Proceedings of the 2021 IISE Annual Conference
A. Ghate, K. Krishnaiyer, K. Paynabar, eds.

Investigating Traffic Crashes Involving Autonomous Vehicles

Jesus Torres, Yueqing Li and Jing Zhang
Lamar University
Beaumont, TX, USA

Abstract

Deaths due to road traffic accidents are one of the leading causes of death in the United States. Furthermore, the economic impact of road traffic accidents accounts for about 3% of the United States' annual gross domestic product (GDP). In the past decade, extensive research has focused on autonomous vehicles (AVs). This technology is said to help prevent traffic accidents while promoting road traffic safety. This study aims to investigate the safety performance of AVs and identify the significant risk factors associated with the AV collisions. The study considers more than 200 crashes involving AVs and includes vehicle factors, environmental factors, collision type and crash severity. Multinomial logistic regression was conducted with collision type. The results showed no statistically significant risk factors to crash severity. However, movement preceding to collision contributes significantly to collision type. When both vehicles are moving, there's a higher likelihood of an angled collision, 47% to be exact. The other scenario which demonstrates a high probability of an angled collision is when the AV vehicle is not moving while the other is moving. The highest probability for a rear-end collision to occur is when the AV vehicle is not moving while the other is moving. This scenario makes up 55% of the entire rear-end collisions. As for the second-highest proportion, both vehicles moving, it consists of 42%. The research shall help reduce AV involved collisions and increase driving safety.

Keywords
autonomous vehicles, logistic regression, risk factors, crash severity

1. Introduction
Deaths due to road traffic accidents are one of the leading causes of death among the age group 1 to 54 in the United States [1]. In 2018, there were 36,560 people killed in motor vehicle traffic crashes on U.S. roadways [2]. Accordingly, every 12 minutes, someone dies in a motor vehicle crash on U.S. roadways [3]. Furthermore, the economic impact of road traffic accidents accounts for about 3% of the United States' annual gross domestic product (GDP) [4]. In retrospect, $616.2 billion lost to road traffic accidents of the $20.54 trillion U.S. GDP for 2018 [5]. Given the severity of the situation in the United States' road traffic safety, there is an urgent need for assistive technology in vehicles.

In the past decade, extensive research has focused on autonomous vehicles (AVs). This technology is said to help prevent traffic accidents while promoting road traffic safety. To better understand how AVs perform, the California Department of Motor Vehicles (DMV) adopted the Autonomous Vehicle Tester (AVT) program in 2014. Regulations of the program stipulate that manufactures are required to report any collision within ten days of the incident. AV collision reports are to be filled by the testing coordinator regardless of the severity or party at fault. In recently conducted studies, the goal has been to investigate the safety performance of AV technology. All the papers, including [6, 7, 8, 9, 10, 11] focused on identifying the risk factors related to AV collision reports. However, the limitations found in those studies prompted the authors of this study to take action.

Given the critical role of AV collision reports, the authors initiated a study to analyze data filed by the CA DMV from January 2018 to August 2020 (*subject to change*). As of August 18, 2020, the DMV has received a total of 265 AV collision reports [12]. However, given that the DMV didn't require additional accident details until January 2018, this study will only focus on the period starting from thereon. Previous studies have identified potential risk factors associated with AV collisions. The issue is that many of these studies didn't have much data available. Therefore, many of their sample sizes were too small. As a consequence, the results of these studies may not be similar to one another, subjecting them to interpretation. This study aims to investigate the safety performance of AVs and identify the significant risk factors associated with the AV collision reports.

2. Literature review

2.1. General overview

Recent studies have focused on identifying the risk factors associated with AV collisions. Most of these studies conducted their analyses using small sample sizes. Consequently, their results may not align with one another, allowing them to be open to interpretation.

2.2. Driving mode

Vehicles in autonomous mode at the time of collision have been reported to have a greater effect on crash severity. Most studies reported that the majority of the AV collisions happen to be in autonomous mode. [10] and [11] concluded that vehicles in autonomous mode are one of the main factors contributing to crash severity in AV collisions. However, [11] also found that the AV driving mode contributes to the collision type. Their results indicate a positive correlation between autonomous driving mode and rear-end collisions.

2.3. Weather

Most researchers suggest that weather conditions have no statistically significant impact on crash severity related to AV collisions. [11] found that more than 90% of reported AV collisions occurred in clear or cloudy weather conditions. Furthermore, [8] concluded that there is no correlation between the severity of damages and weather. It's important to note that California is a top-ranked state in having the sunniest climate nationwide. Given the state's sunny climate it's probable that AVs may demonstrate differing collision statistics if field testing is conducted in different environments. In fact, [6] made the revelation that when a vehicle is in autonomous mode it demonstrates a high likelihood of adverse weather collisions. Data acquisition from different testing environments is necessary to fully understand the role weather conditions play in AV related accidents.

2.4. Lighting

Only one paper, [6], evaluated the effect lighting conditions have on AV collisions. In using a Bayesian latent class model, the authors found that classes associated with dark lighting conditions demonstrate a higher crash severity level. Although no other study calculated the effect lighting conditions have on AV collisions, [6] address a risk factor that may require further consideration.

2.5. Roadway factors

2.5.1 Roadway surface

AV's performance over different roadway surfaces was measured by [8]. The study developed multilayer perceptron models with discrete testing and training partitions. All the models found that changes in the roadway surface are the most challenging situation for AVs. No other study mentioned the roadway surface in their results.

2.5.2 Roadway conditions

The effect of roadway conditions on AV crash severity was also analyzed by [8]. They used a Pearson Chi-Square test to evaluate the association between roadway conditions and AV crash severity. Results demonstrated no statistical significance indicating any connection between roadway conditions and AV crash severity. [8] further stated that roadway conditions are the least important risk factor since they don't affect AVs driving in autonomous mode.

2.5.3 Type of road

AV collisions occur across all types of roads. However, the role that each particular road type has on crash severity is different. [10] identified highways as the most recurrent road type correlated to high crash severity. On that note, [7] found that only 20% of all reported AV collisions occurred in highways and expressways. Suburban roads are the most common road type where 48% of AV collisions happened, followed by 32% in city roads. One-way roads are another main factor in which [11] determined to be contributing to the crash severity level related to AV collisions.

2.6. Type of collision

Across all studies conducted the variable, type of collision is closely analyzed. Moreover, all the papers including [6, 7, 8, 9, 10, 11] concluded that the predominant type of collision is rear-end collisions followed by side-swipe collisions. [7] and [11] described the common vehicle movements which preceded the AV involved crashes as the autonomous vehicle being at a full stop while the conventional vehicle behind proceeds forward. [11] also indicated that rear-end collisions' severity level is higher than that of any other type of collision.

Given the high severity level associated with rear-end collisions, [11] further investigated what factors were contributing to the variable, type of collision. Consequently, they found both AV driving mode and AV movement preceding collision to be significant factors affecting the type of collision. Other studies also investigated the potential factors contributing to the type of collision. [9] attribute the high number of rear-end AV collisions to be the second party's fault, that is the CV drivers. In comparing the drivers' maneuvers preceding an accident with only CVs to that of accidents with AVs, the study concluded that in rear-end AV involved collisions the second party, CV drivers, were at fault because of driving maneuvers such as "driving too closely" and "unsafe speed". On the other hand, [10] identified that AVs are at fault and greater risk of rear-end collisions more often when operated in conventional mode. That is to say, when AVs operate under conventional driving mode there is a higher likelihood of rear-end collisions where the AV is at fault. Human error as mentioned by [9] could be a contributing factor to the observations made by [10]. The variable, type of collision, is a significant risk factor identified across all studies, therefore, providing the most insight into the AV collisions involved.

2.7. Limitations of previous research

One limitation seen across the published studies involving AV collisions is their small sample size. A crucial disadvantage of small sample sizes is their variability. Variability increases with a smaller sample size resulting in less accurate results. Take, for example, [9] only used a total of 53 reported AV collisions ranging from January 2015 to December 2017. At that time, the California DMV had not revised their AV collision reports. Starting in January 2018, all AV collision reports required additional accident details apart from the written descriptions.

Therefore, another limitation identified across some studies is their use of outdated AV collision reports, which only include written descriptions. These reports provided no supporting accident details allowing written descriptions to be open for interpretation. In other words, bias may very well have guided the formation of those written descriptions. A potential drawback to the studies which used outdated AV collision reports is the retrieval of inaccurate data. For example, [9], apart from using a small sample size, acquired partisan data from the written descriptions. The credibility of some studies will be subject to scrutiny. However, [6], the latest article with the largest sample size, shares similar results to [9]. Both studies identified rear-end collisions as the leading collision type. Not all of them used the same independent variables. [9] only used one of them, type of collision, while [6] used five other independent variables. The variable, type of collision, is seen as part of all studies' analyses, which demonstrated to be of statistical significance.

2.8. Research question

The goal of this study is to investigate the safety performance of AVs and identify their risk factors. This study's process is to work more efficiently because it'll avoid the limitations found in previous research. Therefore, it'll include the largest sample size to date, updated AV collision reports, and accurate statistical analyses.

3. Materials and methods

3.1. Data

The CA DMV continues to be the only one that provides public access to its AV collision database. Back in 2014, they adopted the Autonomous Vehicle Tester (AVT) program to assess safety performances. Regulations of the program require reports of all AV collisions within ten days of the incident regardless of severity and fault. This study will analyze all reports from January 2018 to October 2020 because of the additional accident details provided. Therefore, a sample size of 209 AV collision reports is a part of this study. The collected information acquired from the AV collision reports includes. To date, this is by far the most completed AV collision reports dataset.

3.2. Risk factors

3.2.1 General overview

All of the following risk factors are directly from the AV collision reports. Each variable represents the data appropriately. A total of eight variables constructed the database used to formulate the statistical data analyses. The following text provides descriptions of each risk factor utilized and their assigned levels, if any.

3.2.2 Vehicle factors

Two major vehicle factors include driving mode and type of collision. AV's driving mode at the time of collision could be a potential risk factor. For that reason, this variable classifies as either being in autonomous or conventional driving mode. Furthermore, the type of collision recorded by the AV collision reports is another potential risk factor. Collision type classifies as other/unknown, head-on, sideswipe, rear-end, broadside, hit object, and vehicle/pedestrian. Also, injury level refers to there being any recorded injuries specified in the AV collision reports. It classifies as either there being no injuries or recordable injuries.

3.2.3 Environmental factors

There are five environmental factors: weather, lighting, roadway surface, day of the week, and traffic status. Weather classifies as unknown, clear, clouding, raining, or fog. Next, lighting conditions classify as daylight, dusk/dawn, or dark/streetlights. Finally, roadway surface classifies as unknown, dry, or wet. As for the day of the week, this variable consists of two classifications. Its first classification is the collision occurred on a weekday or second during the weekend. Traffic status is another risk factor used to classify AV collisions. This study identified both the morning and evening rush hours for San Francisco county. Morning rush hour is from 07:00 - 09:00 AM, evening rush hour from 05:00 - 07:00 PM, and all other times classify as normal conditions.

3.3. Methodology

This research used multinomial logistic regression to investigate the risk factors impacting the different collision types. The continuous dependent variable, collision type, consists of 4 levels, including 0 - Rear end, 1 - Sideswipe, 2 - Angled, and 3 - Others. In conducting the regression for this study, the base level used was 3 - Others. There were seven independent variables analyzed. They include traffic status, day of the week, driving mode, weather, lighting, roadway surface, and the movement preceding a collision.

4. Results

4.1. Collision Type Logistic regression

To identify the predictors statistically correlated to the dependent variable collision type requires a multinomial logistic regression. The results indicate only one predictor to be of statistical significance. That predictor being the movement preceding the collision. Results show the following p-values; angled collisions = 0.0077, sideswipe collisions = 0.0575, and rear-ending collisions = 0.0005. Accordingly, the two statistically significant collision types are angled and rear-ending collisions.

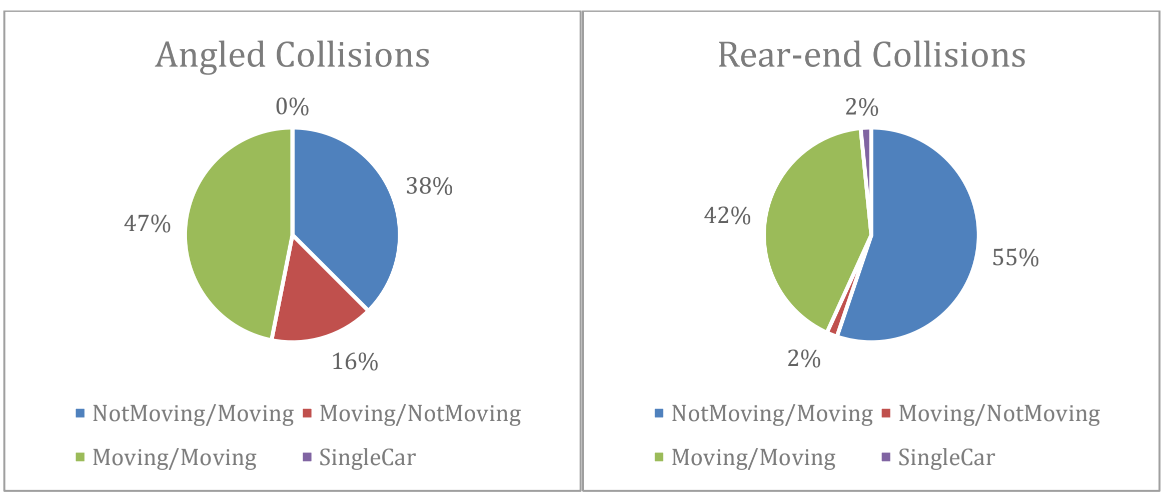

Figure 1: Angled Collisions (Left), Rear-end Collisions (Right)

Fig. 1 illustrates the proportions of the different movements preceding both angled and rear-end collisions. Angled collisions with a p-value of 0.0077 are the second most statistically significant. As shown in Figure 1, when both vehicles are moving, there's a higher likelihood of an angled collision, 47% to be exact. The other scenario which demonstrates a high probability of an angled collision is when the AV vehicle is not moving while the other is moving. Rear-end collisions with a p-value of 0.0005 are the most statistically significant. Figure 1 also demonstrates the different movements preceding rear-end collisions. The highest probability for a rear-end collision to occur is when the AV vehicle is not moving while the other is moving. This scenario makes up 55% of the entire rear-end collisions. As for the second-highest proportion, both vehicles moving, it consists of 42%.

5. Discussion

In conducting the multinomial logistic regression, the dependent variable collision type statistically correlates to one predictor. The predictor "movement preceding collision" demonstrates the following p-values for its three different levels: angled collisions = 0.0077, sideswipe collisions = 0.0575, and rear-ending collisions = 0.0005. In short, rear-ending and angled collisions are the only two levels most affected by the movement preceding a collision.

The results found in this paper differ from others in that no other study analyzed collision type as the dependent variable. Therefore, in most other studies, their dependent variable was crash severity or none. For instance, [9] didn't use a dependent variable. The authors decided to measure the differences in the distribution of types of collisions. Although their approach did provide some results, it's not the same as following a more analytical method. With that said, [8] did follow through in using a dependent variable. [8] found no correlation with the predictor collision type in using crash severity as their dependent variable. Instead, they found a significant correlation between crash severity and roadway surface. [8] reported a p-value of 0.01 in their results. Even then, the study didn't explicitly describe what factors they used to measure a collision's severity. In comparison, our research found a collision's preceding movement to be of statistical importance. Furthermore, our p-values are far more significant than that of [8]. Still, the thought behind changes in roadway surfaces having a positive correlation with crash severity provides some insight. A valid interpretation of this correlation is that autonomous vehicles' performance changes when on different roadway surfaces. AVs may be more vulnerable on wet surfaces than dry ones [8].

6. Conclusion

This paper conducted a multinomial logistic regression to identify the risk factors impacting different collision types. It was all based on the data published by the CA DMV. Our research indicates that the movement preceding a collision correlates with what type of crash is to occur. While other studies [6, 8, 9] only identified rear-end collisions as the most common, our research pinpoints the most significant risk factors. For instance, rear-end collisions are more likely to occur when the AV is not moving, and the other vehicle involved is moving. [7] is the only study which described the possible scenarios affecting rear-end collisions. Whereas, in our research, the movement preceding a collision showed a p-value of 0.0005 for rear-ending impacts. Another factor identified is that for angled collisions with a p-value of 0.0077, there's a higher likelihood of impact when the movement preceding the collision involves both vehicles moving. With these results at hand, AV manufacturers should continue to develop their safety features. Providing AVs with a 360-degree camera system could help minimize the number of rear-end collisions by allowing them to detect their surroundings safely when applying automatic emergency braking systems. As a result, AVs become more aware of their surroundings allowing them to better calculate their braking maneuvers.

Acknowledgments

This study was partially supported by the Natural Science Foundation (1726500). The findings and conclusions of this paper are those of the authors and do not necessarily represent the official position of NSF.

References

[1] "Road Traffic Injuries & Deaths: A Global Problem," 18 December 2019. [Online]. Available: https://www.cdc.gov/injury/features/global-road-safety/index.html.
[2] "2018 Fatal Motor Vehicle Crashes: Overview," National Highway Traffic Safety Administration, 2019.
[3] "Occupational Safety and Health Administration," 2005. [Online]. Available: https://www.osha.gov/Publications/motor_vehicle_guide.html.
[4] "Road Traffic Injuries," 7 February 2020. [Online]. Available: https://www.who.int/en/news-room/fact-sheets/detail/road-traffic-injuries.
[5] "GDP - United States," 2019. [Online]. Available: https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?locations=US.
[6] S. Das, A. Dutta and I. Tsapakis, "Automated Vehicle Collisions in California: Applying Bayesian Latent Class Model," IATSS Research, 2020.
[7] F. M. Favaro, N. Nader, S. O. Eurich, M. Tripp and N. Varadaraju, "Examining Accident Reports Involving Autonomous Vehicles in California," PLoS ONE, vol. 12, no. 9, pp. 1-20, 2017.

[8] S. H. Leilabadi and S. Schmidt, "In-Depth Analysis of Autonomous Vehicle Collisions in California," in 2019 IEEE Intelligent Transportation System Conference, 2019.

[9] D. Petrovic, R. Mijailovic and D. Pesic, "Traffic Accidents with Autonomous Vehicles: Type of Collisions, Maneuvers and Errors of Conventional Vehicles' Drivers," Transportation Research Procedia, vol. 45, pp. 161-168, 2020.

[10] S. Wang and Z. Li, "Exploring the Mechanism of Crashes with Automated Vehicles Using Statistical Modeling Approaches," PLoS ONE, vol. 14, no. 3, pp. 1-16, 2019.

[11] C. Xu, Z. Ding, C. Wang and Z. Li, "Statistical analysis of the patterns and characteristics of connected and autonomous vehicle involved crashes," Journal of Safety Research, vol. 71, pp. 41-47, 2019.

[12] "Autonomous Vehicle Collision Reports," 19 August 2020. [Online]. Available: https://www.dmv.ca.gov/portal/vehicle-industry-services/autonomous-vehicles/autonomous-vehicle-collision-reports/.