Modelling riders' intervention behavior during high-level autonomous driving under extreme conditions: Insights from a VR-enabled simulation study

A. Zheng Xu[a], B. Nan Zheng[a,*], C. Yihai Fang[a], D. Hai L. Vu[a]

[a] Department of Civil Engineering, Monash University, Melbourne, VIC, Australia
a.Zheng.Xu3@monash.edu, b.Nan.Zheng@monash.edu
[*] Corresponding author

Extended abstract submitted for presentation at the Conference in Emerging Technologies in Transportation Systems (TRC-30)
September 02-03, 2024, Crete, Greece

April 28, 2024

Keywords: Autonomous driving system, Intervention behavior, Extreme conditions, Virtual reality

1 INTRODUCTION

The development of the autonomous driving system (ADS) predominantly focuses on technical aspects aimed at accident prevention to enhance the overall safety in self-driving (Nair & Bhat, 2021). Although mainstream research on ADS safety has achieved notable improvements in the safety performance of autonomous vehicles (AVs) (Feng *et al.*, 2023), a significant gap remains between the intelligence of AVs and user acceptance of autonomous driving, particularly under extreme conditions (de Melo *et al.*, 2019). For instance, our previous study revealed that most human riders are likely to intervene in autonomous driving scenarios they perceive as risky, even though the ADS assesses them as safe (Xu *et al.*, 2024a). This suggests that continuous enhancements to the ADS may not necessarily contribute to the widespread adoption of AVs, as riders' intervention behavior could hinder the seamless operation of the ADS.

In this study, we seek to highlight the differences in decision-making between the ADS and passengers in AVs under extreme crash conditions, to identify the specific factors that trigger intervention actions. We reproduce three critical fatal road accidents from Australian roads, as documented by Australian Car Crash Media (CarCrashMedia, 2021), within a high-fidelity VR environment. In these simulations, we substitute the vehicles involved in the crashes with fine-tuned AVs operating with high-level autonomy (i.e., no human supervision or intervention is required) to assess if the collisions would unfold similarly. We then invite 60 participants with varied demographic backgrounds to partake in a human-in-the-loop analysis by experiencing these scenarios within the simulated AVs to explore their perceptions and reactions. Through this analysis, we aim to address the following research questions: 1) What are the primary differences in decision-making between the ADS and human drivers under extreme conditions? 2) To what extent do riders in a high-level autonomous vehicle trust the autonomous driving system's operations under extreme conditions? 3) What are the influencing factors that induce riders to intervene in the autonomous driving under extreme conditions?

2 METHODOLOGY

2.1 Scenario reproducing

We examined three critical scenarios involving crashes between road users under varying environmental conditions: I. A collision between two passenger cars traveling in the same direction on a freeway under fine weather condition, caused by sudden lane-changing behavior. II. A collision involving a van and a Type 3 ambulance at a signalized intersection during nighttime, triggered by urgent passing through. III. A collision at a signalized intersection between a passenger car and a motorcyclist under fog condition, also due to urgent passing through.

To accurately recreate the traffic environment, we used a co-simulation strategy combining VISSIM with CARLA, enhanced by a VR headset and driving simulator. Each 30-second, case-specific scenario modelled background traffic with VISSIM and replaced crash-involved vehicles with those controlled by finely-tuned ADS algorithms, varying from one to all cars. These ADS were trained using the Deep Deterministic Policy Gradient (DDPG) algorithm in CARLA (Xu *et al.*, 2024b). We manually recreated the 3D environments based on crash videos, extending the road network as needed for ADS training purposes. Fig. 1 provides illustrative examples of the key elements involved in this study.

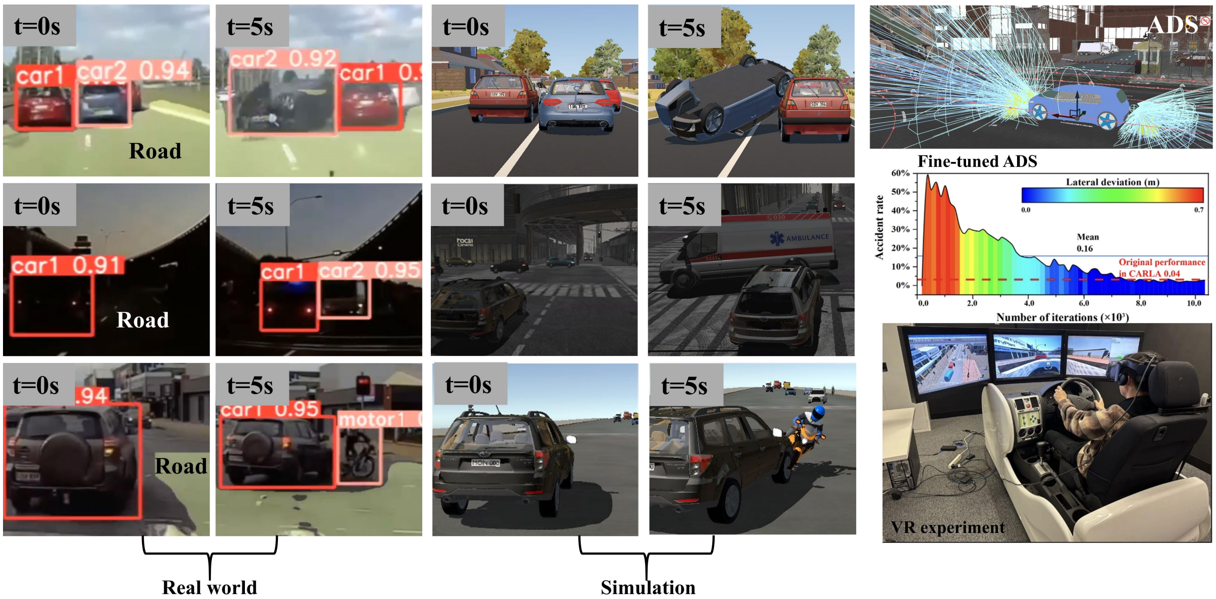

Figure 1 – *Key elements involved in the proposed research.*

2.2 Human-in-the-loop analysis

A total of 60 participants, comprising 38 males and 22 females aged between 17 and 48 years (mean age = 28.44, SD = 6.02), were recruited for the VR experiment. The participants were divided into three groups of 20, with each group having a relatively homogeneous distribution based on demographic characteristics such as gender, age, driving experience (incl. non-drivers), and VR experience. The experiment was conducted using the SIMREX CDS driving simulator, a simulated cabin outfitted with a steering wheel, brake and accelerator pedals, a dashboard, and a vibration system under the seat. The HTC VIVE Pro headset was used to provide an immersive VR experience (see bottom right of Fig. 1).

Participants were briefed on a 30-second autonomous driving session where they could intervene only if feeling unsafe, either by stepping on the brake pedal or turning the steering wheel. Each participant randomly experienced scenarios in crash-involved cars, and each group experienced in different orders. Data collection during the experiment utilized both objective (i.e., participants' reactions) and subjective (i.e., questionnaire surveys) methods. Specifically, intervention behaviors are recorded by the simulator, and participants' levels of simulator sickness are measured using a self-reported Simulator Sickness Questionnaire (SSQ) (Kennedy *et al.*, 1993), where they could report 16 symptoms ranging from none to severe. Additionally, at the end of the experiment, participants completed a simulator acceptance questionnaire (Van Der Laan *et al.*, 1997) coupled with a self-awareness survey driven by a talk-aloud protocol (Afflerbach, 2022) to capture their direct attitudes toward the experiment.

3 Results

3.1 Comparison of decision-making between human drivers and the ADS

Compared to the performance of human-driven vehicles involved in collisions, the ADS can achieve collision-free outcomes following the same operational trajectories, regardless of environmental factors such as road conditions, lighting, traffic flow, interactions with other road users, traffic control measures, etc. It was also found that there is a positive correlation between the number of AVs and reduced accident rates (Fig. 2). For instance, replacing one human-driven vehicle with an AV equipped with an autonomous driving system (ADS) could reduce the accident rate from 100% to 43%, 47%, and 59%, respectively, in three selected scenarios. When all vehicles are driven by ADS, there are no accidents.

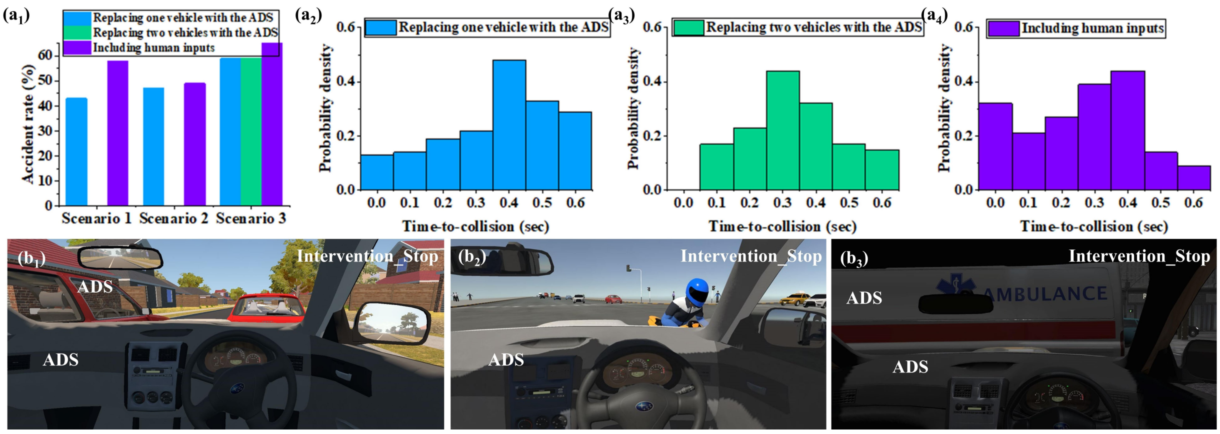

Figure 2 – *(a) Results with respect to the impact of replacing human-driven cars to the ADS on accident prevention; (b) Typical intervention cases from the first-person perspectives.*

The mechanism behind the avoidance of collisions is primarily due to the superior control of the ADS over longitudinal acceleration, which grants the AV greater flexibility in mitigating conflicts with interacting road users. This becomes particularly evident when just one human-driven vehicle is replaced by an ADS-equipped vehicle. Furthermore, when all human-driven vehicles are replaced by AVs, there are no fluctuations in acceleration observed, even though these scenarios are associated with relatively small time-to-collision (TTC) values.

3.2 Participants' reactions to extreme conditions during autonomous driving

The experimental results indicate that riders' intervention behaviors could disrupt idealized scenarios, leading paradoxically to collisions despite participants' intentions to avoid crashes. For example, in Scenario 1, a typical intervention-induced collision occurred when a participant turned the steering wheel to evade an impending crash but inadvertently collided with a vehicle on the adjacent lane. Subsequently, the inertia caused the participant's vehicle to collide again with the original interacting vehicle.

This phenomenon can be explained by considering the vehicles' relative positions alongside participants' perceptions. It was observed that participants with 1-10 years of driving experience were more likely to preemptively intervene in the ADS's decisions, whereas those with more extensive driving experience were less inclined to intervene at the same points, showing greater reliance on the ADS. Most notably, non-drivers did not execute any interventions throughout the experiment, demonstrating complete trust in the ADS across all scenarios.

All participants involved in this study successfully completed the tasks. They expressed appreciation for the simulation approach, and only 7% of participants reported experiencing mild-level sickness. This proofs that their performance was not negatively affected due to simulation factors.

4 Discussion

The significance of this study lies in its potential to bridge the gap between the technical advancements in ADS and user beahvior in AVs. It demonstrates further that riders' intervention behavior, often triggered by their perception of risk even when the ADS assesses the situation as safe, can hinder the seamless operation of the ADS and potentially lead to collisions. Specifically, we found that despite its stability, the consistent and homogeneous pace of movement maintained by the ADS is a critical factor that contributes to riders feeling unsafe. This perception arises because it seems as though the AV did not realize the danger and not react or stop.

Looking ahead, this study opens up several avenues for future research as pointed by prior works (Van Brummelen *et al.*, 2018). It would be valuable to explore the psychological and cognitive factors underlying riders' intervention behavior in more depth. This could involve investigating the role of individual differences, such as risk perception, trust in technology, and prior experiences with AVs, in shaping intervention decisions. Also, future studies could examine the effectiveness of various design strategies in enhancing riders' trust of ADS under extreme conditions. This could include exploring the impact of different human-machine interface designs, feedback mechanisms, and communication strategies on riders' perceptions and behaviors.

References

Afflerbach, Peter. 2022. *Teaching readers (not reading): Moving beyond skills and strategies to reader-focused instruction*. Guilford Publications.

CarCrashMedia. 2021. *Australian Car Crashes*.

de Melo, Celso M, Marsella, Stacy, & Gratch, Jonathan. 2019. Human cooperation when acting through autonomous machines. *Proceedings of the National Academy of Sciences*, **116**(9), 3482–3487.

Feng, Shuo, Sun, Haowei, Yan, Xintao, Zhu, Haojie, Zou, Zhengxia, Shen, Shengyin, & Liu, Henry X. 2023. Dense reinforcement learning for safety validation of autonomous vehicles. *Nature*, **615**(7953), 620–627.

Kennedy, Robert S, Lane, Norman E, Berbaum, Kevin S, & Lilienthal, Michael G. 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, **3**(3), 203–220.

Nair, Gopindra S, & Bhat, Chandra R. 2021. Sharing the road with autonomous vehicles: Perceived safety and regulatory preferences. *Transportation research part C: emerging technologies*, **122**, 102885.

Van Brummelen, Jessica, O'brien, Marie, Gruyer, Dominique, & Najjaran, Homayoun. 2018. Autonomous vehicle perception: The technology of today and tomorrow. *Transportation research part C: emerging technologies*, **89**, 384–406.

Van Der Laan, Jinke D, Heino, Adriaan, & De Waard, Dick. 1997. A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C: Emerging Technologies*, **5**(1), 1–10.

Xu, Zheng, Zheng, Nan, Lv, Yisheng, Fang, Yihai, & Vu, Hai L. 2024a. Analyzing scenario criticality and rider's intervention behavior during high-level autonomous driving: A VR-enabled approach and empirical insights. *Transportation research part C: emerging technologies*, **158**, 104451.

Xu, Zheng, Wang, Changshuai, Jiang, Tanghan, & Zheng, Nan. 2024b. Impediments to Environmental Awareness in Autonomous Driving Systems and Its Effect on User Adoption. *IEEE Transactions on Intelligent Vehicles*.