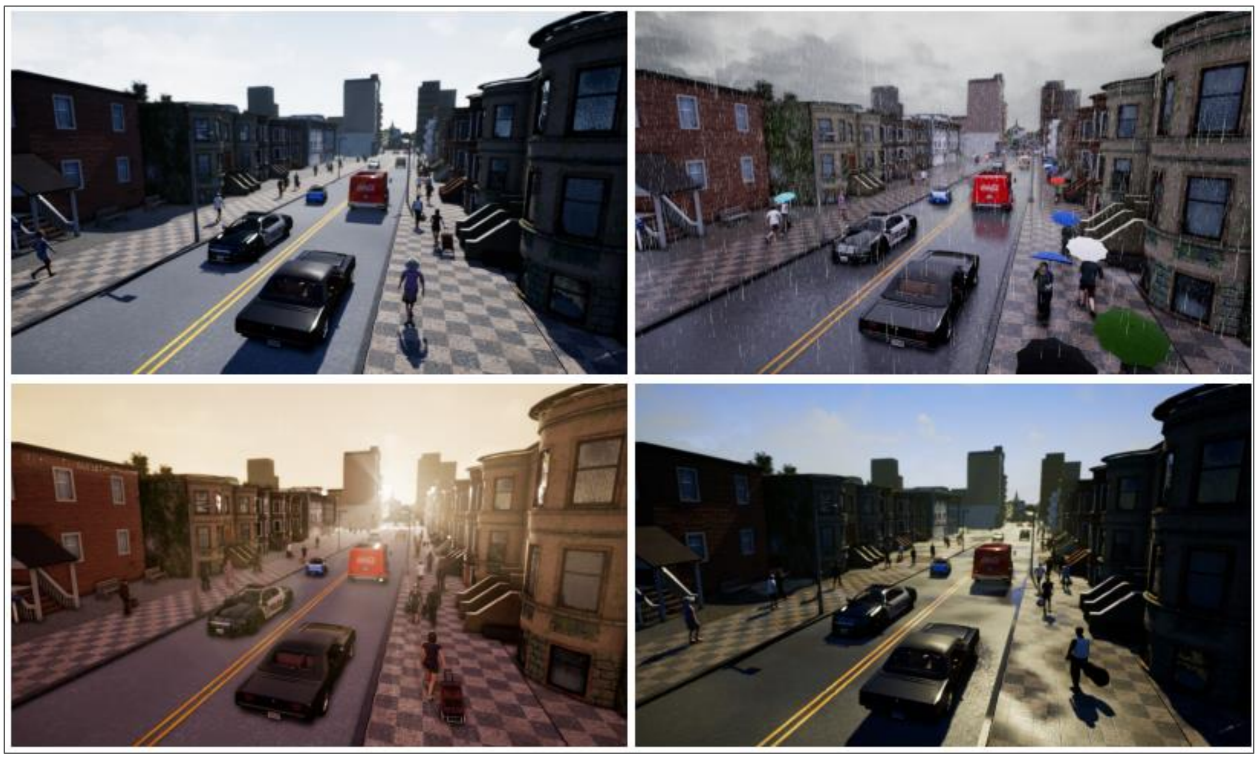

Figure 2 - CARLA[1] Simulation

  - *Moreover, if the errors resulting from the simulated and real-life tests are identical, one can replace the former with the latter to save cost and resources. The underlying criteria depends on the specific task; some useful metrics we've seen our clients utilize are speed, velocity, and distance to stop.*

[1] CARLA: Open Source Simulator for Autonomous Driving Research (Carla.org)

Another useful metric is time to acceleration and deceleration. This is the time from when the ADS-DV sensors or cameras perceive an event or object to when the ADS-DV begins to accelerate or decelerate. This is an example of a non-physical requirement that can be easily determined under simulation but is quite difficult to evaluate in the real world. In particular, measuring the time to acceleration and deceleration over long distances under different road conditions.

The test for the 'sufficiency' of an ADS-DV system consist of more than a rote application with a single binary test and static set of circumstances. Instead, it should provide insight into how the ADS-DV system makes decisions as well as its strengths, weaknesses, and potential errors.

In sum, the non-physical requirements appropriate for simulation ideally entail measuring behavior across a large number of scenarios under different environmental conditions.

D. Question 14 Response

- Will all ADS-DVs without traditional manual controls be capable of receiving and acting upon simple commands not consisting of a street address based destination, such as “drive forward or backwards a distance of 10 feet and stop”; “shift from park to drive and accelerate to 25 mph”; “drive up onto a car hauler truck trailer”; etc.? Please explain projected challenges for ADS-DVs without traditional manual controls to complete discrete driving commands and tasks.

- **Answer:**

  - *In order for ADS-DVs to be capable of receiving and acting upon simple commands without a street address strong natural language processing (NLP) capabilities would be required. Critical in this process is the ability to understand* ***how*** *the ADS-DV system interprets specific instructions such that gaps in the NLP model can be eradicated. Identifying the nuance of commands such as ‘drive up onto a car hauler truck trailer’ is critical for safety.*

    Moreover, translating verbal commands into correct underlying actions requires identifying biases in the data and algorithms that power the self-driving vehicle, as well as the root-causes of errors in the decision-making process of these algorithms. One can leverage explainability tools to extract the underlying rationale in how decisions are made by such algorithms when interpreting commands.

E. Question 15 Response

- How would NHTSA ensure that the performance of the ADS-DV during testing is consistent with how the vehicle would perform during actual normal use?

- **Answer:**

    - *NHTSA can ensure performance by employing dynamic simulation during testing and adopting continuous explainability beyond the confines of an individual series of tests. Moreover, there is an opportunity to review AI Primitives to ensure that the underlying models are free of bias as is possible by providing data and model auditing insights. This combination of examining outcomes ('black-box' testing) and examining the inner workings of models and their primitives ('white-box' testing) will ensure that the performance observed in simulation is consistent with performance during actual normal use.*

    - *For reference, training, validation and test data can be used when evaluating ADS-DV.*
        - *Training data: used to train a model during the ADS-DV development process. Manufacturers will typically have vast datasets spanning numerous scenarios*
        - *Validation data: used to validate the performance of the model during the ADS-DV development process on the types of scenarios that the ADS-DV will be testing. For example, this could include a single image of a cityscape*
        - *Test data: used for the actual performance testing*

    - ***Dynamic Simulation**: The benefit of virtual simulations is that they can be dynamically generated. Instead of relying on a fixed set of scenarios, it's possible to have a fine-grain control to tune any number of parameters in the virtual simulations to best mimic the real world.*
    - *This is achieved by NHTSA providing the validation data to manufacturers but withholding the test data. The validation data would be sufficient to help the manufacturers ensure that their ADS-DV systems can be tested. The provided validation set should include diverse scenarios across all test criteria categories but should be much smaller in size than the actual withheld test data to encourage fairness and un-biasness in the way manufacturers build their models to confirm with test criteria.*
    - *The simulator can add many different permutations to the test data. For example, it can take the same street and add in any number of pedestrians, other vehicles, and even change sunlight and weather to better reflect the full scope of scenarios that would be visible in the real-world. Moreover, the simulation itself could be adapted on -the-fly based on how the ADS-DV is driving. For example, if the ADS-DV is erring towards lower speeds, the simulation could try to introduce other vehicles that overtake the ADS-DV to see how it responds.*

    - ***Continuous Explainability**: A potential paradigm shift is the notion of 'continuous explainability'. In most cases, traditional 'diagnostic explainability' examines how the system makes decisions and behaves based on historical data and identifying root causes. However, 'continuous explainability', implies collecting data during normal use and explaining immediately whether the results are correct. Such capabilities would allow car manufacturers to explain actions in real-time for actual use.*

- *NHTSA could then audit decisions made by the ADS-DV retrospectively to understand how the ADS-DV performs on real-world data. Through 'continuous explainability' NHTSA could identify problematic scenarios where the vehicle either behaves differently despite the similarity to the scenario to what it was tested on (e.g., car decides to accelerate towards a parked garbage truck, despite the fact that one of the previous test scenarios where there is a parked firetruck and the car successfully passed that test by decelerating), or where the scenario deviates significantly from all past test scenarios so one can observe new behavior that is rare but can occur in the real world (e.g., tree suddenly falling at a distance in front of the vehicle). NHTSA could share these problematic scenarios back with manufacturers to help them understand where to improve their systems. These scenarios could then be fed back into simulations so that all manufacturers can benefit.*

- ***AI Primitives Review**: An ADS-DV's performance is constrained by the training data that was used to develop these vehicles. NHTSA can provide model tools for manufacturers to audit their data and models to gain greater insights into the biases that can occur under different scenarios, such as identifying where such biases lie and their probable causes such that manufacturers can better account for them and improve their data and model.*

F. Question 31 Response

- Are there objective, practicable ways for the agency to validate simulation models to ensure their accuracy and repeatability?

- **Answer:**

  - *An objective, practical way for the agency to validate simulation models to ensure they behave consistently with real-life expectations is to leverage explainability tools. Such tools can provide concrete insights into the behavior of such models and their decision-making processes across a wide variety of scenarios, conditions, and environments.*

  - Figure 3*, for example, illustrates a dog hanging out of a car window. If the ADS-DV has been trained to stop every time it perceives a dog, it may stop erroneously when encountering this car driving on the road.*

Figure 3 - Dog ‘edge case’ example (Source: Creative Commons[2])

  - *If the agency tested ADS-DV systems against such ‘edge cases’, companies would have to ensure that such systems were trained on simulated data that resemble, or ideally were identical to, real-life scenarios.*

[2] Image Attribution: *"Furry Marley" by flubberwinkle is licensed under CC BY-NC-SA 2.0*

Sincerely,

Dr. Alexander Wong, Ph.D.
Canada Research Chair in Artificial Intelligence and Medical Imaging
Co-founder and Chief Scientist, DarwinAI

Date: August 27, 2019