Road Traffic Safety Assessment in Self-Driving Vehicles Based on Time-to-Collision with Motion Orientation

Fernando M. Ortiz[a], Matteo Sammarco[b], Marcin Detyniecki[b], Luís Henrique M. K. Costa[a,*]

[a] *GTA/PEE/COPPE – Universidade Federal do Rio de Janeiro – Rio de Janeiro, Brazil*
[b] *AXA Group – Paris, France*

Abstract

Traffic conflict analysis based on Surrogate Safety Measures (SSMs) helps to estimate the risk level of an ego-vehicle interacting with other road users. Nonetheless, risk assessment for autonomous vehicles (AVs) is still incipient, given that most of the AVs are currently prototypes and current SSMs do not directly apply to autonomous driving styles. Therefore, to assess and quantify the potential risk arising from AV interactions with other road users, this study introduces the $\mathrm{TTC}_{\mathrm{mo}}$ (Time-to-Collision with motion orientation), a metric that considers the yaw angle of conflicting objects. In fact, the yaw angle represents the orientation of the other road users and objects detected by the AV sensors, enabling a better identification of potential risk events from changes in the motion orientation and position through the geometric analysis of the boundaries for each detected object. Using the 3D object detection data annotations available from the publicly available AV datasets nuScenes and Lyft5 and the $\mathrm{TTC}_{\mathrm{mo}}$ metric,

*Corresponding author
Email addresses: fmolano@gta.ufrj.br (Fernando M. Ortiz), matteo.sammarco@axa.com (Matteo Sammarco), marcin.detyniecki@axa.com (Marcin Detyniecki), luish@gta.ufrj.br (Luís Henrique M. K. Costa)

we find that at least 8% of the interactions with objects detected around the AV present some risk level. This is meaningful, since it is possible to reduce the proportion of data analyzed by up to 60% when replacing regular TTC by our improved TTC computation.

Keywords: Surrogate Safety Measures, Autonomous Vehicles, Smart Mobility, Road Safety.

1. Introduction

The smart mobility revolution with the introduction of autonomous vehicles (AVs) does not only impact car manufacturing industry only, but also linked businesses like insurance. In fact, the way the vehicle is driven no longer depends on the human driver behavior, but on the Artificial Intelligence (AI) system controlling the vehicle and relying on a multitude of sensors. This new approach is not infallible and there are already reported accidents with vehicles with some level of autonomy (Betz et al., 2019) which raises the liability issue. As the traditional risk assessment does not apply anymore, it becomes important to investigate new metrics that can model the behavior of the AV to, ultimately, help to define the insurance premium.

Vehicle manufacturers are equipping their vehicles with driver assistance and support systems, which already constitute partial automation systems. The American SAE (Society of Automotive Engineering) has defined a classification of vehicle autonomy levels, SAE J3016 (SAE, 2018). Numbered from 0 to 5, higher autonomy levels mean greater AI complexity, as well as intensive use of sensors in the vehicle. Different autonomy levels represent different combinations between driver involvement and the complexity or maturity of the AI system controlling the vehicle. Broadly speaking, AVs have control systems to detect and respond to events in the presence of objects around them. Nevertheless, there are limitations related to the type of situations and how they are managed by the AI system. Such limitations are intrinsic to the randomness of the road infrastructure and object around, as well as the weather and lighting conditions inherent to the environment in which the vehicles operate. Therefore, it is necessary to analyze operational factors of the vehicle that allow identifying risk events for the vehicle itself, passengers, pedestrians and other road users.

Traffic risk events are often evaluated through Surrogate Safety Measures (SSMs) (Tarko et al., 2009). SSMs are not used to prevent or avoid accidents, but to assess and analyze the probability of risk events and their severity based on movement parameters of the ego-vehicle and vehicles around. Currently, safety analysis for regular vehicles (without autonomous functions) takes place through the acquisition of data from off-the-shelf (OTS) devices and vehicle's proprioceptive sensors (Ortiz et al., 2022). On the other hand, AVs contain a set of technologies that aim to improve the perception of the environment outside the vehicle, allowing a safety analyzes which includes road users interacting with the AV, perceived through a variety of exteroceptive sensors, such as cameras and LiDARs (Ortiz et al., 2022).

Currently, risk assessment in AVs is still incipient, given that most of these vehicles are still under development. Therefore, to investigate the potential risk arising from AV interactions with other road users, this study uses public AV dataset from nuScenes (Caesar et al., 2020) and Lyft5 (Kesten et al., 2019) to evaluate the Time-to-Collision (TTC) indicator that considers the yaw angle as an additional parameter for the calculation. Our goal is to identify potential risk events from changes in the motion orientation and position through the geometric analysis of the boundaries for each object detected by the AV. Data annotations from the 3D bounding boxes dimensions (weight w, length l, and height h) and coordinates x, y, z available in the datasets are used to determine the proximity with the AV. The calculation of the yaw orientation is based on the camera intrinsics, i.e., parameters that characterize the optical, geometric, and digital characteristics of the camera (using it as a coordinate system origin), and data from the rotation and translation which corresponds to the motion of objects observed by the vehicle's camera driving video.

With technological advances in terms of sensing and autonomy, we aim to explore the potential of using data from AV prototypes to develop strategies for traffic risk events assessment. Thus, it is possible to monitor AVs through variables that enable policymakers to customize services for stakeholders. For this, data from real AVs in circulation on roads are used. Although AV dataset characteristics are limited in time and crash events (Wang et al., 2017), these allow to describe diverse patterns related to the vehicle's abilities to interact with different challenge events in a rapidly changing environment like the vehicular. The analyzed data were filtered and processed according to the proposed methodology.

In a nutshell, the contributions of our work are:

- We introduce the yaw angle in the TTC calculation of each object whose orientation/position converges to the AV on a collision course. We analyze diverse trajectories (following, head-on, and crossing scenarios) of objects converging to the AV.

- We evaluate the vehicle risk based strictly on the sensor variations and the evasive actions taken by the AV and, thus, provide the basis for an AV driving profile model.

- We reduce the data volume analyzed in risk assessment by considering the geometry of the boundaries used for object detection in the AI system controller. The goal is to discard all the detected objects that do not represent a real risk for the AV.

This paper is organized as follows. Section 2 reviews related works. Section 3 describes the data collection, preparation and analysis used to calculate motion properties and dynamics of both AV and detected objects. Section 4 shows the TTC calculation based on yaw orientation and motion properties of both AV and objects in collision course. Section 5 presents and discuss the results, and finally, Section 6 concludes the paper and presents future work.

2. Related work

Different safety indicators have been designed for risk assessment in traffic conflicts (Mahmud et al., 2017). Indeed, these indicators are characterized by the fact that they allow to quantify the severity of traffic risk events. Additionally, it is possible to estimate the level of risk in scenarios where historical crash data is unavailable. This work focuses on SSMs as a technique to assess risk. SSM use in this work is briefly discussed in the following.

2.1. Surrogate safety measures

SSMs are defined as measurements that are used to describe the relationship between road users pairs in a traffic risk event to quantify the crash probability or the potential traffic conflict severity in a meaningful way. Traffic conflicts analysis can be based on evasive actions or temporal/spatial proximity (Zheng et al., 2014). In particular, we aim to describe traffic conflicts based on temporal and spatial proximity using the Time-to-Collision (TTC) metric, a safety estimation indicator based on distance and speed variations. Through the analysis of these factors, it is possible to estimate and argue the severity of risk events associated with the vehicle. However, traffic conflicts do not depend only on the vehicle operation, and therefore, the analysis of risk events is subject also to the nature of the decisions by the drivers in the presence of any traffic risk event. An example of this is the reaction time, actions to minimize accidents, the veracity of evasive actions, as well as the intensity of evasive actions.

2.2. Time-to-Collision (TTC)

TTC is defined as the time it would take for the ego-vehicle to collide with an object ahead, if the current relative speed was maintained from the previous advance along the same path (Hayward, 1972). This is a continuous measure of safety that can be calculated at any moment as long as the ego-vehicle and the object are in a conflict area, i.e., an instantaneous situation where two or more road users interact in a road segment, with some uncertain, non-zero probability of colliding in the near future. Thus, TTC enables the collision course analysis for vehicles and predicts how is the vehicle's motion related to other users of the road infrastructure. Moreover, TTC is the simplest and most effective analytical metric for collision risk assessment in according to their study (Tak et al., 2018).

Equation 1 defines the TTC as the relation of the distance between the ego-vehicle and objects ahead ($d_{(ego,obj)}$) and speed difference between both ego-vehicle (v_{ego}) and an object ahead (v_{obj}); for simplicity in this case we assume the object is another vehicle. Typically, the TTC value indicates the minimum time to collide, calculated continuously through the detection process of a potential traffic risk event. In the situation of imminent collision, TTC values assume finite decreasing values as the severity of the traffic risk event increases. It is worth noting that the TTC value allows inferring the amount of reaction time available for evasive maneuvers as a measurement of the risk level.

$$TTC = \begin{cases} \frac{d_{(obj,ego)}}{v_{ego}-v_{obj}}, & \text{if } v_{ego} > v_{obj} \\ \infty, & \text{otherwise} \end{cases} \tag{1}$$

Due to TTC limitations (it ignores evasive actions, speed restrictions of the ego-vehicle direction related to the object ahead), several modifications have been proposed to improve the accuracy of this metric.

2.2.1. *Modified Time-to-Collision (MTTC)*

Modified Time-to-Collision (MTTC) (Ozbay et al., 2008) uses acceleration as a parameter to analyze the vehicle trajectory and its conflict discrepancies due to acceleration/deceleration. However, MTTC depends on both the acceleration of the following vehicle and the leading vehicle, the latter being difficult to measure or obtain, from the ego-vehicle. Furthermore, MTTC by itself does not allow the severity of potential risk events to be quantified, since various combinations of distance/velocity/acceleration may produce similar MTTC values. For this, the authors propose a Crash Index (CI) that uses kinematic variation factors to estimate the severity of risk events (Ozbay et al., 2008).The authors conclude that CI can effectively model the temporal distribution of accidents to the same extent as MTTC.

2.2.2. *Enhanced Time-to-Collision (ETTC)*

Another TTC variation is the Enhanced Time-to-Collision (ETTC) (Kiefer et al., 2005). ETTC assumes that following and leading vehicles do not change their courses until a collision occurs. Moreover, deceleration in leading vehicle is considered until it stops. On the other hand, following vehicle's deceleration is considered to zero when the brake onset. Thus, ETTC calculation allows to define thresholds for "near" and "far" perception in Forward Collision Warning systems.

2.2.3. Time-to-Collision with Disturbance (TTCD)

Time-to-Collision with Disturbance (TTCD) analyzes collision risks product of disturbances in the leading vehicles (Xie et al., 2019). TTCD also can capture rear-end conflict risks in car-following scenarios where the leading vehicle may have higher speed. TTCD considers the deceleration product of the disturbance, and the critical deceleration rate imposed by the leading vehicle deceleration.

2.2.4. Time Exposed TTC (TET) and Time Integrated TTC (TIT)

On the other hand, to determine safety evaluations based on TTC in time intervals, other indicators have been proposed to describe micro-levels of safe and safety-critical events derived from the TTC value analysis. The Time Exposed Time to Collision (TET) is an indicator proposed in (Minderhoud and Bovy, 2001) which analyzes the time period that a vehicle remains exposed to high-risk events based on TTC values. These time periods analyze TTC measurements by thresholds defining the risk level. Thus, TET represents the duration of the exposition of safety-critical TTC values over a specified time duration. Thus, all of the instants in which the driver is following the leading vehicle, which $0 < TTC < TTC*$ must be summed. Nonetheless, this indicator takes into account a single threshold TTC* (i.e., safety/safety-critical events), and therefore, it does not consider the variation between lower TTC values. To reduce the impact of low TTC values do not affect the TET indicator, Minderhoud *et al.* (Minderhoud and Bovy, 2001) propose the Time Integrated Time to Collision (TIT) metric which integrates the TTC to define the safety level for each TET interval analyzed in each driver's profile. Thus, TTC values below

TTC* is also considered in the calculation process.

In addition to their improvements, TTC's variation metrics (MTTC, ETTC, TTCD, TET and TIT) also have disadvantages. These metrics are limited by the absence of motion analysis of the road users interacting with the ego-vehicle (e.g., evasive maneuvers, motion orientation, among others) when they are in a collision course. Table 1 shows a comparison of the approaches to improve TTC calculation.

Table 1: Summary of previous approaches using TTC.

Approach	Methodology	Advantages	Disadvantages
TTC (Hayward, 1972)	Calculation based on constant speed.	• Simple calculation based on distance and speed variations.	• Ignores motion characteristics of the ego-vehicle and road users.
MTTC (Ozbay et al., 2008)	Calculation uses acceleration in TTC general formulation.	• MTTC considers the acceleration of ego-vehicle and other vehicles during collision course.	• Ignores motion characteristics of the ego-vehicle and road users.
ETTC (Kiefer et al., 2005)	Calculation uses deceleration behavior of the objects ahead.	• ETTC considers characteristics of the objects ahead and their behavior when deceleration events occurs.	• Ignores motion characteristics of the ego-vehicle and road users.
TTCD (Xie et al., 2019)	Calculation considers the effects of disturbing events in vehicles ahead of the ego-vehicle.	• TTCD Analyzes reactions of the objects ahead that can affect the ego-vehicle.	• Ignores motion characteristics of the ego-vehicle and road users. • It is not clear how to apply the TTCD in diverse scenarios.
TET—TIT (Minderhoud and Bovy, 2001)	Calculation considers time duration and extension for the ego-vehicle drives in high-risky situations.	• Measures consider time intervals for safety analysis.	• Ignores variations occurred in TTC analysis.
TTC_{mo}	Calculation considers motion orientation of the objects ahead with respect to the ego-vehicle.	• This metric considers motion orientation on the ego-vehicle's motion axis. • TTC_{mo} considers just the objects ahead in collision course with the ego-vehicle.	• Depends on accuracy from semantic segmentation classification and bounding boxes processing in the ego-vehicle.

2.3. SSMs based on motion dynamics

Some studies analyze unrestricted road users' motion as part of the dynamics in vehicular environments. Miller *et al.* (Miller and Huang, 2002) develop a collision warning system that analyzes traffic risk events and evasive actions, sharing the location and kinematic measures from the ego-vehicle and the surrounding vehicles. The algorithm analyzes the time to collision and the time to avoidance in a parametric way. Laureshyn *et al.* (Laureshyn et al., 2010) propose a theoretical analysis of SSMs in collision course to determine the severity of traffic risk events. Given that interactions between road users are continuous, the authors suggest some strategies to calculate TTC for conflicts of different angles at constant speed. The authors stated that in potential collisions, a corner of one of the vehicles touches one side of the other vehicle. Thus, a new concept for TTC is developed, which calculates TTC between a moving line section of the ego-vehicle and a point in the other vehicle, in a time instant t. Next, the coordinates of the line section ending after t seconds based on a constant speed motion. Some assumptions about parallel motion are defined, depending on gradient of the line. Jiménez *et al.* (Jiménez et al., 2013) make an improved calculation of TTC in (Miller and Huang, 2002), assuming the vehicle geometry to be rectangular. In addition to the simplified calculation, the system analyzes the dimensions of the vehicles involved in the interaction, and the areas involved in a potential traffic conflict. In this way, Qu *et al.* (Qu et al., 2018) proposes a TTC method with motion orientation based on GPS coordinates to analyze cross-collision events. The authors use GPS data to calculate speed and distance, as well as the heading and the orientation angles of the target vehicle. The authors use a rectangle model to represent the shape of the target vehicles. The experiments are carried out in a simulated environment with two test vehicles. The results show that rectangular model enables the TTC calculation more accurately, and can also have superior performance when the angle between two vehicles is small, reducing false alarms. Ward *et al.* (Ward et al., 2015) analyze the interactions between vehicles to define a prediction system and avoidance of collisions in vehicle-to-vehicle (V2V) communication systems. The method analyzes TTC for vehicles without motion restrictions. The authors calculate TTC in 2D, based on the relative vehicle motion and a looming method (a technique for gating predictions based on the relative motion of the vehicles), which considers the relationship of the vehicle roll angle, linear and angular velocity, and the yaw rate vector. Wachenfeld *et al.* (Wachenfeld et al., 2016) propose a Worst-Time-To-Collision (WTTC) metric to identify risk events related to the mobility dynamics of objects. The authors do a physical analysis of vehicle motion using the Kamm's circle (a theory about the transferable forces from the tire to the road surface) and entering the yaw angle.

Differently from these studies, this paper analyzes the motion orientation of diverse road users that surround the ego-vehicle, detected through exteroceptive sensors, which enables the analysis not only with vehicles, but also with pedestrians and two-wheelers. Table 2 shows a comparison of the approaches involving motion orientation to improve TTC calculation.

2.4. SSMs based on data analysis

To analyze multiple interactions through SSMs, several works have developed studies based on software simulation. Papadoulis *et al.* (Papadoulis et al., 2019) and Virdi *et al.* (Virdi et al., 2019) performs a safety assessment for autonomous

Table 2: Summary of works considering motion orientation in the TTC calculation.

Approach	Methodology	Advantages	Disadvantages
Miller and Huang (2002)	The authors propose a collision warning system based on calculation of intersection points.	• The system includes an algorithm for intersection collision warning detection and considers communication strategies.	• Ignores motion characteristics of the ego-vehicle and road users.
Laureshyn et al. (2010)	Calculation of TTC based on convergence in different angles at constant speed.	• The framework enables to calculate collision probability based on TTC in sideswipe conflicts.	• Limited by disregarding motion characteristics of the ego-vehicle and road users.
Jiménez et al. (2013)	The authors make an improved calculation of TTC based on methodology proposed in Miller and Huang (2002).	• The framework considers vehicle geometry to be rectangular. • The tool considers also the dimensions of vehicles involved in the conflict.	• The framework is not tested on a real scenario.
Qu et al. (2018)	The authors propose a methodology to analyze cross-collision events based on GPS data.	• The system considers vehicle geometry to be rectangular. • The system uses GPS data to analyze orientation and heading angles of the target vehicle.	• The system is limited by the GPS precision and randomness of target vehicles.
Ward et al. (2015)	The authors propose an indicator that generalizes TTC to the planar case, mapping vehicle trajectories on the road to predict traffic conflicts.	• Planar analysis relies heavily on the relative positions of other traffic participants at the moment of predict the risk of a traffic conflict between vehicles. • The model considers uncertainties by communication (V2V).	• The model ignores other road users in the ego-vehicle vicinity.
Wachenfeld et al. (2016)	The authors propose a method to reduce the amount of data to estimate the criticality of a conflict.	• The method considers the motion orientation through yaw angles.	• WTTC can define uncritical events as potential risky, e.g., vehicles travel side by side. • WTTC does not consider other road users.
Our proposal (TTC_{mo})	Calculation considers motion orientation of the objects ahead with respect to the ego-vehicle.	• This metric considers motion orientation on the ego-vehicle's motion axis. • TTC_{mo} considers just the objects ahead in collision course with the ego-vehicle. • TTC_{mo} also discards other objects out the ego-vehicle's path.	• Depends on accuracy from semantic segmentation classification and bounding boxes processing in the ego-vehicle.

and connected vehicle fleets through the SSAM simulation tool (Surrogate Safety Assessment Model). The authors observed that most AV conflicts occur at intersections, and concluded that depending on the degree of AV penetration, the conflict rate may decrease. Zhang *et al.* (Zhang et al., 2020) evaluate the safety of connected autonomous vehicles by analyzing lane switching and exclusivity through the simulation tool PTV-VISSIM. On the other hand, Alghodhaifi and Lakshmanan (Alghodhaifi and Lakshmanan, 2020) analyze SSMs as a basis for a pedestrian protection system, through simulations in Matlab/Simulink.

Other works described in the literature analyze SSM metrics in data collection from naturalistic conduction studies. These testbeds use diverse exteroceptive sensors such as radars, cameras, GNSS, or V2X communication devices, to detect objects around the vehicle. Data sources, such as 100-Car (Dingus et al., 2006) and SHRP2 (Campbell and , U.S.) have been extensively studied via TTC to formulate safety metrics, analyze risk events, and compare simulated and real environments (Montgomery et al., 2014; Markkula et al., 2016). In the same way, Safety Pilot Model Deployment (SPMD) used around 3,000 human-driven vehicles, equipped with V2V communication devices and Mobileye sensing devices (Nodine et al., 2015). He *et al.* (He et al., 2018) evaluate SSMs from SPMD data. The authors implement three metrics: TTC, MTTC, and the Deceleration Rate to Avoid Collision (DRAC). The authors observed that the MTTC presented the best overall performance. Kusano *et al.* (Kusano et al., 2014) develop a methodology to identify situations where the ego-vehicle driver generates an evasive braking action. The authors use radar data and kinematic measures from the ego-vehicle (Dingus et al., 2006) to calculate the TTC as metric to activate warning actions. Five car-following scenarios are identified to implement the algorithm: scenarios where the leading vehicle or lack of leading vehicle lack is correctly identified by the algorithm; scenarios where the leading vehicle is detected but it is not in collision course with the following vehicle; and scenarios where the algorithm failed to identify the leading vehicle or detects other objects different of the visual analysis. The authors conclude that the algorithm can identify 91.8% of the braking events when verified visually.

On the other hand, analysis of SSMs in self-driving vehicles is limited. However, multiple AV developers have made available traces of their vehicles in the test phase (Ortiz et al., 2022). Beauchamp *et al.* (Beauchamp et al., 2022) make an analysis of safety measures considering collected video data captured by automated shuttles in two cities. The authors defined five possible interactions with other road users based on the collision angle and the parallelism angle: head-on, rear-end, side parallel and leaving. The safety indicators computed in this work were speed, acceleration, TTC and PET. The authors conclude that all the analyzed interactions were safe, due to the limited speed of the shuttle compared to other road users around it. De Ceunynck *et al.* (De Ceunynck et al., 2022) perform behavioral observations of two automated shuttles in Norway at intersections with various road users, such as pedestrians, scooters or bicycles. The authors conclude that more than 90% of the interactions with pedestrians are not dangerous, while there were some inconsistencies in the recognition of cyclists when turning. Alozi *et al.* (Alozi and Hussein, 2022) propose a framework based on Extreme Value Theory (EVT) to assess the safety of AV-pedestrian interactions by quantifying potential conflicts between them. The authors use data from nuScenes and Lyft5 AVs, and data from manual driven vehicles to test the accuracy of the EVT approach in relation to vehicle-pedestrian accident data. The goal is to reduce the pedestrian accident rate per million vehicle kilometers travelled. The analysis uses TTC and PET to evaluate the interactions. The authors estimate that the AV-pedestrian accident rate is between 4.041 and 5.499 per million vehicle kilometers traveled, which is a high value considering the safety of pedestrians.

Differently from these studies, this paper analyzes real data from AVs involving the orientation and position of the detected objects in order to accurately describe the motion of road users with the respect to the AV. Table 3 shows a comparison of the approaches using diverse data sources.

Table 3: Summary of works considering diverse data sources for safety analysis.

Approach	Methodology	Advantages	Disadvantages
Software simulation-based	Analysis of multiple scenarios with penetration rate capacity.	• It allows to implement flexible scenarios both in scalability and in time execution.	• It is not possible to simulate all the factors that are involved in driving, from human factors to unpredictable events such as fog.
Naturalistic data	Analysis of real-world data from the drivers (profiling) and the external environment.	• Naturalistic driving provides sensor setups to analyze both driving behavior and external environment related with the events that may occur.	• A limitation in self-driving vehicles is inferred by the AVs' penetration degree in the road infrastructure.
Beauchamp et al. (2022)	The authors make an analysis of safety measures considering video frames captured camera sensor in automated shuttles.	• The methodology considers collision and parallelism angles that enable the analysis considering how is the interaction with other road users.	• The authors do not consider the heading orientation in the calculation of safety metrics used to analyze the interactions between the shuttle and the road users. • Analysis depends on semantic data and categorization.
De Ceunynck et al. (2022)	The authors make an analysis of the impact of conflicts between AVs and pedestrians considering data collected by automated shuttles.	• The authors make a behavioral analysis of the automated shuttle interacting with other road users as pedestrians, bicycles or scooters.	• The authors do not consider the heading orientation in the calculation of safety metrics used to analyze the interactions between the shuttle and the road users. • Analysis depends on semantic data and categorization.
Alozi and Hussein (2022)	The authors propose a method to assess safety in the AV-pedestrian interactions.	• The method considers all the interactions with pedestrians.	• The authors do not consider the heading orientation in the analysis of AV-pedestrian interactions. • Analysis depends on semantic data and categorization.
Our proposal (TTC_{mo})	Calculation considers motion orientation of the objects ahead with respect to the ego-vehicle.	• This metric considers motion orientation on the ego-vehicle's motion axis. • TTC_{mo} considers just the objects ahead in collision course with the ego-vehicle. • TTC_{mo} also discards other objects out the ego-vehicle's path.	• Analysis depends on semantic data and categorization.

2.5. SSMs based on exteroceptive sensors

Studies on the evaluation of TTC through exteroceptive sensors have been developed to recognize the various entities with which a vehicle can interact. Aycard *et al.* (Aycard et al., 2011) propose a risk assessment system at intersections. The authors use data fusion from camera and LiDAR sensors to detect and establish the dynamics of detected objects. For risk quantification, the

TTC is used as a collision risk indicator. The authors conclude that risk assessment through environmental perception can enhance safety applications in the automotive industry. Kilicarslan and Zheng (Kilicarslan and Zheng, 2019) analyze vehicle collisions through TTC using video cameras. The authors analyze the divergence of horizontal and vertical movement in video frames without relying on bounding boxes. To this aim, TTC analysis is based on the size variations of the detected object in the video, divided by the size changes in time intervals. The analysis of the algorithm proposed by the authors is used in videos of naturalistic driving without accidents. Results show 94% accuracy and 93% precision in the relationship between the computed system and the actual video. Meanwhile, compared to the detection of the LiDAR sensor in the KITTI dataset (Geiger et al., 2012), the authors observe that LiDAR-based measurements depend on the depth of detection, discontinued detection, in addition to requiring 3D analysis. In this sense, video frame analysis is robust and can have a higher degree of accuracy.

The analysis of road safety metrics is closely related to the collection of image data from specific areas (mostly intersections), or video analysis in vehicles with embedded devices. Unlike these works, this study explores the potential of using data generated by AVs to develop road safety analysis solutions based on the vehicles' own sensing. Specifically, we focus on the TTC analysis with emphasis on the road users' motion orientation. Depending on the road users' orientation, TTC must be evaluated differently to accurately validate traffic conflicts involving the AV. This paper analyzes TTC based on the road users' orientation and position related to the AV. For that, nuScenes AV dataset (Caesar et al., 2020) and Lyft5 dataset (Kesten et al., 2019) are used in this study to analyze the motion orientation and position of the detected objects by the AV while it is moving. The goal is to analyze the TTC based on the yaw angle of the detected object and its position with respect to the AV through data analysis from exteroceptive sensors' data readings in AVs. To the best of our knowledge, this is the first analysis considering orientation for the TTC calculation based on data from AVs.

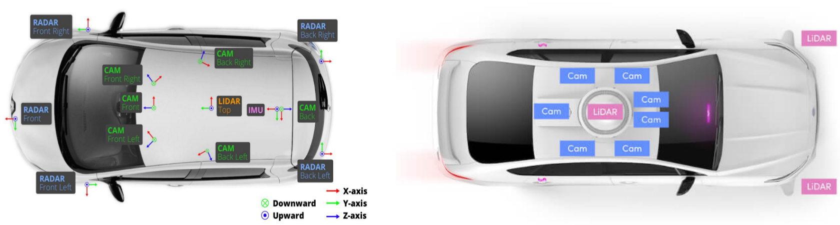
3. Methodology

Some experimental AV dataset are publicly available. In this work, we use two datasets including semantic data, nuScenes (Caesar et al., 2020) and Lyft5 (Kesten et al., 2019). As described in (Ortiz et al., 2022), these datasets have various characteristics that can be analyzed for braking and sudden acceleration analysis. In the following, we first describe how we extract data from the dataset and how TTC metrics are computed.

3.1. Dataset overview

nuScenes: nuScenes (Caesar et al., 2020) is a public large-scale dataset of autonomous driving traces which includes images from camera, point clouds (PC) from LiDAR, and radar signals detected by the sensors installed on the vehicle. This dataset also provides data from the vehicle internal sensors (e.g., acceleration or speed). In total, the dataset includes almost 6 hours of data gathered by two AVs, one in Boston (US), the other one in Singapore (SG). The internal sensing data is acquired from the CAN bus.

Lyft5: Lyft5 (Kesten et al., 2019) is another public large-scale dataset of AV traces, which contains images from cameras and LiDAR PCs. The perception dataset consists of 2.5 hours of data gathered by twelve vehicles in Palo Alto (PA) divided into 180 scenes of 25 seconds each. Unlike the nuScenes AV dataset, Lyft5 does not provide CAN bus data from the vehicle.



(a) nuScenes AV sensors.

(b) Lyft5 AV sensors.

Figure 1: Sensor setup for nuScenes (Caesar et al., 2020) and Lyft5 (Kesten et al., 2019).

Table 4 summarizes characteristics of nuScenes and Lyft5 datasets. We analyze the training data available for both datasets. As perception datasets, the raw data is processed by a perception system that uses sensory systems and software to perform multiple behavioral observations and interactions from different objects around the ego-vehicle, i.e., infrastructure and road users (Houston et al., 2020). Each detected object is described as an instance and it can have multiple interactions with the AV. Each instance is marked with a 3D bounding box, and categorization and attribute labels; each interaction of that instance with the AV is recorded in a log. Examples of categorization are vehicle type, two-wheelers, pedestrians, road infrastructure, among others, and attributes are vehicles or pedestrians stopped, in motion, among others. The nuScenes AV dataset contains day/night scenes with different weather conditions, 23 categories and 9 attributes for in-motion objects. On the other hand, Lyft5 contains less scenes, but the proportion of 3D bounding boxes annotations is similar to that of nuScenes AV dataset. Similarly to nuScenes, Lyft5 defines 9 categories and 18 attributes.

Table 4: Statistics of the two AV datasets.

	Scenes	**Vehicles**	**Images**	**LiDAR PCs**	**Radar PCs**	**Bounding Boxes**	**Day/ Night**	**Weather**	**Categories/ Attributes**
nuScenes	850	2	1.4 M	400 k	1.3 M	1.4 M	Yes	Yes	23/9
Lyft5	180	12	323 k	46 k	0	1.3 M	No	No	9/18

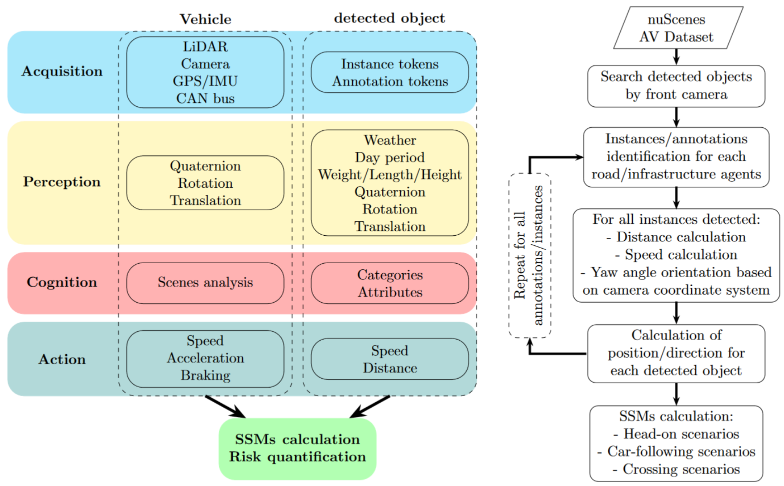
Both nuScenes and Lyft5 datasets include data from keyframes (i.e., synchronized samples among LiDAR, Radar and camera data, at 2 Hz and 5 Hz, respectively (Caesar et al., 2020; Kesten et al., 2019)), and data from each sensor sweeps, based on the sampling frequency of each one. Metadata of all samples are available in JSON files format. Moreover, the datasets provide training data, that is, data with sample annotations used to describe diverse characteristics of the object itself around the ego-vehicle, based on LiDAR PCs and JPEG images from the cameras.

3.2. Data preparation

Analysis on safety assessment requires to use data resulting from functional areas of AVs (IEEE Electronics Packaging Society, 2019). Thus, acquisition data (e.g., raw data from sensors like camera, GPS/IMU, among others), perception data (e.g., object detection, location, environment), cognition data (e.g., motion planning, maneuvers, among others) and action data (e.g., speed, brakes, steering, among others). Figure 2(a) shows the data used for both AV and detected objects through the exteroceptive sensors, with respect to functional areas for the AV performance analysis. Thus, AV and detected objects metadata are used to assess the safety of AV interactions with various road users and infrastructure. It is worth noting that it is possible to assess safety with respect to users other than vehicles, such as pedestrians and two-wheelers. Therefore, data from all AV autonomy phases are used to assess risk events for the categories of detected objects in the dataset, in order to establish a standard of AV driving with respect to the road users' motion. Although traffic accidents are unexpected and rare events that can be associated with multiple causing factors, this analysis can help to explain more clearly potential traffic accidents since any collision describes a convergence approach between the users involved in the collision, as described in Figure 2(b).

3.2.1. Motion orientation and position angle

Our goal is to identify the allocentric and egocentric spatial relations between detected objects and the ego-vehicle, defined by relative directions, distances and bearings (Meilinger and Vosgerau, 2010). Allocentric data is based on relations object-to-object. On the other hand, egocentric data is based on relations self-to-object (Meilinger and Vosgerau, 2010). Thus, our analysis uses the heading orientation (yaw angle) as an allocentric representation of the motion orientation (direction) of each detected object with respect to the ego-vehicle. Motion orientation enables to describe when an object is converging to or diverging



(a) Key variables used in the data analysis. (b) Data analysis procedure.

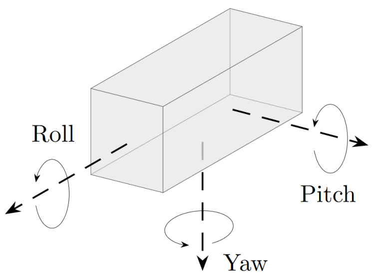
Figure 2: Summary of data used to analyze traffic risk events based on motion orientation.

of the ego-vehicle's course. In the meantime, the position angle is used as egocentric representation to describe the location of the objects with respect to the ego-vehicle. Position angle enables to infer the severity of a risk event conditioned by the position and the orientation of each detected object while interacting with the ego-vehicle. From the analysis of the dynamics of road users and the ego-vehicle, it is possible to evaluate metrics inherent to the objects' motion. For that, we use the nuScenes and Lyft5 devkits (nuTonomy, 2018; Lyft SDK, 2019), which provides a set of libraries to manipulate their datasets. We compute the bounding box orientation, swapping the sensor coordinate frame $[x,\ y,\ z]$ of the LiDAR ($[1,\ 0,\ 0]$) for the camera ($[0,\ -1,\ 0]$), according to the coordinate frames defined for each sensor. In this way, we set the yaw angles for each object detected based on the sensor coordinate frame of the frontal camera, as observed in Figure 1.

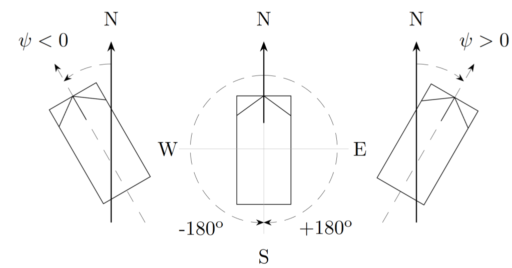
Furthermore, to describe the spatial orientation of the vehicle and the detected objects the yaw angle is used, as shown in Figure 3(a). Yaw angles (ψ) also indicate the orientation of each detected object. North is $0\,^{o}$ (ψ_0), east is $90\,^{o}$ (ψ_1), west is $-90\,^{o}$ (ψ_2) and south is $\pm 180\,^{o}$ (ψ_3). Objects have positive heading in clockwise direction and negative value in counterclockwise direction. About the ego-vehicle, we assume that yaw angle is $\psi_{ego} = 0\,^{o}$. Thus, detected objects with yaw angle between $\psi_2 < \psi_0 < \psi_1$ indicate that the direction on z-axis is forward the ego-vehicle; meanwhile, yaw angles between $\psi_1 < \psi_3 < \psi_2$ indicate that the direction is opposite to the ego-vehicle, as shown in Figure 3(b). On the other hand, position angles (θ) indicate the location of an object with respect to the ego-vehicle. Position angles $0\,^{o} < \theta < 90\,^{o}$ indicate object locations at right-side with respect to the driving direction, while $-90\,^{o} < \theta < 0\,^{o}$ at left-side, as shown in Figure 3(c). Bounding box centroid coordinates (x, z) are used to determine θ. Thus, it is possible to establish when the ego-vehicle path is converging with detected objects.

3.2.2. Geometric analysis of objects and ego-vehicle

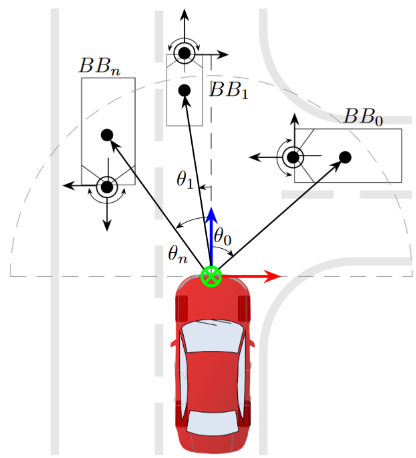
As shown in Figure 4, we use the ego-vehicle size specification to obtain a geometric representation and to analyze the interaction with surrounding objects. The width and length of the detected objects, available from the bounding boxes, are considered in the geometric analysis. In this sense, each vertex is labeled to determine its location and orientation when the object moves and rotates. The ego-vehicle is also represented as a bounding box. It is important to note that since $\psi_{ego} = 0$, the position of its vertices will always be the



(a) Roll, pitch and yaw rotations.



(b) Illustration of the yaw orientation in road users.



(c) Object position with respect to the ego-vehicle.

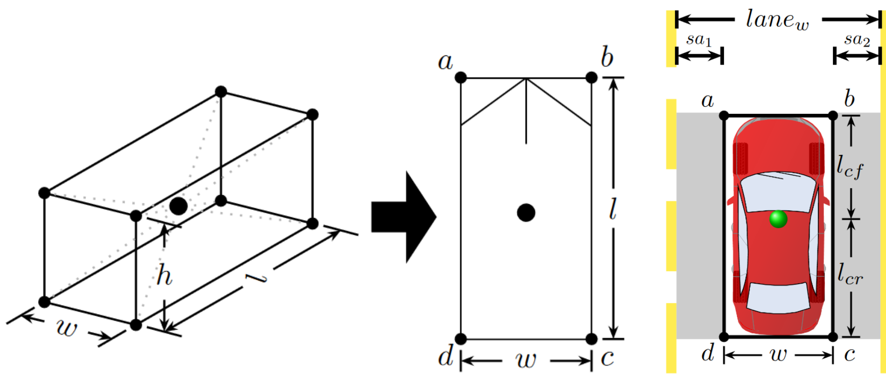
Figure 3: Relationship between AV and detected objects via motion orientation and position angle.

same for the analysis. In addition, the remaining space between the lane width and the ego-vehicle width is used as a safety area (sa_n, where n is an object identifier), to identify objects adjacent to the AV that may represent potential traffic conflicts. The lane width is based on the respective road city regulations. Vehicle size specifications are reported in Table 5.

To determine which objects are in collision course with the ego-vehicle, we aim to identify adjacent or overlapping trajectories between the ego-vehicle and other objects through motion orientation analysis. In this way, we identify be-

Table 5: nuScenes and Lyft5 vehicle overall dimensions. The width (w) includes external mirrors. The length between camera and vehicle front-side (l_{cf}) and the length between camera and vehicle rear-side (l_{cr}) are based on the camera location on the vehicle's rooftop.

Dimensions	**Vehicle**	
	nuScenes (Renault Zoe)	**Lyft5 (Ford Fusion)**
w [m]	1.945	2.121
l [m]	4.087	4.871
h [m]	1.562	1.478
l_{cf} [m]	1.810	2.302
l_{cr} [m]	2.277	2.569



(a) Bounding box geometry.

(b) Ego-vehicle geometry.

Figure 4: Geometric representation for an object (a) and the ego-vehicle (b).

havior indicators according to the AV reaction in several possible interactions with the objects around. Thus, we process AVs data to identify these interactions. In this analysis, we consider data annotations of the camera's coordinate system as reference. Information of the bounding box like the yaw rate (ψ), centroid position data in the image ($x,\ y,\ z$), and the size ($w,\ l,\ h$) are extracted from each annotation. Each vertex of a bounding box and the AVs are calculated by the relationship between sizes and the centroid coordinates, as described in Equation 2:

$$
\begin{aligned}
a_x, d_x &= x - \frac{w}{2}, \\
b_x, c_x &= x + \frac{w}{2}, \\
a_z, b_z &= z + \frac{l}{2}, \\
c_z, d_z &= z - \frac{l}{2}.
\end{aligned}
\tag{2}
$$

Table 6 shows the vertices calculation for both the ego-vehicle and bounding boxes. We also model the ego-vehicle as a bounding box to analyze the interaction of each corner of it with the detected objects. Thus, we consider the position of the camera on the vehicle's rooftop as the origin x, z. It is important to note that the camera position does not correspond to the vehicle's centroid, and therefore, it is necessary to calculate l_{cf} and l_{cr}, as shown in Figure 4(b). Furthermore, we assume that $\psi = 0$ since we analyze the interactions with objects detected from images captured by the AV front camera.

Table 6: Relation between the centroid position in the bounding box and the ψ_{obj} rotation.

Vertices	**Bounding Box**		**nuScenes/Lyft5**	
	x_{obj}	z_{obj}	x_{ego}	z_{ego}
a	$a_{x_{obj}} cos(\psi_{obj}) + a_{z_{obj}} sin(\psi_{obj}) + x_{obj}$	$-a_{x_{obj}} sin(\psi_{obj}) + a_{z_{obj}} cos(\psi_{obj}) + z_{obj}$	$a_{x_{ego}}$	$z_{ego} + l_{cf}$
b	$b_{x_{obj}} cos(\psi_{obj}) + b_{z_{obj}} sin(\psi_{obj}) + x_{obj}$	$-b_{x_{obj}} sin(\psi_{obj}) + b_{z_{obj}} cos(\psi_{obj}) + z_{obj}$	$b_{x_{ego}}$	$z_{ego} + l_{cf}$
c	$c_{x_{obj}} cos(\psi_{obj}) + c_{z_{obj}} sin(\psi_{obj}) + x_{obj}$	$-c_{x_{obj}} sin(\psi_{obj}) + c_{z_{obj}} cos(\psi_{obj}) + z_{obj}$	$c_{x_{ego}}$	$z_{ego} - l_{cr}$
d	$d_{x_{obj}} cos(\psi_{obj}) + d_{z_{obj}} sin(\psi_{obj}) + x_{obj}$	$-d_{x_{obj}} sin(\psi_{obj}) + d_{z_{obj}} cos(\psi_{obj}) + z_{obj}$	$d_{x_{ego}}$	$z_{ego} - l_{cr}$

Next, we analyze when an intersection exists between ego-vehicle vertices and bounding boxes converging to the AV path. For this, data from the detected object vertices and ego-vehicle vertices are analyzed to determine the interactions between them. For this analysis, a line segment is defined as the line connecting the adjacent vertices of the bounding box. We define line's equation for each selected bounding box segment of both the object and the ego-vehicle and potential intersections are calculated, as shown in Equation 3:

$$
\begin{aligned}
A_{ego}x + B_{ego}z &= C_{ego},\\
A_{obj}x + B_{obj}z &= C_{obj},
\end{aligned}
\tag{3}
$$

where A, B, and C correspond to the line's equation values for each segment of the bounding box (object) interacting with the ego-vehicle. These values are given by a set of conditions that depend on the detected object's orientation.

Once the line equations have been calculated, the resulting values are used to compute the intersection coordinates at x, z:

$$
\begin{aligned}
x_{ego \cap obj} &= \frac{(B_{ego}C_{obj}) - (B_{obj}C_{ego})}{(A_{ego}B_{obj}) - (A_{obj}B_{ego})},\\
z_{ego \cap obj} &= \frac{(A_{obj}C_{ego}) - (A_{ego}C_{obj})}{(A_{ego}B_{obj}) - (A_{obj}B_{ego})}.
\end{aligned}
\tag{4}
$$

Then, the distance d is calculated between the potential conflict vertices and segments between the detected object and the ego-vehicle:

$$
d = \sqrt{(x_{seg_{ego}} - x_{ego \cap obj})^2 + (z_{seg_{ego}} - z_{ego \cap obj})^2}. \tag{5}
$$

The geometric analysis of the ego-vehicle in relation to any detected object, using the motion orientation of the latter, allows us to evaluate diverses eventualities:

- **Identification of the first impact point**: This methodology enables to evaluate the first point of impact of the ego-vehicle with any detected object that is on the collision course at a given instant of time.

- **Approach type**: The geometric analysis also enables evaluating points of potential impact on the detected object. Furthermore, how the approximations occur can help to understand how the AV decision-making occurs. On the other hand, it is also possible to analyze other sensors, such as the side cameras; In this work we only analyze the interactions detected in the front camera. Side cameras analysis is not part of this work.

- **Interaction with other road users**: In addition to analyzing other vehicles interacting with the AV, it is also possible to evaluate how interactions with other road users occur, e.g., pedestrians, two-wheelers, objects, animals, among others. However, this analysis depends on the semantic data and categorization of objects detected by the AV.

Finally, it is possible to identify the location of objects around the AV. As shown in Figure 5, this methodology allows observing the potential impact point on the ego-vehicle ($x_{seg_{ego}}, z_{seg_{ego}}$), on the detected object ($x_{ego \cap obj}, z_{ego \cap obj}$), as well as the distance (d) between those potential impact points. This analysis also enables evaluating potential impact points with other bounding box segments from detected objects, e.g., segments $\overline{\mathrm{ab}}$, $\overline{\mathrm{bc}}$, $\overline{\mathrm{cd}}$, $\overline{\mathrm{da}}$. In the ego-vehicle, just the segment $\overline{ab}$ is analyzed. Moreover, in order to determine not only moving objects, as reported in (Kusano et al., 2014), the goal is to define also when movable/static objects (e.g., vehicles parked, traffic signals, among others) can provoke AV evasive actions that may represent potential risk events immediately. Thus, this work aims to evaluate the interactions between vertices of both AV and movable/static objects. This is important considering that although the proximity of the AV to other objects is inherent in the vehicular environment (e.g., adjacent vehicles, crosswalks, crossing vehicles, among others), and therefore some risk events can result in false positives.

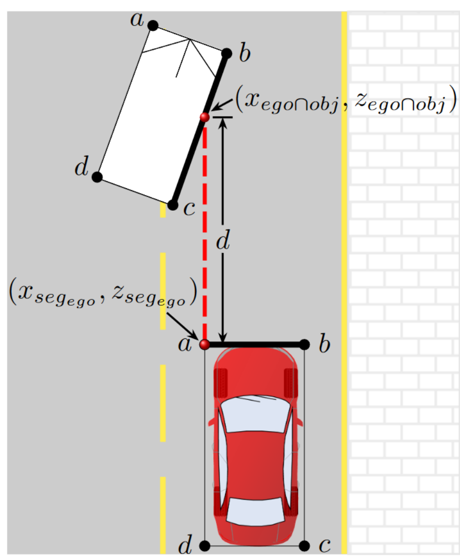


Figure 5: Geometric analysis of line segments, intersections and distance between the ego-vehicle and the detected objects. The segment analyzed in the detected object is $\overline{bc}$, and segment $\overline{ab}$ in the ego-vehicle.

This analysis allows describing various interactions with surrounding objects detected by the AV. Nevertheless, it is necessary to quantify the risk when the AV is on a collision course. For that, this work uses the TTC considering the detected objects' orientation as a metric to improve the analysis of traffic risk events involving the AV. The goal is to propose an improved TTC and test it with real data from AVs.

4. Time-to-Collision with Motion Orientation

From the analysis of camera images, it is possible to determine the position of objects. We can derive both the absolute location of the object and the position in the image through projections from 2D camera frames. As shown in Figure 6, it is possible to analyze the mapping between the world coordinate system and camera coordinate system that corresponds to the coordinate system used for vehicle navigation. Also, the object's speed related to the ego-vehicle is calculated by measuring the time difference between the sending and rebounding laser pulses from the LiDAR sensor.

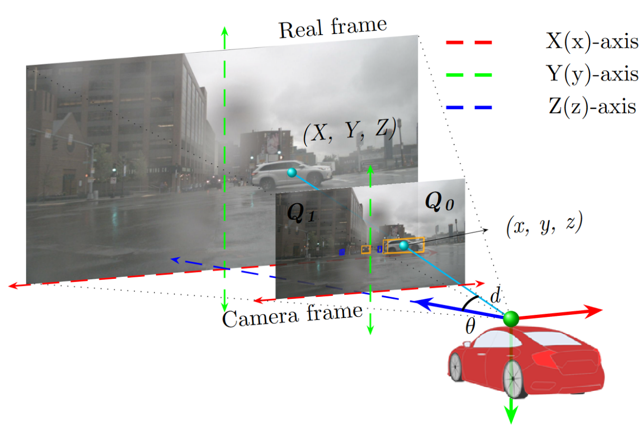


Figure 6: Mapping between a real frame and the camera frame.

To reduce the shortcomings of SSMs proposed in the literature, we include the motion orientation and position of objects detected by the AV as a parameter for the TTC calculation. The goal is to improve the accuracy of TTC to assess risk events for AV. Equation 6 summarizes the computation of TTC_{mo}:

$$TTC_{mo} = \begin{cases} \frac{d_{(obj,ego)}}{v_{ego} - v_{obj}\cos(\psi_{obj})}, \text{ if } (v_{ego} - v_{obj}\cos(\psi_{obj})) > 0, \\ \infty : \begin{cases} \text{if } (b_{ego} + sa_2) < (a,b,c,d)_{obj} < (a_{ego} - sa_1), \\ \text{or } (v_{ego} - v_{obj}\cos(\psi_{obj})) < 0, \\ \text{or } v_{ego} = 0, \end{cases} \end{cases} \tag{6}$$

where d is the distance between the segment/vertex on the ego-vehicle's course and the front-side of the ego-vehicle, v_{ego} is the speed of the ego-vehicle, v_{obj} is the speed of the detected object, and ψ_{obj} is the yaw angle of the detected object. The product of v_{obj} and ψ_{obj} captures the influence of the speed component on the same axis of the ego-vehicle shift (z-axis), since the geometric analysis uses the camera's reference system. On the other hand, $\mathrm{TTC}_{\mathrm{mo}}$ tends to infinity when none of the bounding box vertices of detected objects are in the path of the AV or invading the safety area (sa). Likewise, it is assumed that when detected objects with speed higher than the AV. Finally, when the AV is stopped, it is inferred that there will be no risk event. The speed values of the detected objects and the ego-vehicle in the AV nuScenes are obtained directly from the dataset. On the other hand, the speed data of the ego-vehicle in the Lyft5 dataset is obtained from the analysis of translation data by means of the haversine formula (Ivis, F., 2006).

As a result, it is possible to differentiate traffic risk events in both car following and head-on scenarios. Therefore, $\mathrm{TTC_{mo}}$ is conditioned to the yaw orientation of each road user detected by the ego-vehicle. Thus, for detected objects whose position indicates that they are on a collision course with the ego-vehicle, as defined in Section 3.2.1, orientation angles between $-90^{\circ} < 0^{\circ} < 90^{\circ}$ indicate that road users heading diverges from the ego-vehicle heading, describing car-following or crossing scenarios. Thus, the speed component $v_{obj}\cos(\psi_{obj}) \geq 0$, and $\mathrm{TTC_{mo}}$ value can only be calculated when a positive speed difference between the vehicles exists (Minderhoud and Bovy, 2001), which corresponds to the general definition of TTC. On the other hand, orientation angles between $90^{\circ} < 180^{\circ} < -90^{\circ}$ describe head-on or crossing scenarios, where the speed component is $v_{obj}\cos(\psi_{obj}) < 0$, and indicate that the road users heading converges with the motion direction of the ego-vehicle. Therefore, different from the general definition of TTC, $\mathrm{TTC_{mo}}$ value is calculated by adding the speeds of the ego-vehicle and the road user, following the definition in (Laureshyn et al., 2010).

As shown in Figure 7, scenarios 7(b), 7(c), 7(d), and 7(e) show the detection of objects which move in the same direction as the AV, in a car following event; scenarios 7(f), 7(g), 7(h), and 7(i) show the detected objects traveling in opposite direction to the ego-vehicle, configuring a head-on event; finally, scenarios 7(j), 7(k), and 7(l) show AV interactions with objects converging to the AV or a next point in a perpendicular trajectory, configuring a crossing event. Thus, it is possible to determine if the AV may be on a collision course with other road users with which the AV interacts continuously.

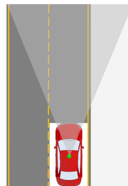
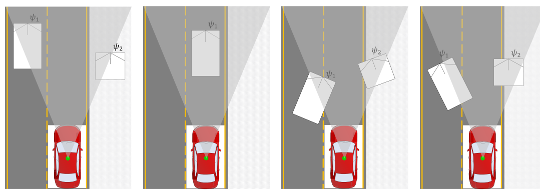
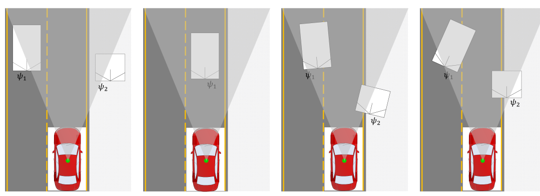
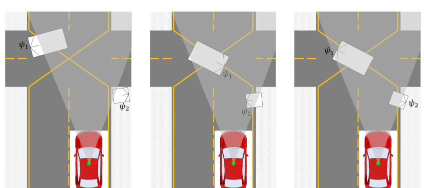
Scenarios	Characteristics
 (a)	Ego-vehicle has not detected objects.
 (b) (c) (d) (e)	**Car-following scenarios** • Ego-vehicle detects objects moving in same direction, but they are not on a collision course with it (Fig. 7(b)). • Ego-vehicle detects objects ahead (Fig. 7(c)). • Ego-vehicle detects objects converging (Fig. 7(d)), or diverging (Fig. 7(e)), still remaining in the vehicle's path.
 (f) (g) (h) (i)	**Head-on scenarios** • Ego-vehicle detects objects moving in opposite direction, but they are not on a collision course with it (Fig. 7(f)). • Ego-vehicle detects objects converging directly to it (Fig. 7(g)). • Ego-vehicle detects objects converging (Fig. 7(h)), or diverging (Fig. 7(i)), still remaining in the vehicle's path.
 (j) (k) (l)	**Crossing scenarios** • Ego-vehicle detects objects in perpendicular direction, but they are not on a collision course with it (Fig. 7(j)). • Ego-vehicle detects objects crossing towards the lane where it is moving. (Fig. 7(k)), still remaining in the vehicle's path. • Ego-vehicle detects objects crossing out the lane where it is moving. (Fig. 7(l)), still remaining in the vehicle's path.

Figure 7: Possible scenarios for detected objects ahead identified by the position and motion orientation.

To quantify the risk level from the $\mathrm{TTC}_{\mathrm{mo}}$ analysis, we employ the risk coefficients proposed in (Li et al., 2017). This criterion gathers values which correspond to the reaction time requirements in AVs, based on the parameters described in (Rydzewski and Czarnul, 2021). Table 7 shows the risk coefficient defined according to the $\mathrm{TTC}_{\mathrm{mo}}$ values.

Table 7: Risk coefficient as a function of TTC.

Severity grade	**$\mathrm{TTC_{mo}}$ [s]**	**Description**	**Risk coefficient**
0	> 4.0	No safety risk	0.0
1	2.5 to 4.0	Accident-to-conflict ratio stable	0.2
2	1.5 to 2.5	Low risk level	0.3
3	1.0 to 1.5	Moderate risk level	0.6
4	≤ 1.0	High risk level	0.8

Motion orientation has a direct impact on the safety analysis. The road users' is random by nature, therefore it is inferred that traffic risk events require a mapping analysis of the detected objects around the ego-vehicle. Next, $\mathrm{TTC_{mo}}$ analysis is used on the AV datasets presented in Section 3.1.

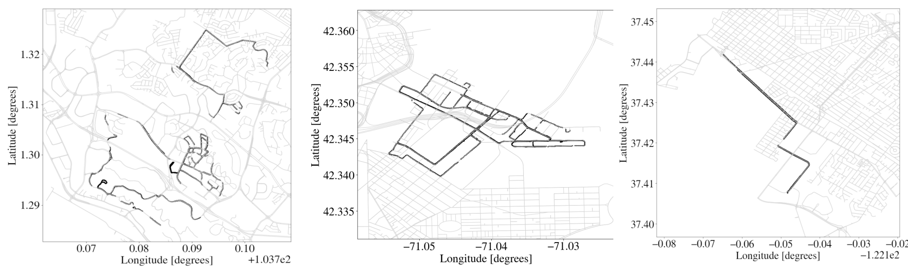
5. Performance Evaluation

We analyze factors that can compromise vehicle and passengers safety. For this, we focused on vehicle tracking, speed limit based on traffic regulations and $\mathrm{TTC_{mo}}$ to estimate the risk of ego-vehicle interactions with other road users.

5.1. Vehicle tracking

The frequency of each event is influenced by the topology of the cities where the AVs circulate, as shown in Figure 8. To analyze the vehicle tracking, we enriched the datasets with data related to road type and speed limit. The ego pose data encoded in translation data are transformed into geodetic coordinates to track the vehicle. Then, we use geodetic coordinates are used to make queries in Nominatim[1] and Overpass API[2].

[1] https://nominatim.org/
[2] https://overpass-turbo.eu/



(a) Singapore. (b) Boston. (c) Palo Alto.

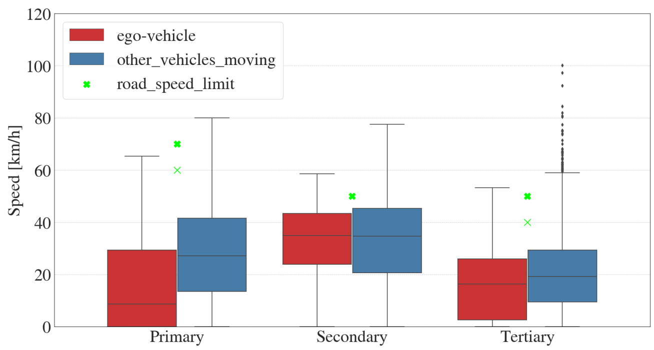
Figure 8: Trajectories of the AVs in the datasets.

5.2. Speed limit analysis

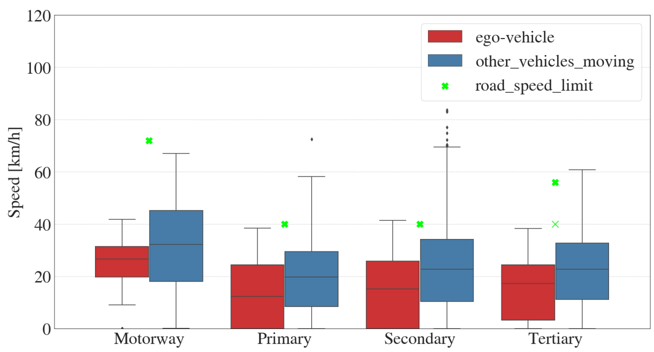
From the vehicle tracking analysis in Section 5.1, the ego vehicle speed profile is verified to ensure compliance with traffic regulations. Figure 9 shows that the ego vehicle maintains an average speed between 15 km/h and 30 km/h in Boston, 20 km/h and 40 km/h in Singapore, and between 30 km/h and 50 km/h in Palo Alto. Likewise, the speed of vehicles moving in front of the ego vehicle is analyzed. It is possible to observe that some samples exceed the threshold speed limit established by the traffic regulations; obviously relevant information given that speeding increases the probability of risky events.

5.3. TTC_{mo} evaluation

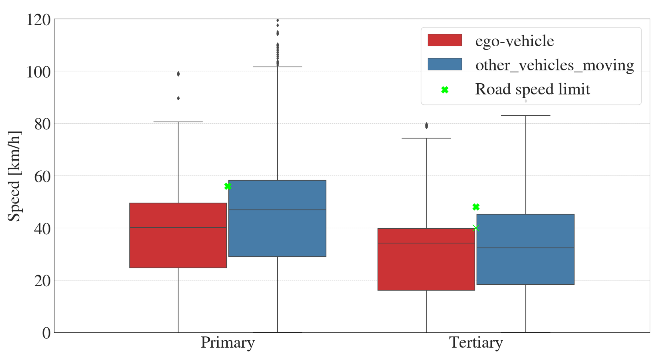
Kinematic measures like speed and distance from the detected objects are used for the $\mathrm{TTC}_{\mathrm{mo}}$ calculation. Speed and distance are estimated through LiDAR measurements, while the images are used for the recognition of the various objects around the AV. Data is available in the datasets in form of annotations and metadata for each instance (object) detected by the AV. Moreover, annotations are identified by categories, each one associated with each object detected.



(a) Ego-vehicle speed vs. road speed limit in Singapore.



(b) Ego-vehicle speed vs. road speed limit in Boston.



(c) Ego-vehicle speed vs. road speed limit in Palo Alto.

Figure 9: Relationship between the ego-vehicle speed and the road speed limit. Green marks describe the maximum speed allowed for each road type. Some road types have different speed limits; these are identified with thick and thin marks

Table 8 shows the observation statistics for the objects detected by the frontal camera.

Table 8: Categories instances in the AV datasets.

Dataset	nuScenes		Lyft5
	Singapore	**Boston**	**Palo Alto**
Images analyzed	14,106	18,617	21,640
Instance annotations	11,308	21,251	10,525
Sample annotations	107,615	225,957	227,043
Vehicles	46,262	137,927	211,287
Two-wheelers	4,373	2,835	7,039
Pedestrians	25,915	36,221	8,672
Animals	36	121	45
Traffic objects	31,029	49,853	–

To analyze potential risk events, AV datasets are examined to assess the driving behavior. For that, annotations made to images captured by the front camera are analyzed. Annotations with no speed data are discarded: 5% from the Lyft5 dataset, 4.3% from the Boston subset, and 1.1% from the Singapore subset. Next, we evaluate the regular TTC defined in Equation 1 for all the valid annotations, in order to observe the proportion of objects interacting with the ego-vehicle. In proportion, approximately 70% of the samples represent some risk level w.r.t. valid ones, as shown in Figure 10. Different from the analysis with the regular TTC, which only discards events when $v_{obj} > v_{ego}$, the $\mathrm{TTC}_{\mathrm{mo}}$ methodology proposed in Section 3 allows to determine which objects may be in the ego-vehicle's course. Therefore, objects that are not in the course of the ego-vehicle, or those whose exceeds the position angle threshold defined in Section 3.2.1 are discarded, since they do not represent a potential traffic conflict. Therefore, Annotations of objects converging to the ego-vehicle's course or the safety zone defined in Section 3.2.2 are analyzed. The proportion of samples representing some risk w.r.t. valid ones corresponds approximately 4% to 5% for Palo Alto and Singapore subsets, and approx. 8% for the Boston subset, as observed in Figure 10.

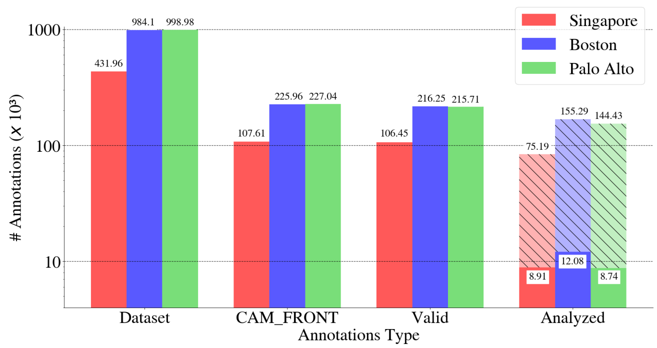


Figure 10: Number of annotations ($\times 10^3$) assessed for the analysis of potential risk events in the AV datasets studied. Hatch pattern bars in Analyzed label on x-axis correspond to the TTC general formulation analysis; solid color bars correspond to the TTC_{mo} proposed in this work.

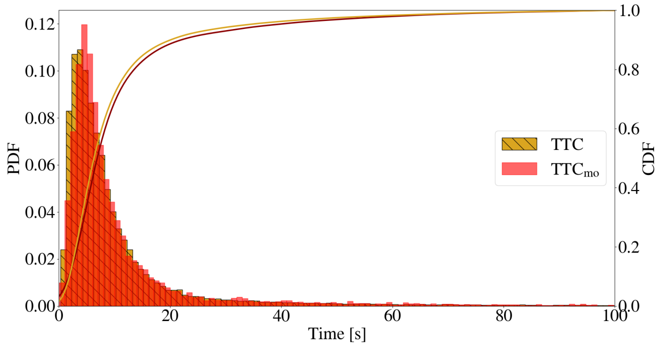
Figure 11 shows the TTC_{mo} and conventional TTC frequency distributions for each analyzed dataset. It is possible to observe that the distribution in all cities is very similar, with distributions skewed to the right. Therefore, the 5^{th} and the 85^{th} percentiles are evaluated, which represent the most pronounced inflection points in the cumulative distribution. Values below the 5^{th} percentile represent TTC values $< 2.4\,\text{s}$ in all datasets. We also note that the bulk of representative TTC samples are concentrated in up to $33\,\text{s}$, with an average of maximum $18\,\text{s}$. On the other hand, comparing the distribution of TTC and TTC_{mo}, it is possible to observe that the TTC_{mo} distribution in Singapore and Boston subsets is smaller than TTC distribution, which allows us to observe a trend towards a decrease in the frequency of high-risk events. Therefore, TTC_{mo} appears to have a more precise collision course compared to TTC, which leads to a stricter definition of conflicts and less data to be analyzed. Meanwhile, the distribution of TTC_{mo} in Palo Alto is contrary to the data trend in Singapore or Boston. It is possible to observe an increment in the frequency of events with time $< 10\,\mathrm{s}$, however, the frequency in time $> 10\,\mathrm{s}$ decreases compared with TTC distribution. This trend can be influenced by interactions with parked vehicles along the AV route.

From the annotations analyzed in Figure 10, it is possible to observe the frequency and the type of events concerning potential risk events, when both the objects and the ego-vehicle are in collision course. Table 9 shows the total frequency of event types based on the course of detected objects, as described in Section 3.2.1, classified as following, head-on, and crossing events. Course analysis can help to analyze the way in which these objects converge with the AVs. These data are important to consider the severity of the event. For example, a car-following event can have a different effect than a head-on event.

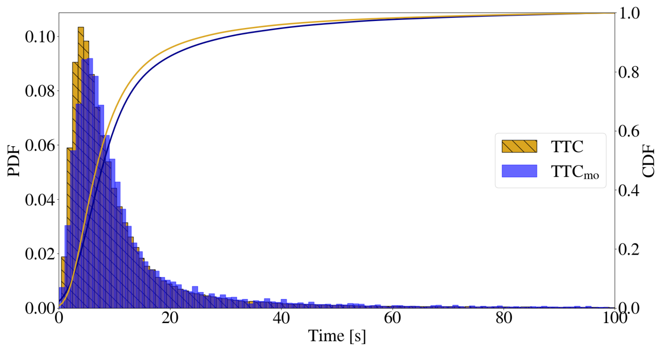
Table 9: Conflict types defined by position and orientation concerning to the ego-vehicle.

Event/City	**Singapore**	**Boston**	**Palo Alto**
Following	3,094	4,452	7,022
Head-on	595	824	147
Lane-change	267	349	104
Crossing	4,947	6,456	1,463
Total events	8,903	12,081	8,736

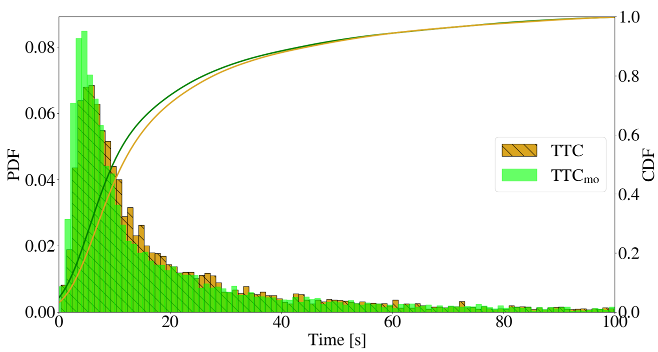
To analyze the risk level of the ego-vehicle interactions with other objects, we use the severity hierarchy based on the level and severity zones proposed by Hydén (Hydén, 1987). Severity level defines a threshold for serious and non-serious conflicts. On the other hand, severity zones quantitatively define severity levels. Both severity level and zones are based on a relationship between time and speed. A fixed threshold to define a high-risk event is based on the



(a) Singapore.



(b) Boston.



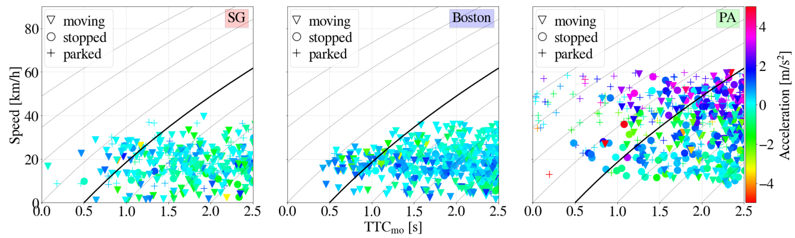
(c) Palo Alto.

Figure 11: Cumulative and Probability Density Functions for $\mathrm{TTC}_{\mathrm{mo}}$ and $\mathrm{TTC} < 100\,\mathrm{s}$ for each dataset.

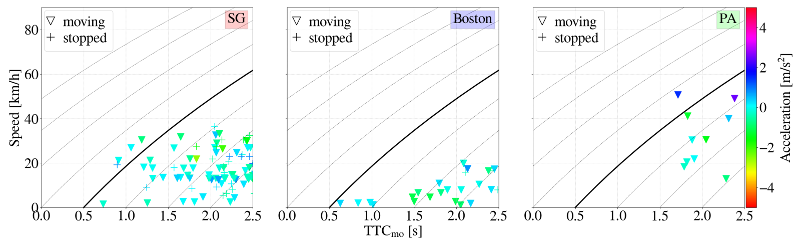
Time-to-Accident (TA) under a traffic conflict. This value was established at 1.5 s (Hydén, 1987), which is consistent with the studies reported in (Rydzewski and Czarnul, 2021), and that corresponds to the response time of the sensors readings, processing, recognition and planning tasks of the AV between the detection of an obstacle and the evasive action.

All interactions that represent some risk level for the ego-vehicle are presented in Figure 12. All interactions within the 5^{th} percentile are plotted, as observed in the cumulative distributions of Figure 11. We note that most of the observed interactions in SG and Boston occur with vehicles and objects. Figure 12(a) shows that interactions with $\mathrm{TTC}_{\mathrm{mo}} < 1.5\,\mathrm{s}$ occur with other moving vehicles, with a deceleration pattern as the $\mathrm{TTC}_{\mathrm{mo}}$ decreases. On the other hand, in PA we observe more interactions with parked vehicles. This characteristic is due to Lyft5 vehicles move along the roadside parking areas, next to the first lane at right, where some parked vehicles are invading the safety area *(sa)* defined for the AV. On the other hand, it is interesting to note that interactions with pedestrians show some events that represent lower risk of collision, as shown in Figure 12(b). The same behavior is observed for the two-wheelers in Figure 12(c). Finally, Figure 12(d) shows the interactions with objects of the vehicular infrastructure like barriers, traffic cones, among others.

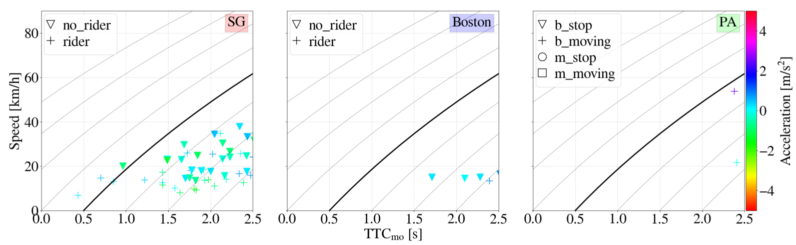
To summarize, the proportion of interactions for all the annotations analyzed represents less than 1% for high-risk events, whereas events with some risk represent approximately 10%. Events that do not represent any risk represent more than 70%, as shown in Figure 13. Compared to valid annotations, the proportion of interactions that represent some risk level is less than 2%. This is consistent with the results observed in (Beauchamp et al., 2022) and (De Ceunynck et al., 2022), where it is observed that most traffic events are not risky



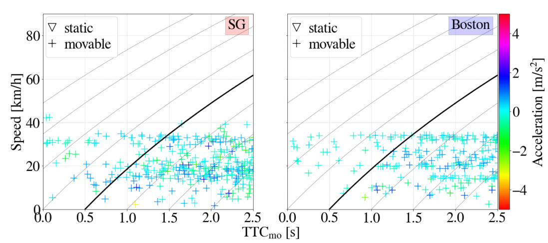
(a) Vehicles.



(b) Pedestrians.



(c) Two-wheelers.



(d) Objects.

Figure 12: TTC$_{\text{mo}}$ 5$^{\text{th}}$ percentile indicators for each scenario in relation to ego-vehicle speed and acceleration. Acceleration changes are shown in heatmap color variations. The columns describe the city where the interactions take place: to the left Singapore (SG), to the center Boston, and to the right Palo Alto (PA). Meanwhile, the rows describe the general category of objects interacting with the AV. Conflicts above the black line on the graphs are ranked as serious; below the black line, non-serious.

for the ego-vehicle. Furthermore, compared to (Li et al., 2017), it is observed that moderate and high risk events have a lower proportion. Nonetheless, the experimentation environment is different, and the results are expressive due to our analysis takes advantage of the sensors mobility and the variability of the scenarios where the vehicles transit.

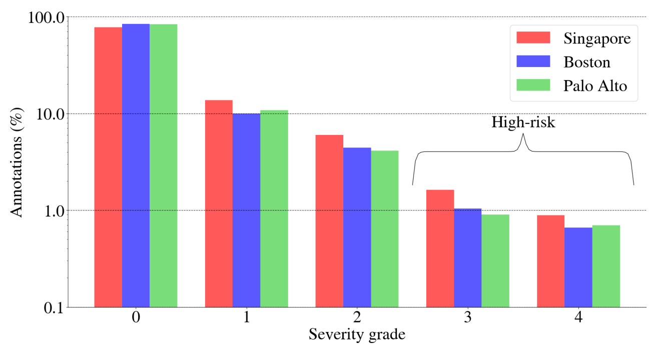


Figure 13: Annotation volume based on severity grade ratio.

The present $\mathrm{TTC}_{\mathrm{mo}}$ analysis allows to assess risk events through the geometric analysis of the boundaries associated with each object detected by the AVs. Thus, it is possible to limit the analysis to objects in a possible collision course. This is relevant for $\mathrm{TTC}_{\mathrm{mo}}$ analysis since it is possible to identify how interactions occur with various road users and objects. Nevertheless, further investigation is needed to establish a pattern of AV behavior with a longer time sequence in the scenes, mainly to obtain more parameters to describe driving behavior patterns related the AI system that controls the vehicle.

An advantage of data analysis through exteroceptive sensors is that risk assessment is not limited to claims related to vehicles only. This is observed in Figure 12, where the $\mathrm{TTC}_{\mathrm{mo}}$ is assessed for various categories and attributes available in the datasets. Moreover, the distribution of risk events was similar among the three datasets, with 85% of the sampling concentrated in less than 33 s, and the highest risk events below 2.4 s, as shown in Figure 11. It is also important to note that the analysis of safety metrics for various road users will depend on the data labeling available. This can be observed for example in Table 8, where the Lyft5 dataset does not have data related to traffic infrastructure objects. It is important to explainability requirements to understand traffic conflicts between detected objects and the ego-vehicle, based on the road users motion.

It should be noted that there are some limitations in the used AV datasets. The sampling time of each scene is limited to a maximum of 25 s (Lyft5), and 20 s (nuScenes), in most cases without sequence, which prevents observing a greater number of events with potential risk. Another limitation is related with the speed of the AVs analyzed, which is much lower than the limit speeds of the road infrastructure. The speed uniformity of the AVs reduces the possibility of observing the effect of the evasive actions by the AV. Finally, the number of vehicles limits the risk assessment analysis since the age and learning experience of the autonomous system may still be limited.

On the other hand, sensor-associated errors can influence the risk analysis of the ego-vehicle. Despite the existence of errors in both translation and speed in both datasets, object detection based on LiDAR and the camera perform well in image-only methods to infer the dimensions of the detected objects and their kinematic measurements (Caesar et al., 2020). In fact, object detection is a challenging area since objects around are not symmetric, contain different footprints, and therefore, the computation of bounding boxes generation is a hard problem. It is important noting that $\mathrm{TTC_{mo}}$ is a metric that depends on the object detection and kinematic variables related to the object, and therefore, requires high precision of the sensors. Otherwise, $\mathrm{TTC_{mo}}$ calculation may result in erroneous measurements influenced by errors of the autonomous system driving the vehicle.

Finally, the calculation of $\mathrm{TTC_{mo}}$ considering the motion orientation of the detected objects reduces the overload generated by the volume of data in the safety analysis. Thus, our improved $\mathrm{TTC_{mo}}$ reduces by up to 60% the proportion of data to be analyzed when compared to the regular TTC. Motion orientation and geometry analysis enable to discard all objects that, despite interacting with the AV, they do not converge on a collision course, and therefore, they do not represent a risk for AV. It is relevant if we consider that safety monitoring requires immediate analysis when exists potential traffic conflicts.

6. Conclusion and Future Work

This study aims to explore the potential of using AV data to identify high-risk events in traffic by analyzing TTC and motion orientation. Real data collected from AVs in different cities was used to identify risk events. A detailed data analysis and processing are presented in this study, in addition to serving as a guide for other researchers who want to use public AV datasets. In particular, we study traditional SSM like TTC considering the motion orientation of the road users detected by the AV. This is a scenario few explored since the information of the road users is limited when there are no direct measurements of them. By the motion orientation, it is possible to analyze diverse scenarios like following, head-on and crossing events. This allows a more intuitive safety analysis related to all detected objects moving in different directions.

As future work, our goal is to describe traffic risk events accurately to improve the risk assessment process in AVs. For this, it is necessary to manage the data to optimize the analysis of the data collection available in each ego vehicle, so that it is scalable, and responds to the immediateness of risk assessments.

Acknowledgments

This work was funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil (CAPES) – Finance Code 001, CNPq (grant #304142/2017-4), FAPERJ (grant #E-26/202.932/2017), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) grant #15/24494-8, and Axa Group – France.

References

Alghodhaifi, H., Lakshmanan, S., 2020. Simulation-based model for surrogate safety measures analysis in automated vehicle-pedestrian conflict on an urban environment, in: Auton. Syst.: Sens. Process. Secur. Veh. Infrastruct., pp. 8–21.

Alozi, A.R., Hussein, M., 2022. Evaluating the safety of autonomous vehicle–pedestrian interactions: An extreme value theory approach. Anal. Methods Accid. Res. 35, 100230. https://doi.org/10.1016/j.amar.2022.100230.

Aycard, O., Baig, Q., Bota, S., Nashashibi, F., Nedevschi, S., Pantilie, C., Parent, M., Resende, P., Vu, T., 2011. Intersection safety using lidar and stereo vision sensors, in: IEEE Intell. Veh. Symp., pp. 863–869. https://doi.org/10.1109/IVS.2011.5940518.

Beauchamp, É., Saunier, N., Cloutier, M.S., 2022. Study of automated shuttle interactions in city traffic using surrogate measures of safety. Transp. Res. Part C: Emerging Technol. 135, 103465. https://doi.org/10.1016/j.trc.2021.103465.

Betz, J., Heilmeier, A., Wischnewski, A., Stahl, T., Lienkamp, M., 2019. Autonomous driving—a crash explained in detail. Appl. Sci. 9. https://doi.org/10.3390/app9235126.

Caesar, H., Bankiti, V., Lang, A.H., Vora, S., Liong, V.E., Xu, Q., Krishnan, A., Pan, Y., Baldan, G., Beijbom, O., 2020. nuScenes: A multimodal dataset for autonomous driving, in: IEEE/CVF Conf. Comput. Vision .Pattern Recognit. (CVPR), pp. 11618–11628. https://doi.org/10.1109/CVPR42600.2020.01164.

Campbell, K., (U.S.), S.S.H.R.P., 2012. The SHRP 2 Naturalistic Driving Study:

Addressing Driver Performance and Behavior in Traffic Safety. Transp. Res. Board.

De Ceunynck, T., Pelssers, B., Bjørnskau, T., Aasvik, O., Fyhri, A., Laureshyn, A., Johnsson, C., Hagenzieker, M., Martensen, H., 2022. Interact or counteract? behavioural observation of interactions between vulnerable road users and autonomous shuttles in oslo, norway. Traffic Saf. Res. 2, 000008. https://doi.org/10.55329/fbhr3456.

Dingus, T.A., Klauer, S.G., Neale, V.L., Petersen, A., Lee, S.E., Sudweeks, J., Perez, M.A., Hankey, J., Ramsey, D., Gupta, S., et al., 2006. The 100-car naturalistic driving study, Phase II-results of the 100-car field experiment. Technical Report. U.S. NHTSA.

Geiger, A., Lenz, P., Urtasun, R., 2012. Are we ready for autonomous driving? the KITTI vision benchmark suite, in: IEEE Conf. Comput. Vision Pattern Recognit. (CVPR), pp. 3354–3361. https://doi.org/10.1109/CVPR.2012.6248074.

Hayward, J., 1972. Near-miss determination through use of a scale of danger. Highway Res. Rec. Report TTSC 7115, 24–34.

He, Z., Qin, X., Liu, P., Sayed, M.A., 2018. Assessing surrogate safety measures using a safety pilot model deployment dataset. Transp. Res. Rec. 2672, 1–11. https://doi.org/10.1177/0361198118790861.

Houston, J., Zuidhof, G., Bergamini, L., Ye, Y., Jain, A., Omari, S., Iglovikov, V., Ondruska, P., 2020. One thousand and one hours: Self-driving motion prediction dataset. CoRR abs/2006.14480. https://doi.org/10.48550/arXiv.2006.14480.

Hydén, C., 1987. The Development of a Method for Traffic Safety Evaluation: The Swedish Traffic Conflicts Technique. Ph.D. thesis. Lund University. Sweeden.

IEEE Electronics Packaging Society, 2019. Automotive, in: Heterogeneous Integration Roadmap. IEEE, USA. chapter 5, pp. 1–22.

Ivis, F., 2006. Calculating geographic distance: Concepts and Methods. https://bit.ly/3wWn7AJ. Accessed September, 2022.

Jiménez, F., Naranjo, J.E., García, F., 2013. An improved method to calculate the time-to-collision of two vehicles. Int. J. Intell. Transp. Syst. Res. 11, 34–42. https://doi.org/10.1007/s13177-012-0054-4.

Kesten, R., Usman, M., Houston, J., Pandya, T., Nadhamuni, K., Ferreira, A., Yuan, M., Low, B., Jain, A., Ondruska, P., Omari, S., Shah, S., Kulkarni, A., Kazakova, A., Tao, C., Platinsky, L., Jiang, W., Shet, V., 2019. Lyft level 5 AV dataset. https://lft.to/3cGvAwk.

Kiefer, R.J., LeBlanc, D.J., Flannagan, C.A., 2005. Developing an inverse time-to-collision crash alert timing approach based on drivers' last-second braking and steering judgments. Accid. Anal. Prev. 37, 295–303. https://doi.org/10.1016/j.aap.2004.09.003.

Kilicarslan, M., Zheng, J.Y., 2019. Predict vehicle collision by TTC from motion using a single video camera. IEEE Trans. Intell. Transp. Syst. 20, 522–533. https://doi.org/10.1109/TITS.2018.2819827.

Kusano, K.D., Montgomery, J., Gabler, H.C., 2014. Methodology for identifying car following events from naturalistic data, in: IEEE Intell. Veh. Symp., pp. 281–285. https://doi.org/10.1109/IVS.2014.6856406.

Laureshyn, A., Åse Svensson, Hydén, C., 2010. Evaluation of traffic safety, based on micro-level behavioural data: Theoretical framework and first implementation. Accid. Anal. Prev. 42, 1637–1646. https://doi.org/10.1016/j.aap.2010.03.021.

Li, Y., Lu, J., Xu, K., 2017. Crash risk prediction model of lane-change behavior on approaching intersections. Discrete Dyn. Nat. Soc. 2017. https://doi.org/10.1155/2017/7328562.

Lyft SDK, 2019. Lyft Dataset SDK. https://github.com/lyft/nuscenes-devkit. Accessed September, 2022.

Mahmud, S.S., Ferreira, L., Hoque, M.S., Tavassoli, A., 2017. Application of proximal surrogate indicators for safety evaluation: A review of recent developments and research needs. IATSS Research 41, 153–163. https://doi.org/10.1016/j.iatssr.2017.02.001.

Markkula, G., Engström, J., Lodin, J., Bärgman, J., Victor, T., 2016. A farewell to brake reaction times? kinematics-dependent brake response in naturalistic rear-end emergencies. Accid. Anal. Prev. 95, 209–226. https://doi.org/10.1016/j.aap.2016.07.007.

Meilinger, T., Vosgerau, G., 2010. Putting egocentric and allocentric into perspective, in: Spatial Cognit. VII, pp. 207–221. https://doi.org/10.1007/978-3-642-14749-4_19.

Miller, R., Huang, Q., 2002. An adaptive peer-to-peer collision warning system, in: IEEE 55th Veh. Technol. Conf. (VTC Spring), pp. 317–321. https://doi.org/10.1109/VTC.2002.1002718.

Minderhoud, M.M., Bovy, P.H., 2001. Extended time-to-collision measures for road traffic safety assessment. Accid. Anal. Prev. 33, 89–97. https://doi.org/10.1016/S0001-4575(00)00019-1.

Montgomery, J., Kusano, K.D., Gabler, H.C., 2014. Age and gender differences in time to collision at braking from the 100-car naturalistic driving study. Traffic Inj. Prev. 15, S15–S20. https://doi.org/10.1080/15389588.2014.928703.

Nodine, E., Stevens, S., Lam, A., Jackson, C., Najm, W.G., et al., 2015. Independent evaluation of light-vehicle safety applications based on vehicle-to-vehicle communications used in the 2012-2013 safety pilot model deployment. Technical Report. U.S. NHTSA.

nuTonomy, 2018. nuscenes-devkit. https://github.com/nutonomy/nuscenes-devkit. Accessed September, 2022.

Ortiz, F.M., Sammarco, M., Costa, L.H.M.K., Detyniecki, M., 2022. Applications and services using vehicular exteroceptive sensors: a survey. IEEE Trans. Intell. Veh. 99, 1–20. https://doi.org/10.1109/TIV.2022.3182218.

Ozbay, K., Yang, H., Bartin, B., Mudigonda, S., 2008. Derivation and validation of new simulation-based surrogate safety measure. Transp. Res. Rec. 2083, 105–113. https://doi.org/10.3141/2083-12.

Papadoulis, A., Quddus, M., Imprialou, M., 2019. Evaluating the safety impact of connected and autonomous vehicles on motorways. Accid. Anal. Prev. 124, 12–22. https://doi.org/10.1016/j.aap.2018.12.019.

Qu, C., Qi, W.Y., Wu, P., 2018. A high precision and efficient time-to-collision algorithm for collision warning based V2X applications, in: 2nd Int. Conf. Rob. Autom. Sci. (ICRAS), pp. 1–5. 10.1109/ICRAS.2018.8443265.

Rydzewski, A., Czarnul, P., 2021. Human awareness versus autonomous vehicles view: comparison of reaction times during emergencies, in: IEEE Intell. Veh. Symp. (IV), pp. 732–739. https://doi.org/10.1109/IV48863.2021.9575602.

SAE, 2018. SAE Standard J3016: Taxonomy and Definitions for Terms Related to On-Road Motor Vehicles. Technical Report. Society of Automotive Engineers (SAE). https://doi.org/10.4271/J3016_202104.

Tak, S., Kim, S., Lee, D., Yeo, H., 2018. A comparison analysis of surrogate safety measures with car-following perspectives for advanced driver assistance system. J. Adv. Transp. 2018. https://doi.org/10.1155/2018/8040815.

Tarko, A., Davis, G., Saunier, N., Sayed, T., 2009. Surrogate measures of safety. Safe Mobility: Challenges, Methodology and Solutions-

(Transport and Sustainability) 11, 383–405. https://doi.org/10.1108/S2044-994120180000011019.

Virdi, N., Grzybowska, H., Waller, S.T., Dixit, V., 2019. A safety assessment of mixed fleets with connected and autonomous vehicles using the surrogate safety assessment module. Accid. Anal. Prev. 131, 95–111.

Wachenfeld, W., Junietz, P., Wenzel, R., Winner, H., 2016. The worst-time-to-collision metric for situation identification, in: IEEE Intell. Veh. Symp., pp. 729–734. https://doi.org/10.1109/IVS.2016.7535468.

Wang, W., Liu, C., Zhao, D., 2017. How much data are enough? a statistical approach with case study on longitudinal driving behavior. IEEE Trans. Intell. Veh. 2, 85–98. https://doi.org/10.1109/TIV.2017.2720459.

Ward, J.R., Agamennoni, G., Worrall, S., Bender, A., Nebot, E., 2015. Extending Time to Collision for probabilistic reasoning in general traffic scenarios. Transp. Res. Part C: Emerging Technol. 51, 66–82. https://doi.org/10.1016/j.trc.2014.11.002.

Xie, K., Yang, D., Ozbay, K., Yang, H., 2019. Use of real-world connected vehicle data in identifying high-risk locations based on a new surrogate safety measure. Accid. Anal. Prev. 125, 311–319. https://doi.org/10.1016/j.aap.2018.07.002.

Zhang, J., Wu, K., Cheng, M., Yang, M., Cheng, Y., Li, S., 2020. Safety evaluation for connected and autonomous vehicles' exclusive lanes considering penetrate ratios and impact of trucks using surrogate safety measures. J. Adv. Transp. 2020, 1–16.

Zheng, L., Ismail, K., Meng, X., 2014. Traffic conflict techniques for road safety analysis: open questions and some insights. Can. J. Civ. Eng. 41, 633–641. https://doi.org/10.1139/cjce-2013-0558.