10-29-2022

Toward a framework for mitigating object detection decision making bias in driverless cars

Swarnamouli Majumdar
University of Massachusetts Lowell, Swarnamouli_Majumdar@student.uml.edu

Amit V. Deokar
University of Massachusetts Lowell, amit_deokar@uml.edu

Toward a Framework for Mitigating Object Detection Bias in Driverless Cars

Research-in-Progress

Swarnamouli Majumdar
Manning School of Business
University of Massachusetts Lowell
Swarnamouli_Majumdar@student.uml.edu

Amit V. Deokar
Manning School of Business
University of Massachusetts Lowell
Amit_Deokar@uml.edu

ABSTRACT

Bias in machine learning is a significant problem that demands industry-wide attention, and in the case of driverless vehicles, life and death are at stake. The debate is whether autonomous vehicles are safe, yet more likely to strike a pedestrian of color than a person of white skin color. It is essential to obtain a greater understanding of the algorithmic bias that occurs during driving-centric object recognition. Major automakers plan to develop cars with a degree of autonomy between Level 4 and Level 5 within the next decade. At Level 5, the system is meant to behave similarly to a human driver; it can drive anyplace lawful and can make independent decisions (SAE International 2016).

The design of autonomous vehicles must undergo a comprehensive ethical analysis. The U.S. government developed a checklist for self-driving vehicles that includes ethical considerations and choices made intentionally and consciously (Kang 2016). These algorithms must be provided to the United States Department of Transportation through the National Highway Traffic Safety Administration in a transparent manner. However, these efforts do not clearly address the problem of bias and how it may emerge in an AV's decision-making process. This research attempts to understand algorithmic bias that emerges during driving-centric object recognition. This project is opportune since several jurisdictions have recently permitted autonomous cars to run on public roads and several businesses are testing their AVs on public roads, often resulting in a notable number of fatalities. Autonomous systems are vulnerable to algorithmic bias because they do not need an active or passive person to identify and correct such biases (Danks & London, 2017). To reduce bias, design and implementation methodologies can be enhanced in accordance with new ethical norms (Arkin et al., 2011). We are researching cognitive heuristics to develop a procedural framework for minimizing object detection bias in AV model training. The aim is to continue efforts to reduce object detection bias that affects autonomous system decision-making to promote safety and increase technology adoption. Autonomous cars must evaluate safe operating options when a collision can't be prevented. This may persist due to the fluidity and unpredictability of human-automobile interactions. Artificial intelligence (AI) systems are biased because to their reliance on data, individuals, and model design. Data bias has been connected to obstacles associated with the deployment of ethical AI (Kohli & Chadha, 2019) and has detrimental effects on underrepresented demographic groups. This study addresses training data bias, algorithmic processing bias, transfer context bias, and interpretation bias. The first two biases relate to the computational aspects of the problem. The latter two relate to the improper deployment of algorithms and autonomous systems. If the training data are condition-specific, the model cannot make precise judgments in the actual world. This study describes a method for identifying, assessing, and mitigating such bias.

Keywords

Machine learning, autonomous systems, object detection, decision-making bias, algorithmic bias, process framework

References

1. Arkin, R. C., Ulam, P., & Wagner, A. R. (2011). Moral decision making in autonomous systems: Enforcement, moral emotions, dignity, trust, and deception. *Proceedings of the IEEE*, *100*(3), 571-589.
2. Danks, D., & London, A. J. (2017). Algorithmic Bias in Autonomous Systems. In *IJCAI* (Vol. 17, pp. 4691-4697).
3. Kirkpatrick, K. (2016). Battling algorithmic bias: how do we ensure algorithms treat us fairly?. *Communications of the ACM*, *59*(10), 16-17.
4. Kohli, P., & Chadha, A. (2019). Enabling pedestrian safety using computer vision techniques: A case study of the 2018 uber inc. self-driving car crash. In *Future of Information and Communication Conference* (pp. 261-279). Springer, Cham.
5. Schwarting, W., Alonso-Mora, J., & Rus, D. (2018). Planning and decision-making for autonomous vehicles. *Annual Review of Control, Robotics, and Autonomous Systems*, 1, 187-210.
6. Forward, S. E. (2009). The theory of planned behaviour: The role of descriptive norms and past behaviour in the prediction of drivers' intentions to violate. *Transportation Research Part F: traffic psychology and behaviour*, 12(3), 198-207.