

Validation and Analysis of Driving Safety Assessment Metrics in Real-world Car-Following Scenarios with Aerial Videos

Duo Lu, Sam Haines, Rider University

Varun Chandra Jammula , Prabin Kumar Rath , Hongbin Yu, Yezhou Yang, Arizona State University

Jeffrey Wishart, Science Foundation of Arizona,

Abstract

Data-driven driving safety assessment is crucial in understanding the insights of traffic accidents caused by dangerous driving behaviors. Meanwhile, quantifying driving safety through well-defined metrics in real-world naturalistic driving data is also an important step for the operational safety assessment of automated vehicles (AV). However, the lack of flexible data acquisition methods and fine-grained datasets has hindered progress in this critical area. In response to this challenge, we propose a novel dataset for driving safety metrics analysis specifically tailored to car-following situations. Leveraging state-of-the-art Artificial Intelligence (AI) technology, we employ drones to capture high-resolution video data at 12 traffic scenes in the Phoenix metropolitan area. After that, we developed advanced computer vision algorithms and semantically annotated maps to extract precise vehicle trajectories and leader-follower relations among vehicles. These components, in conjunction with a set of defined metrics based on our prior work on Operational Safety Assessment (OSA) by the Institute of Automated Mobility (IAM), allow us to conduct a detailed analysis of driving safety. Our results reveal the distribution of these metrics under various real-world car-following scenarios and characterize the impact of different parameters and thresholds in the metrics. By enabling a data-driven approach to address driving safety in car-following scenarios, our work can empower traffic operators and policymakers to make informed decisions and contribute to a safer, more efficient future for road transportation systems.

Introduction

Ensuring road safety is paramount in today's increasingly interconnected and technologically advanced world. Traffic accidents and dangerous driving behaviors remain significant challenges, and analyzing these issues with precision is a crucial step toward increasing road safety. Meanwhile, the rapid advancement of intelligent transportation systems demands innovative solutions for enhancing traffic situation awareness and vehicle-infrastructure coordination. Additionally, with the ongoing advancement of vehicles equipped with Automated Driving Systems (ADS), often denoted as "Automated Vehicles (AVs)", there exists a growing imperative among industrial stakeholders to establish a reliable procedure for assessing the operational safety performance of these vehicles. This procedure should offer a consistent, impartial, and technology-agnostic evaluation to instill public trust as automated vehicles become increasingly integrated into our transportation systems.

However, there are two main challenges in driving safety assessment. First, the primary roadblock is the development of objective and quantitative safety assessment methodologies that can be computed efficiently at the vehicle trajectory level (e.g., surrogate safety metrics [1]) instead of incident level (e.g., counting and analyzing

accidents). These metrics often require precise measurement of the location and speed of a set of traffic participants all the time, which is difficult to obtain. Second, these established methodologies need to be validated in various real-world traffic scenarios instead of simulations, especially from the perspective of traffic operators and policymakers instead of automated vehicle developers. This requires flexible, fine-grained data acquisition methods that can be easily deployed at any chosen traffic scene and datasets that cover a variety of traffic scenes. Such a dataset is difficult to construct. As a result, for traffic researchers, the deficiency in data and analysis tools has hindered the advancement of road safety efforts.

To address these challenges, in 2020, the Institute of Automated Mobility (IAM) introduced a comprehensive set of operational safety assessment (OSA) metrics [1], representing a pioneering contribution that encompasses both automated vehicles and human-driven vehicles. Subsequent best practices [2] were proposed by the Automated Vehicle Safety Consortium (AVSC) in 2021, aligning philosophically with many IAM OSA metrics. This industry-wide consensus has enabled the Society of Automotive Engineers (SAE) Vehicle and Vehicle (V&V) Task Force to embark on the development of a Recommended Practice for OSA metrics. These metrics form the cornerstone for the evaluation process, along with the development of the OSA Methodology [3] by the IAM, slated for inclusion in future standards documentation. Building upon this foundation, in 2021 and 2022, IAM researchers measured and evaluated a subset of the OSA metrics in simulation [4] as well as using real-world data collected at an intersection [5][6]. The focus is on safety envelope-type metrics within selected car-following scenarios from the list of pre-crash scenarios published by the National Highway Traffic Safety Administration (NHTSA) [7].

At the heart of the OSA metrics research lies the recognition that without accurate vehicle localization and adequate vehicle trajectories, it is exceptionally challenging to calculate these metrics and understand the complexities of road safety, especially in the context of car-following scenarios. In these scenarios, where vehicles closely trail one another, even the most minute driving behaviors can have profound safety implications. To fill this critical data gap, IAM researchers developed camera-based traffic data acquisition methods from the perspective of infrastructure [8][9] and drones [10] in collaboration with Arizona State University and the University of Arizona. With advanced Artificial Intelligence (AI) algorithms and deep neural network models trained from large amounts of data, these methods can obtain the vehicle trajectories on a map, which facilitates further analysis of driving behavior. Particularly, in [10], a drone data processing framework and a drone dataset of about 100 hours of high-resolution traffic videos collected from various diverse sites were constructed. This enables researchers to extract vehicle trajectories at the decimeter level and further analyze OSA metrics.

Based on these prior works, in this paper, we developed a method to automatically detect leader-follower vehicle pairs using a computer

program, which enables us to process more than 1.2 million data samples from 5,433 pairs of vehicles efficiently. This validation and analysis work is a follow-up to [4] and [5], but our analysis delves into the distribution of a few driving safety metrics across various real-world car-following scenarios under a wide range of different scenes rather than one or a few scenes with limited data points. Moreover, our metrics measurement data shed light on how different parameters and thresholds impact these metrics. Furthermore, by integrating our findings into the broader context of driving safety metrics for automated vehicles, our work strengthens the foundation for a comprehensive, data-driven approach to assess driving safety in car-following scenarios. The insights garnered from our research not only empower traffic operators and policymakers to make informed decisions but also play a pivotal role in shaping the future of safety assessment of road transportation systems with automated vehicles.

Traffic Scene Dataset

In this research, we obtained accurate vehicle trajectories at 12 different traffic scenes in the Phoenix metropolitan area, covering four different scene categories, including roundabouts, intersections, local road segments, and highway segments, with three scenes for each category. For each scene, we flew a consumer-grade drone (DJI Mavic Air 2) at about 120 meters to obtain a video track of about 21 minutes. During each data acquisition flight, the drone stayed still to maintain its position and camera view angle. All 12 traffic scenes are shown in the Appendix.

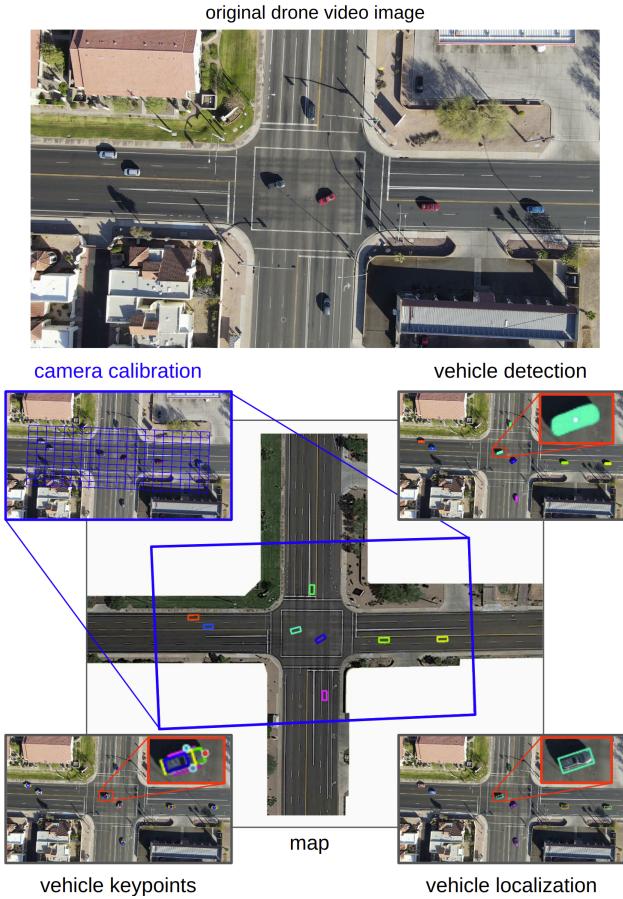


Figure 1. Example traffic scenes from four scenario categories.

The video data is processed using the CAROM Air framework [10], and an overview is shown in Figure 1. This framework consists of a pipeline to track vehicles on drone videos and obtain their trajectories in the 3D space. There are four steps detailed as follows.

Step 1) Camera calibraiton. We obtained the satellite map image of each traffic scene from a screenshot through popular online map services such as Google Maps or Microsoft Bing Maps. After that, we manually calibrated the first image in the drone video and the satellite map image to set up pixel-to-pixel coordinate transformation. For convenience, we assume the vehicle-traversable road areas covered by the drone video are on flat ground for all the traffic scenes. Based on our experience, this assumption only causes negligible vehicle localization errors. This step also sets up a 3D reference frame relative to the satellite map. Since the satellite image is to scale and has a relatively high resolution (6 ~ 7 cm per pixel), locations and distances can be measured precisely at the decimeter level. Once the video and the map are set up, the pipeline processes each video image in the sequence. To compensate for drifts and other motions of the drone during the video recording, the pipeline tracks the lane marker features on the ground, estimates the 3D pose of the drone camera for each video image, and corrects the coordinate system transformation relative to the manually annotated first image of the video.

Step 2) Vehicle detection. We trained a deep neural network based on Mask R-CNN [11] to detect the 2D bounding box, instance segmentation mask, and key points of each vehicle on every video image. These key points consist of observable features such as corners of the front and rear windshields, lights, bumpers, mirrors, etc. In total, 19 key points are detected. Given the same vehicle detected on two adjacent video images, we check their bounding boxes overlap and color consistency to associate them. As a result, each detected and associated vehicle is assigned a unique vehicle ID. Experimental evaluation shows that our pipeline can reliably detect about 94% of vehicles.

Step 3) Vehicle model fitting. We construct a morphable vehicle model with 33 skeleton vertices from statistical analysis of 200 vehicle 3D models collected from the Internet. These 3D models are manually built by artists for games and 3D animations. Our morphable model can be controlled by five parameters to change its shape to approximate a variety of real-world vehicles, including sedans, coupes, SUVs, mini-vans, vans, and pickup trucks. After that, we fitted it onto each detected vehicle on every 2D video image following the camera perspective geometry. Essentially, we developed an algorithm to solve the shape, location, and orientation of the morphable vehicle model such that the projected 3D skeleton vertices of the model best match the 2D key points detected on the image. Utilizing such a 3D morphable vehicle model is the key to allowing our pipeline to achieve decimeter-level vehicle localization accuracy in most cases, which further enables driving safety metrics calculation and analysis.

Step 4) Vehicle state estimation. We track each detected vehicle in the 3D space and continuously estimate its location and orientation, image by image, using an Extended Kalman Filter (EKF) in a simplified vehicle kinematic bicycle model. This step is crucial to allow our pipeline to recover from occasional misdetection of vehicle instances on some video images. Moreover, it further allows us to obtain other vehicle motion states, such as velocity and acceleration, beyond just location and orientation. Given the vehicle's kinematic states, we save them into a file. Once a video is processed, we use the saved kinematic states of vehicles to reconstruct the trajectory and export them together with the map as our traffic scene datasets.

Vehicle Trajectory Postprocess

To compute the driving safety assessment metrics in the car following scenarios, we need to derive vehicle leader-follower pairs and vehicle motion states (e.g., speed and acceleration) beyond the trajectory of individual vehicles. We developed three postprocessing stages to achieve this goal: trajectory smoothing, automated leader-follower identification, and pair filtering. They are detailed in the following subsections.

Trajectory Smoothing

Given a vehicle trajectory with vehicle locations and orientations, we smoothed each trajectory point and estimated the speed and acceleration by fitting a quadratic curve on the trajectory within a sliding window. Examples of trajectories and smoothing results are shown in Figure 2.

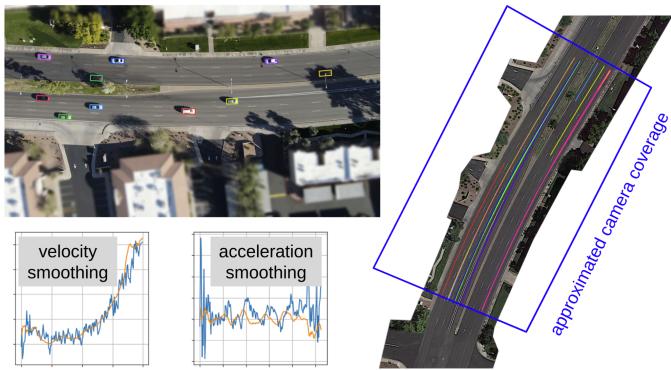


Figure 2. Example trajectories and smoothing results.

Specifically, there are three steps. First, given a vehicle trajectory consisting of a sequence of location points, for each location point, we take a segment of the trajectory centered at that location point. Since the trajectory is obtained from one video image, the length of the segment is set up to the frame-per-second of the video (i.e., 30 in our case) such that the vehicle will move for one second on that segment. Second, we assume that the vehicle will maintain its speed and acceleration during this one-second window, and we fit a quadratic curve on this segment to compute the speed and acceleration. The computed speed and acceleration are assigned only to the center point of the window. Third, we slide the window and repeat this step for each location point on the trajectory. After that, we save the smoothed speed and acceleration together with the trajectory so that they can be used for metrics calculation.

Although this method has a fundamental flaw without considering the dynamic properties of vehicles, e.g., the speed cannot be constant if the acceleration is nonzero in the window, in reality, it performs reasonably well with our data since the vehicles generally do not abruptly accelerate or decelerate within a short period in our traffic scenes. However, because of the sliding window, the computed speed may be skewed slightly when the vehicles start to move or brake hard in the intersection scenes, which may cause the calculated safety metrics to be biased. The extent of the influence requires further investigation. Meanwhile, we plan to develop better trajectory smoothing algorithms in future work.

Automated Leader-Follower Identification

When calculating safety metrics within an interaction scenario, it is essential to determine the vehicle pairs engaged in interactions that impact each other's decisions. To extract leader-follower pairs from vehicle trajectories, we developed a method to detect such pairs from the data and the map automatically. Different from our prior works [4][5] with one or a few limited scenes, we desire to develop a general method that can work for any scene. Our method consists of four steps, which are detailed as follows.

First, we manually segment the map into areas. Each area is assigned a unique area ID to facilitate the next step. An example of the semantic segmentation of the map is shown in Figure 3. These segments generally follow the lane markers and traffic rules. Our area resembles the "lanelet" [12] concept used in high-definition maps for road traffic simulation and autonomous driving algorithm development. However, we only need the areas without vectorized lane boundaries and further complicated annotation of traffic rules. Hence, the segmentation is conducted on the 2D satellite map instead of the 3D point cloud map, and segmenting a map costs about one hour, which is less expensive than constructing high-definition maps for autonomous driving. Moreover, for each scene, the map only needs to be segmented once, regardless of the amount of video data. Furthermore, the segmented map is also useful for other tasks, such as detecting safety-critical traffic incidents or traffic rule violations.

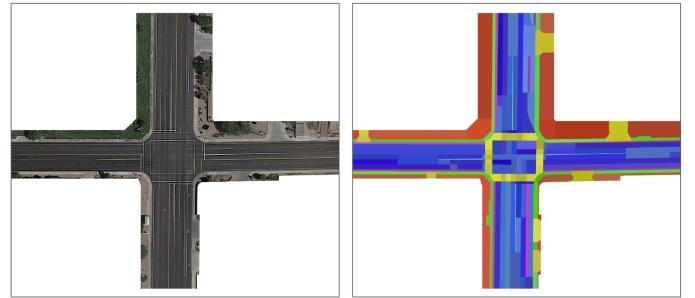


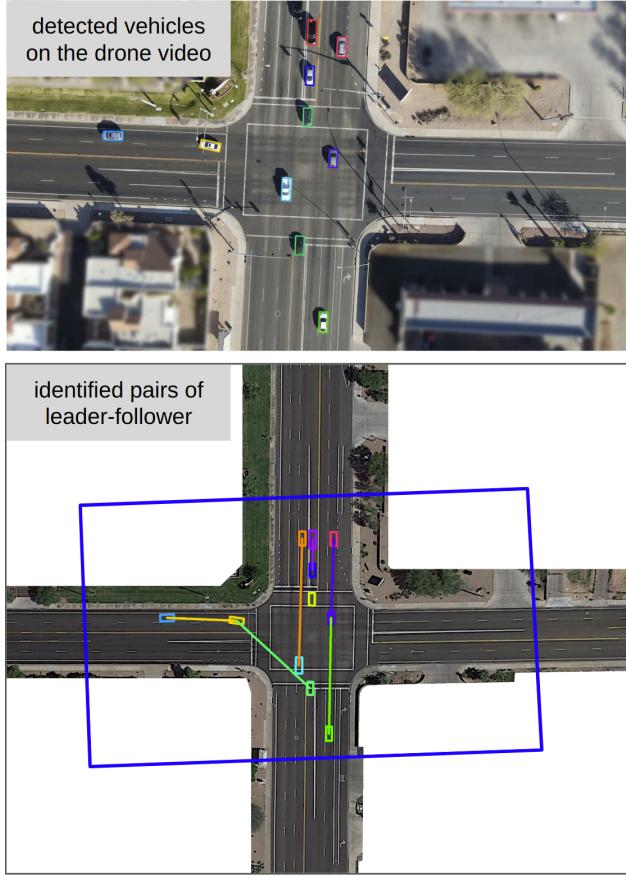
Figure 3. An example of the map of an intersection scene from a satellite image (left) and the semantic segmentation to lane areas (right).

Second, we convert each trajectory to a sequence of areas passed by that trajectory, which helps us combine those trajectories that pass the same area sequence into sets. If two vehicles traverse the same sequence of areas, potentially, they may form leader-follower pairs. Moreover, if two trajectories overlap for a subsequence of areas, these two vehicles may also form leader-follower pairs on the overlapped part. These overlapped trajectory segments are also combined into sets. These sets are constructed on the fly for every pair of vehicles that appear on the same video image. Since we convert each vehicle trajectory to a string of traversed area IDs, these sets can be efficiently computed using a substring algorithm.

Third, we use the trajectory sets and trajectory segment sets from the previous step to determine the "lane mates" for each vehicle in every video image. For a given vehicle A in an image, we identify another vehicle B on the same image that meets three conditions to be considered a "lane mate" of A. The first condition checks if A and B are in the same "lane", and there are two cases. On the one hand, if both vehicles' trajectories are in the same trajectory set, they satisfy this condition. On the other hand, if their trajectories are not in the same set, we check if they have an overlapping segment and if they are on the same overlapped trajectory segment. If both checks pass, the first condition is also satisfied. The second condition assesses

how much B deviates from A's trajectory. We find the closest point of B's current position on A's trajectory and consider it acceptable if the distance is less than a set threshold (2 meters in our implementation). The third condition is the reverse of the second: A should not deviate too much from B's trajectory, computed similarly.

Fourth, given any vehicle A on a video image, once we have all "lane mates" of vehicle A that satisfy all these three conditions, we project their current positions onto the trajectory of vehicle A using the closest point computed in the second condition of the previous step. This essentially allows us to convert the two-dimensional locations of other vehicles on a map to one-dimensional offsets on the trajectory of vehicle A. These offsets are relative to the current position of vehicle A, and hence, they can be sorted easily. Finally, the "lane mate" vehicle with the smallest positive offset is identified as the leader of vehicle A. An example of the identified leader-follower vehicle pairs is shown in [Figure 4](#).



[Figure 4. An example of identified leader-follower vehicle pairs.](#)

Pair Filtering

Despite the development of our automated leader-follower identification algorithm, it's important to acknowledge that complex vehicle configurations can pose unique challenges for automated identification algorithms, making data filtering and manual verification essential to maintain the precision of leader-follower pair identification. The data filtering is primarily required in two distinct cases. First, human intervention is crucial to correct any misidentifications in situations where a vehicle is not tracked accurately due to various factors, mainly failures in the neural

network vehicle detector and sensor limitations (e.g., overexposure). For this case, we watched every video and created a list of wrongly tracked vehicles, which is used to filter out vehicle pairs. Given our 3-hour videos at 12 scenes, approximately 6% of detected vehicles are filtered due to various tracking issues.

The second case that demands extra processing pertains to vehicles with trailers since we do not want to wrongly consider a truck with a trailer as a pair of leader-follower vehicles. For this case, we developed an algorithm to detect and filter trailers. There are three steps. First, we particularly annotated hundreds of video images with trailers to train the neural network vehicle detector so that it has the capability to generate a type code for each vehicle instance, including the trailer as one type. Second, for a pair of leader-followers identified in the previous step, if the follower vehicle is detected as a trailer and the distance between the pair of vehicles is less than a threshold, the pair is filtered. In our implementation, the threshold is set to the maximum of the leader and follower vehicle lengths. Third, we watched the video and manually verified the detected trailers. Surprisingly, all trailers on our videos are correctly detected. However, a few trucks were wrongly detected as trailers. Finally, the filtered data is used for metrics calculation and analysis.

Vehicle Pair Statistics

In this section, we delve into the statistical analysis of our identified vehicle leader-follower pairs dataset, with a particular focus on key properties such as vehicle velocity and headway distance between these pairs. Understanding the statistical characteristics of these pairs is paramount in gaining insights into driving behaviors and safety in car-following scenarios.

[Table 1. Statistics of vehicle pairs and the number of data samples.](#)

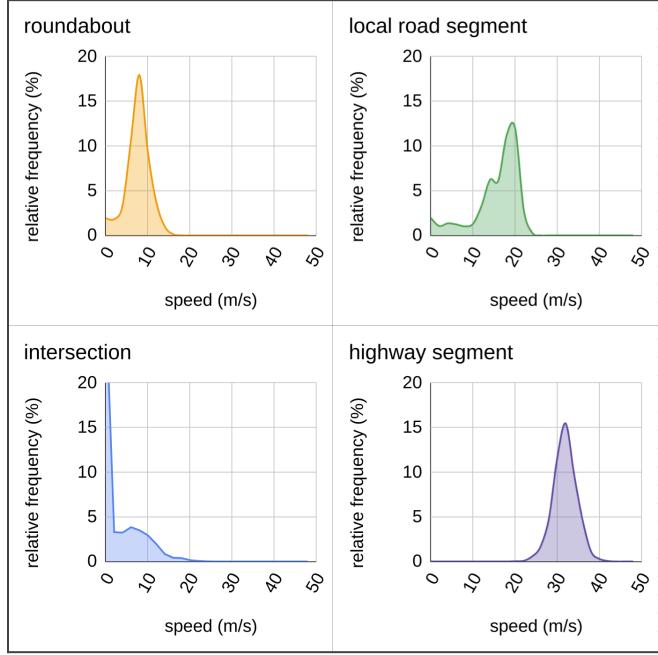
scenario category	# of vehicle pairs	# of data samples
roundabout	1,002	208,995
intersection	1,663	779,271
local road segment	795	99,975
highway segment	1,973	116,202
total	5,433	1,204,443

With our 3-hour drone video dataset, in total, 5,433 vehicle pairs are identified after filtering and manual verification. Most leader-follower pairs last for at least a few seconds when they pass the covered area of our drone camera. Hence, for each drone video image, we can obtain one data sample of the vehicle pair. In total, about 1.2 million data samples were obtained. [Table 1](#) lists the number of vehicle pairs and the number of data samples for each scenario category. Throughout the rest of this paper, we use the term "pair" to denote "a data sample of a leader-follower vehicle pair".

To balance our dataset, we intentionally selected the three roundabout sites with relatively heavy traffic and also deliberately recorded the video during rush hours. This enables us to obtain a sufficient number of vehicle pairs for driving safety metrics analysis. Normally, a roundabout will not be as busy as those in our dataset. For the same reason, we selected an intersection and two highway segments with moderate traffic load. Among all the four traffic scenario categories,

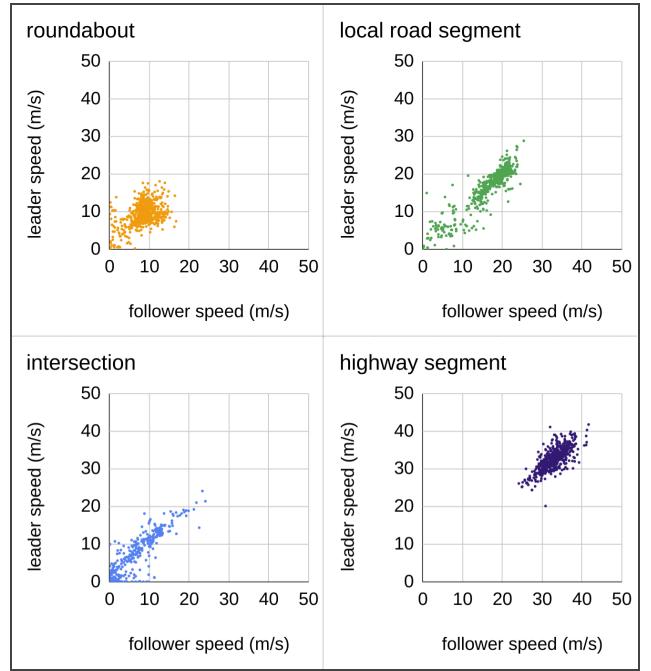
the local road segments have the least number of vehicle pairs, partially because the road segments are inherently not designed for heavy traffic load and partially because the headway distance between two vehicles is relatively large under light traffic load such that many pairs are not within the coverage of the drone camera. However, for the intersections, since many pairs of vehicles will stop to wait for the traffic light at the inbound buffer zone of intersections, they will be captured for long periods of time, with a significant number of data samples generated. Unfortunately, for most of the data samples under this situation, both the leader vehicle and the follower vehicle are stopped, and the data samples do not carry much information for metrics analysis.

In [Figure 5](#), we show vehicle speed distribution for the four scenario categories. Vehicle speeds exhibit notable variations across diverse scenarios, with higher velocities typically observed on highway segments, while lower speeds are common in the other three scenarios, reflecting the complex dynamics of each environment.



[Figure 5. Distribution of vehicle speed.](#)

We further show the speed distribution of leader-follower vehicle pairs in [Figure 6](#). To construct each subplot of this figure, we randomly sampled 200 pairs from each scene. In total, for each scenario category with three scenes, there are 600 sampled pairs. In the majority of cases, leader-follower vehicle pairs tend to exhibit similar speed changes, with the follower's speed closely tracking that of the leader; however, in specific scenarios like roundabouts and intersections, vehicles may slow down upon entering and accelerate upon exiting, resulting in a discernible lag between the follower's speed and the leader's speed, which can be observed in the lower left corner of the plot. The average absolute speed differences for the four scenario categories, i.e., roundabout, intersection, local road segment, and highway segment, are 2.45 m/s, 1.27 m/s, 1.47 m/s, and 1.48 m/s, respectively. 90% of vehicle pairs' speed difference is less than 5 m/s.



[Figure 6. Distribution of speed of leader-follower vehicle pairs.](#)

Finally, we show the distribution of the headway distance of leader-follower vehicle pairs in [Figure 7](#). The headway distance is measured as the difference of position offset along the trajectory of the follower vehicle instead of Euclidean distance in the 3D space, which is important for trajectories that are not straight. Headway distances between leader-follower vehicle pairs vary across scenarios, generally shorter in high-density environments like intersections and roundabouts but longer on local road segments and highways, reflecting the impact of traffic conditions on the spacing between vehicles. The average absolute speed differences for the four scenario categories, i.e., roundabout, intersection, local road segment, and highway segment, are 33 m, 15 m, 41 m, and 52 m, respectively. However, as the maximum coverage of our drone camera is about 145 m to 160 m on the diagonal line of the video image, depending on the drone's flying height (typically at 110 m to 120 m), any headway distance larger than the coverage will not be considered in our data analysis.

The headway distance between leader-follower vehicle pairs is influenced by a multitude of dynamic factors. These include the speed of the vehicles, traffic density, driver behavior, road geometry, and environmental conditions. In high-density traffic or congested environments, headway distances tend to be shorter due to the need for closer following to maintain traffic flow. Conversely, on open highways or roads with lower traffic volumes, headway distances often increase. Driver aggressiveness, visibility, and reaction times also play a role in shaping these distances, as do road features such as curves, intersections, and obstacles. Weather conditions such as rain or fog can further impact headway distances by necessitating greater spacing for safety. These complex interactions highlight the dynamic nature of headway distances in real-world driving scenarios, making their analysis a key element in understanding traffic behavior and enhancing road safety.

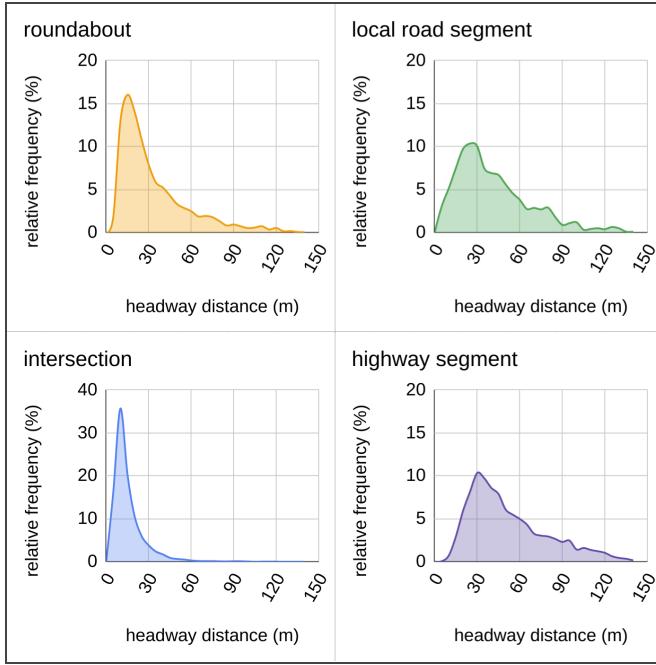


Figure 7. Distribution of headway distance of leader-follower vehicle pairs.

By examining the distribution and trends in vehicle velocity and headway distance, we can uncover valuable information about how vehicles interact and influence each other's movements on the road, shedding light on the dynamics of traffic flow and safety considerations.

Analysis of Longitudinal Safety Envelope

In this section, we compute and analyze the longitudinal safety metrics. The Minimum Distance Safety Envelope (MDSE) [1] is a pivotal metric for evaluating driving safety, particularly in the context of leader-follower vehicle pairs during sudden braking scenarios. This safety envelope encompasses three critical components: first, the distance traveled by the follower during their response time; second, the braking distance of the follower; and third, the braking distance of the leader. Formally, it is defined as follows.

$$MDSE = \left[v_F \rho + \frac{1}{2} a_F \rho^2 + \frac{(v_F + \rho a_F)^2}{2b_F} - \frac{(v_L)^2}{2b_L} \right]_+$$

Here is a list of notations in the equation.

- v_F and v_L are the longitudinal speeds of the follower vehicle and the leader vehicle, respectively. In our analysis, we assume that vehicles do not slip sideways, and hence, v_F and v_L are the vehicle speeds derived from the trajectory smoothing step.
- a_F is the longitudinal acceleration capability of the follower vehicle, which is set to 1.8 m/s^2 .
- b_F and b_L denote the longitudinal deceleration capability of the follower and leader vehicles, which are set to 3.6 m/s^2 and 6.1 m/s^2 .
- ρ is the response time of the follower vehicle, which is set to 0.2 seconds.

The values of a_F , b_F , b_L , and ρ are set up following our prior work [4], i.e., with the calibration of the naturalistic driving study (the NDS setup in [4]).

Furthermore, MDSE metrics are implemented to measure the occurrence of a behavior and/or event with a given spatio-temporal safety envelope formulation. These formulations are essential in helping identify potential hazardous driving conditions and play an important role in ensuring the safety and reliability of transportation networks, as they define the time and space the subject vehicle has for performing maneuvers and responding to actions of nearby objects, with the aim of reducing the risk of a collision occurring. The unique characteristics of various driving scenarios, including roundabouts, intersections, local road segments, and highway segments, introduce significant variations in the dimensions of this safety envelope. We show the distribution of computed MDSE in Figure 8 for the four scenario categories.

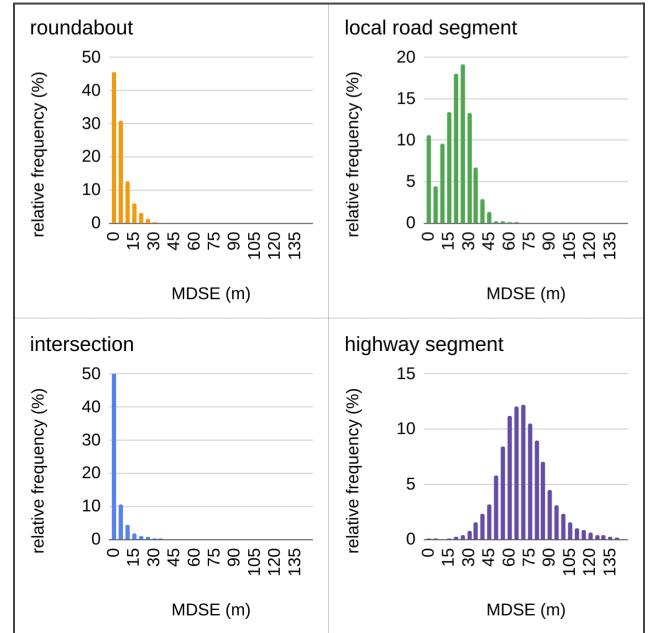


Figure 8. Distribution of Minimum Distance Safety Envelope (MDSE).

In environments like roundabouts and intersections, where vehicles may need to decelerate abruptly, the MDSE tends to be relatively compact, reflecting the need for quick and precise responses. This is further depicted in Figure 8. In road segments characterized by intersections and roundabouts, there is a notable bias among manual vehicle operators to exhibit a heightened frequency of speed reduction occurrences. In contrast, road segments such as highways and local roads tend to consistently maintain frequencies of elevated speeds that adhere to prescribed limit ranges. This reflects the nuanced driving behavior and speed dynamics of distinct road infrastructure. However, since the vehicle speed is relatively low in these two scenarios, as shown in Figure 5, vehicle pairs generally maintain their headway distance larger than the MDSE, which can be seen by comparing Figure 7 and Figure 8. Conversely, on local road segments and highways, where traffic flow is typically more stable, the safety envelope tends to be more extensive, allowing for greater reaction time and braking distance due to the larger vehicle speed. Understanding these variations is crucial for enhancing driving safety and developing effective safety measures tailored to specific scenarios.

An MDSE violation occurs when the computed MDSE is less than the instantaneous headway distance between the leader vehicle and the follower vehicle. To better capture the relation between MDSE and the headway distance, we compute the MDSE ratio [1], which is formally defined as the ratio of the headway distance to the calculated MDSE between the follower vehicle and the leader vehicle. Formally, it is defined as follows.

$$MDSE \text{ ratio} = \frac{\text{headway distance}}{MDSE}$$

With this definition, MDSE violations can be simply referenced by the vehicle pairs with an MDSE ratio of less than one.

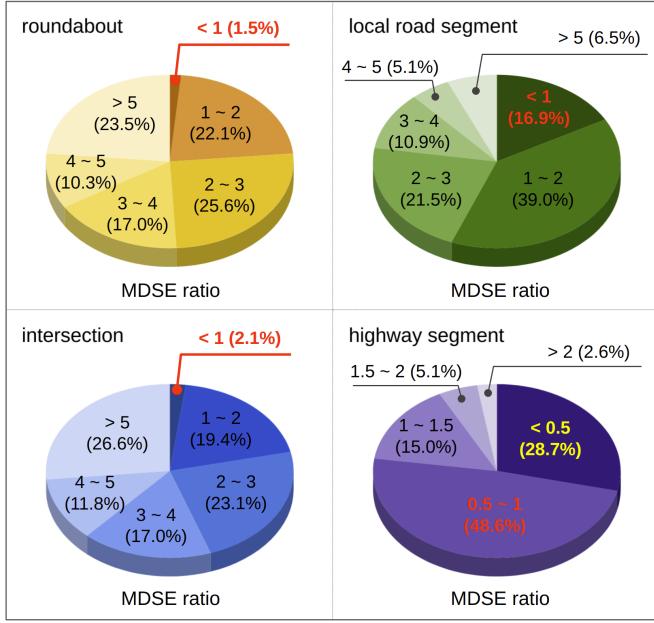


Figure 9. Distribution of MDSE ratio.

Figure 9 shows the distribution of the MDSE ratio for the four scenario categories. Particularly, the MDSE violation part is labeled in red and bright yellow. The substantial disparity in MDSE violations among different scenario categories is a notable observation. In particular, the remarkably high percentage of MDSE violations on highway segments, reaching 77.4%, is a cause for concern. On the one hand, this statistic suggests a heightened risk of safety breaches in scenarios characterized by high-speed travel and potentially reduced reaction times. On the other hand, we speculate that the 0.2-second response time is relatively short for high-speed scenarios since the general suggestion of headway time should be at least one to two seconds when driving on the highway [13] so as to leave enough response time. However, even if the response time is set to one second, still 38% of vehicle pairs have MDSE violations under the highway segment scenario. The elevated MDSE violations on highway segments emphasize the critical need for enhanced safety measures and increased awareness of the unique challenges presented by these scenarios. Conversely, the lower percentages of MDSE violations in roundabouts, intersections, and local road segments, at 1.5%, 2.1%, and 16.9%, respectively, indicate relatively more favorable safety conditions in these environments. These findings underscore the importance of tailoring safety interventions and policies to address the specific safety dynamics in different types of road scenarios, with a particular focus on mitigating the pronounced MDSE violations on highway segments to enhance driving safety.

Finally, we show the headway distance relative to the follower vehicle speed in Figure 10, with the MDSE violations highlighted in red. It is evident from the figure that the violations tend to occur when follower vehicles are traveling at relatively higher speeds while failing to maintain a sufficient headway distance from the leader vehicle. Particularly, the violations are more frequent when the follower vehicle's speed is larger than 15 m/s. This trend observed in this data aligns with established principles of safe driving, emphasizing the importance of maintaining a suitable buffer between vehicles, especially at higher speeds. In addition, the MDSE ratio reveals the extent to which followers exhibit elevated speeds and reduced headway distances, which further indicates that the MDSE ratio can serve as a clear reminder of the heightened risk associated with tailgating and insufficient spacing between vehicles, particularly when driving at higher speeds on the road, to mitigate the potential for rear-end collisions and enhance overall road safety.

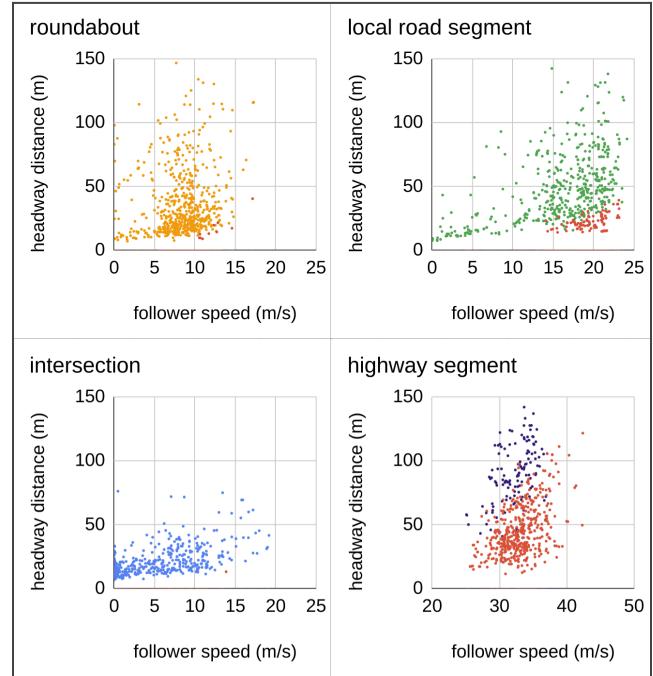


Figure 10. Distribution of headway distance with respect to the follower vehicle speed, with MDSE violations highlighted in red.

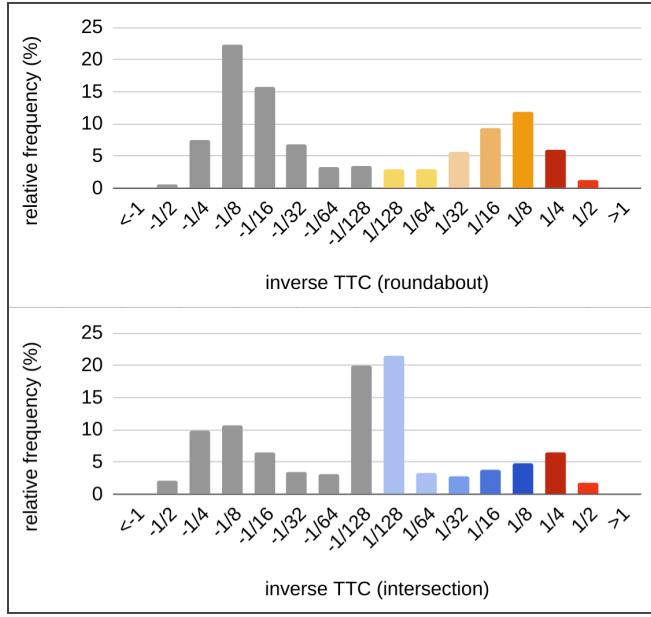
Analysis of Time-based Longitudinal Metrics

In this section, we examine two longitudinal safety metrics based on time, i.e., Time-To-Collision (TTC) [1] and Modified Time-To-Collision (MTTC) [1]. These two metrics are typically categorized as traditional Surrogate Safety Measures (SSMs), which offer an alternative to accident-based indicators. SSMs are valuable tools used in road safety analyses to quantitatively assess various hazardous traffic situations.

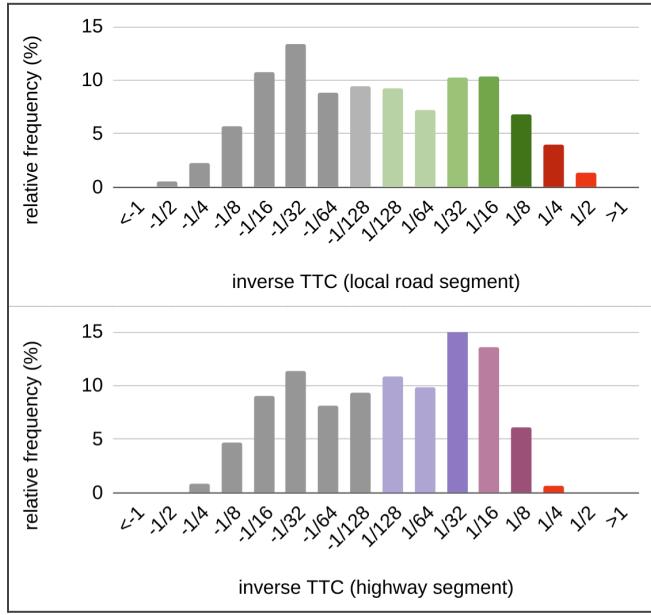
TTC is defined as the time of collision between two entities in a given scenario in which both entities continue with present velocities in the current environment in the same direction. Formally, it is defined as follows.

$$TTC = \frac{X_L - X_F}{v_F - v_L}$$

The variables needed to calculate TTC are relative positions (i.e., $X_L - X_F$), and speed between the two vehicles (i.e., $v_F - v_L$). Although it was originally defined only for vehicles traveling in a straight line, we extended it so that the relative positions X_L and X_F are the offsets on the trajectory of the follower vehicle, even when the trajectory is not straight but following a curved lane. Moreover, the TTC formulation only considers the speed at which the subject vehicles are traveling, implicitly assuming that the response time and acceleration of both objects are zero. Hence, it can have negative values if the speed of the leader vehicle is larger than the speed of the follower vehicles, and it can even be infinite if the two vehicles have the same speed. For the convenience of analysis, we show the histogram of the inverse of TTC instead of TTC in [Figure 11](#) and [Figure 12](#). The histogram bins are intentionally selected at log scale to better capture the portion of vehicle pairs with TTC in different ranges.

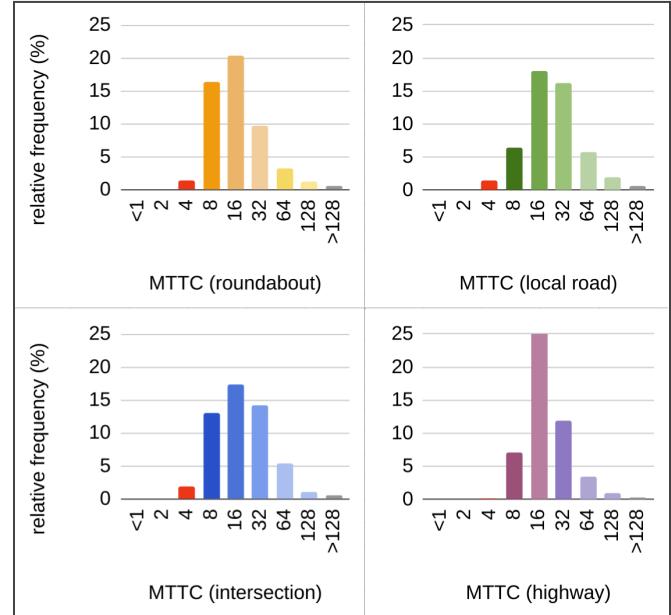


[Figure 11. Histogram of inverse TTC for roundabouts and intersections.](#)



[Figure 12. Histogram of inverse TTC for local road and highway segments.](#)

From these two figures, we can obtain a few insights. First, although the speed of the leader vehicle and the follower vehicle are generally in the same range with slight variance, as shown in [Figure 6](#), the TTC metric can capture the details of the variance. This also means that the vehicle speed measurements need to be precise. While this is generally not an issue for our drone videos, for a system on the roadside or on the road infrastructure, it might be challenging, and slight sensor noise may cause large fluctuation of TTC, especially when the speeds of the vehicle pairs are close. Second, the inverse TTC plots seem symmetric, which indicates that the speed variance of the vehicle pairs may not be able to capture the driver's intent. Hence, even statistical results of TTC can show how dangerous the traffic is, but it is difficult to interpret why it is dangerous with temporal analysis of the change of TTC. Third, given our dataset, for the four scenario categories, i.e., roundabout, intersection, local road segment, and highway segment, the percentage of vehicle pairs with a TTC less than four seconds is 5.8%, 6.4%, 1.0%, and 0.5%, respectively. These are also highlighted in red in [Figure 11](#) and [Figure 12](#). We further manually confirmed that for both roundabouts and intersections, the cause of these small TTC values is pairs of vehicles entering the roundabouts or the intersections, i.e., the delayed deceleration of the follower vehicle with respect to the leader vehicle. Although it seems like these small TTC values indicate higher chances of tailgate collision, it is counterintuitive that the low-speed scenarios generate more small TTC values than the high-speed scenarios. As a result, it might not be appropriate to directly compare the percentage of TTC violations or the extent of TTC violations among different scenarios. Instead, a calibration from a naturalistic driving dataset like ours might be needed to interpret TTC results.



[Figure 13. Histogram of MTTC for the four scenario categories.](#)

As TTC only considers velocity, MTTC introduces enhancements to traditional collision risk assessment by considering an additional factor, i.e., the relative acceleration. Formally, it is defined as follows.

$$MTTC = \frac{-\bar{\Delta V} \pm \sqrt{\bar{V}^2 + 2\Delta AD}}{\Delta A}$$

Here is a list of notations in the equation.

- $\bar{\Delta V}$ is the relative speed, i.e., $v_F - v_L$.
- D is the headway distance, i.e., $X_L - X_F$.
- $\bar{\Delta A}$ is the relative acceleration, i.e., $a_F - a_L$, which is not considered in TTC.

The modification by considering relative acceleration provides a more comprehensive evaluation of safety in traffic situations. In [Figure 13](#), we show the histogram of calculated MTTC for the four scenario categories. Given our dataset, for the four scenario categories, i.e., roundabout, intersection, local road segment, and highway segment, the percentage of vehicle pairs with an MTTC less than four seconds is 1.4%, 1.9%, 1.5%, and 0.1%, respectively. As the results show, in scenarios characterized by abrupt speed changes or complex driver behavior, MTTC offers a more nuanced perspective on collision risk, making it a valuable tool for capturing the intricacies of real-world traffic dynamics. However, the MTTC metric shares the same properties as the TTC metric. Particularly, the usage of acceleration makes it even more challenging for roadside or road infrastructure-based systems to calculate it in real time with enough accuracy.

Discussions

The utilization of artificial intelligence has led to increased popularity of traffic monitoring video-based road traffic safety analysis within academic and research communities [\[14\]](#) [\[15\]](#). In [\[5\]](#), using the infrastructure-based cameras and the localization method mentioned in [\[9\]](#), kinematics information for car-following scenarios was extracted where a pair of vehicles approach or leave a traffic intersection. Since the cameras are ground-based, it is not always possible to extract accurate localization information as these data could have scenarios where one or both vehicles could be occluded. In this work, the data is extracted from drone videos, which provides an advantage over ground-based data in such scenarios. Additionally, in [\[4\]](#) and [\[5\]](#), time-based and distance-based metrics were measured for car-following scenarios in simulation and for real-world data, respectively. Both of these works emphasize metric violations and their durations to study the robustness and relevance of the proposed metrics. In this work, the emphasis is on analyzing the magnitude distribution of the safety metrics for different scenario categories.

Besides cameras on the ground or road infrastructure, naturalistic driving data can also be obtained from onboard sensors, such as the VTTI dataset [\[16\]](#). However, to study driving safety metrics defined on pairs of leader-follower vehicles, such a naturalistic driving dataset may not be the best option due to the limitation of the data for vehicle pairs. Instead, a drone dataset can provide more samples of vehicle pairs with better diversity. There are also drone datasets available for academic research, notably the LeveXData [\[17\]](#). Compared to these drone datasets, we are interested in sampling data at specific traffic scenes designated by us and comparing the metrics under different scenes in the same city. Hence, we believe it is more valuable to develop a flexible data acquisition method than using existing open datasets.

Conclusions and Future Work

In this study, we introduce a novel dataset and conduct an in-depth analysis of driving safety metrics specifically designed for car-following scenarios. We leverage cutting-edge technology, utilizing drones to capture high-resolution video data from 12 distinct

traffic scenes in the Phoenix metropolitan area. Advanced Artificial Intelligence (AI) algorithms are employed to extract precise vehicle trajectories, and semantic maps are used to identify leader-follower relationships among vehicles. By incorporating a set of metrics based on prior work on Operational Safety Assessment (OSA) metrics by the Institute of Automated Mobility (IAM), we examine and analyze three driving safety metrics (i.e., MDSE, TTC, and MTTC) in real-world traffic scenes. Our data uncovers the distribution of these metrics and compares them in different scenarios, which provides insights into the impact of various parameters and thresholds on these metrics. From the results obtained for our car-following scenarios, we draw the following conclusions.

- Time-based metrics like TTC or MTTC rely on accurate measurement of velocities and accelerations of the vehicles, which presents challenges for traffic operators with equipment mounted on the roadside or on the road infrastructure. Instead, they are potentially useful for self-driving vehicles with onboard sensors that directly measure the difference of velocity and acceleration.
- The Minimum Distance Safety Envelope (MDSE) considers velocity, the capability of acceleration, and the capability of deceleration, response time, and headway distance, which is robust and interpretable. Instead of using a threshold for violation detection, we prefer the MDSE ratio to capture the severity better. Additionally, we recommend using different response time and acceleration capability parameters under different traffic scenarios (e.g., shorter response time on local roads vs. longer response time on highways).

We believe that our data-driven approach has the potential to empower traffic operators and policymakers and equip them with valuable information to make informed decisions and enhance the safety and efficiency of future road transportation systems. However, there are a few notable limitations in our method. First, our dataset does not contain safety-critical incidents such as vehicle tailgate collisions (for general car-following scenarios) or T-bone collisions (for roundabouts). Hence, it complements prior metrics studies such as [\[4\]](#) but lacks flexibility. In the future, we plan to develop a "replay-and-simulation" program with our dataset so that researchers can replay our data in simulation while being able to instrument vehicles and change their behavior. Particularly, we want to answer questions like "How safe is it if the vehicle is 10% faster?". Second, due to regulations, the maximum height of our drone is about 120 meters, which leads to limited coverage. This is especially restrictive when studying high-speed scenarios since the headway distance can easily exceed the coverage of our drone. In the future, we plan to fly two drones together and pitch the camera view angle slightly to cover a larger area. Third, as a research project, we have limited resources to annotate and train our neural network vehicle detector, and hence, the number of tracking errors is not negligible (about 6% of all vehicle instances), which requires human intervention. In the future, we plan to develop better AI models and algorithms to reduce the amount of human effort in data processing. Fourth, the current driving safety metrics are mainly defined on car-following scenarios, which ignores many factors that need to be considered, such as lateral safety margin, pedestrians, and rate of turning. In the future, we plan to develop different fine-grained metrics for different scenarios and better quantify the severity of violation.

References

1. Wishart, J., Como, S., Elli, M., Russo, B. et al., "Driving Safety Performance Assessment Metrics for ADS-Equipped Vehicles," SAE Technical Paper 2020-01-1206, 2020.
2. Automated Vehicle Safety Consortium(AVSC), "AVSC Best Practice for Metrics and Methods for Assessing Safety Performance of Automated Driving Systems (ADS)," SAE Industry Technologies Consortia, 2021.
3. Como, S., Wishart, J., "Evaluating Automated Vehicle Scenario Navigation Using the Operational Safety Assessment (OSA) Methodology," SAE Technical Paper 2023-01-0797, 2023.
4. Elli, M., Wishart, J., Como, S., Dhakshinamoorthy, S. et al., "Evaluation of Operational Safety Assessment (OSA) Metrics for Automated Vehicles in Simulation," SAE Technical Paper 2021-01-0868, 2021.
5. Jammula, V.C., Wishart, J., and Yang, Y., "Evaluation of Operational Safety Assessment (OSA) Metrics for Automated Vehicles Using Real-World Data," SAE Technical Paper 2022-01-0062, 2022.
6. Srinivasan, A., Mahartayasa, Y., Jammula, V.C., Lu, D., et al., "Infrastructure-based LiDAR Monitoring for Assessing Automated Driving Safety," SAE Technical Paper 2022-01-0081, 2022.
7. Wassim, N., Ranganathan, R., Srinivasan, G., Smith, J., et al., "Description of Light-Vehicle to-Vehicle Communications for Safety Applications Based on Vehicle-to-Vehicle Communications," National Highway Traffic Safety Administration, Report No. DOT HS 811 731, 2013.
8. Altekar, N., Como, S., Lu, D., Wishart, J., et al., "Infrastructure-based Sensor Data Capture Systems for Measurement of Operational Safety Assessment (OSA) Metrics," SAE Technical Paper 2021-01-0175, 2021.
9. Lu, D., Jammula, V.C., Como, S., Wishart, J., et al., "CAROM - Vehicle Localization and Traffic Scene Reconstruction from Monocular Cameras on Road Infrastructures." In 2021 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2021.
10. Lu, D., Eaton, E., Weg, M., Wang, W., et al., "CAROM Air - Vehicle Localization and Traffic Scene Reconstruction from Aerial Videos," In 2023 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2023.
11. He, Kaiming, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. "Mask R-CNN." In Proceedings of the IEEE International Conference on Computer Vision (ICCV), IEEE, 2017.
12. Poggenhans, F., Pauls, J.H., Janosovits, J., Orf, S., et al., "Lanelet2: A High-Definition Map Framework for the Future of Automated Driving." In 2018 21st International Conference on Intelligent Transportation Systems (ITSC), IEEE, 2018.
13. Ayres, T.J., Li, L., Schleuning, D. and Young, D., "Preferred Time-Headway of Highway Drivers." In Proceedings of 2001 IEEE Intelligent Transportation Systems, IEEE, 2001.
14. Sivaraman, S. and Trivedi, M.M., "Looking at Vehicles on the Road: A Survey of Vision-based Vehicle Detection, Tracking, and Behavior Analysis." IEEE Transactions on Intelligent Transportation Systems, IEEE, 2013.
15. Ali, G., "Characterizing Human Driving Behavior Through an Analysis of Naturalistic Driving Data." PhD diss., Virginia Tech, 2023.
16. Virginia Tech Transportation Institute, "VTTI Data", online: <https://www.vtti.vt.edu/PDFs/publications/VTTI-Data.pdf>, VTTI.
17. Krajewski, R., Bock, J., Kloeker, L. and Eckstein, L., "The highD dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems." In 2018 21st International Conference on Intelligent Transportation Systems (ITSC), IEEE, 2018.

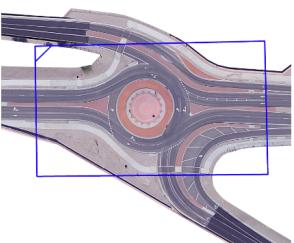
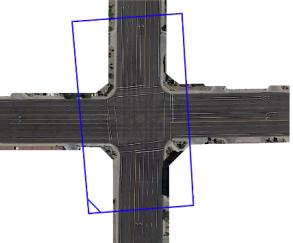
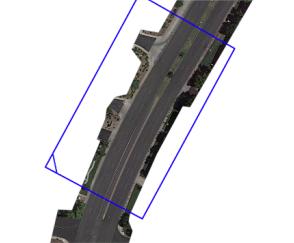
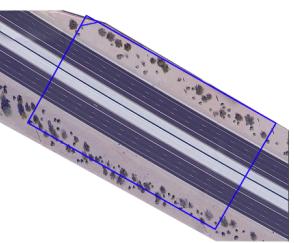
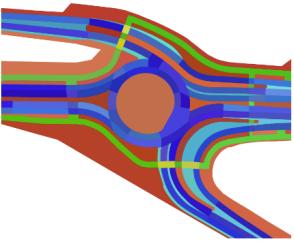
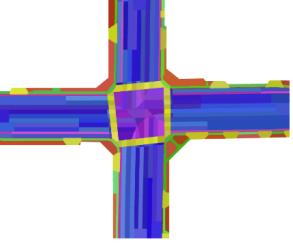
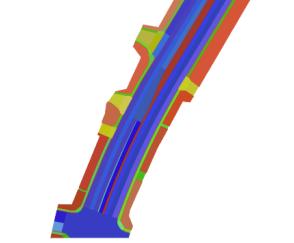
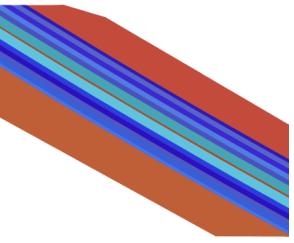
Acknowledgments

This work was sponsored by the Arizona Institute for Automated Mobility (IAM), Rider University, and Arizona State University. The authors are grateful to the Institute of Automated Mobility (IAM) for generously funding this work and the Maricopa County Department of Transportation (MCDOT) for enabling the infrastructure and structural resources required for this work. Duo Lu and Yezhou Yang acknowledge the support from the National Science Foundation (grant 2329780). Hongbin Yu and Prabin Kumar Rath acknowledge the support from the National Science Foundation (grants 1624842, 1839485, 1933742, 2137295), the Industry University Cooperative Research Center (IUCRC) for Efficient Vehicles and Sustainable Transportation Systems (EVSTS) at Arizona State University.

Appendix A: A List of Traffic Scenes in this Study

	Roundabout A2	Intersection B2	Local Road C1	Highway D0
Drone Image				
Satellite Map				
Lane Segments				

	Roundabout A4	Intersection B3	Local Road C2	Highway D2
Drone Image				
Satellite Map				
Lane Segments				

	Roundabout A5	Intersection B5	Local Road C3	Highway D5
Drone Image				
Satellite Map				
Lane Segments				

In this appendix, we list all 12 traffic scenes in this study. They are grouped into four main categories, as in the four columns of the table. The names of the scenes follow those in the CAROM Air dataset [10]. For each scene, we show the drone image, the satellite map, and the lane segments. The satellite maps are obtained from screenshots of Microsoft Bing Maps (for scene D0) or Google Maps (for all other scenes), with manual editing by removing any vehicles and shadows. The field-of-view of each drone video is also shown on the satellite map as a blue quadrilateral, and the blue triangle of the quadrilateral indicates the top-left corner of the drone video. The lane segments are color-coded in different semantic categories, such as curb area (red), sidewalk (green), crosswalk (yellow), lane (blue and purple), and lane buffer space (cyan).