Asian Transport Studies 10 (2024) 100138

Contents lists available at ScienceDirect

Asian Transport Studies

journal homepage: www.sciencedirect.com/journal/asian-transport-studies

Traffic flow characteristics and traffic conflict analysis in the downstream area of expressway toll station based on vehicle trajectory data

Qiaoqiao Ren [a], Jie He [b], Ziyang Liu [b], Min Xu [a,*]

[a] Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, 999077, Hong Kong, China
[b] School of Transportation, Southeast University, 2 Si Pai Lou, Nanjing, Jiangsu, 210096, China

ARTICLE INFO

Keywords:
Downstream area of toll station
Vehicle trajectory data extraction
Traffic flow characteristics
Traffic conflicts
Traffic safety

ABSTRACT

The limited road space and overlapping driving decisions may cause frequent speed and lane changes in the downstream area of expressway toll stations, which easily cause high accident risks. In this paper, micro vehicle trajectory data in the Huai'an South Toll Station was automatically extracted from aerial videos using the novel YSKT algorithm framework to analyze traffic flow characteristics. To further evaluate vehicle collision risks, the Extended Time-To-Collision (ETTC) and Post Encroachment Time (PET) indicators were employed. The results demonstrated that the longitudinal and lateral velocity and acceleration, the distribution of lane-changing points, lane-changing times, and travel time varied with road sections and vehicle types. Notably, the ETC vehicles had a higher risk of traffic accidents compared to MTC vehicles since their greater initial speed. The findings could provide valuable references for traffic managers to comprehensively understand traffic characteristics and evaluate traffic safety of similar complex road nodes.

1. Introduction

The mileage of expressways has increased rapidly in China since the implementation of the policy of "building roads with loans and paying back with fees". In 2021, the Ministry of Transport of China issued the statistical bulletin on the development of the transportation industry, which pointed out that the total mileage of expressways in China has reached 161,000 km and ranks first in the world (Ministry of Transport of China, 2021). As the key nodes of expressways, the toll stations directly affect the overall service level, traffic capacity, and road safety. Many studies have confirmed that toll stations will increase the risk of accidents on highways (Saad et al., 2019; Xing et al., 2020).

The downstream area of the expressway toll station is the exit where vehicles drive out of expressways and then drive into the next road section (Chen et al., 2013). The movement of vehicles in this confluence area is a complex process. Specifically, vehicles driving out of different toll lanes compete for road space and change speed to enter the target lanes, resulting in interleaving in the transition section (Xing et al., 2020). Subsequently, the traffic flow on the connecting section will be disturbed again due to the influence of lane-changing signs or intersection signal controls (Vahidi and Sciarretta, 2018). At present, the most typical form of toll stations in China is the hybrid toll station with the coexistence of Electronic Toll Collection (ETC) lanes and Manual Toll Collection (MTC) lanes (Abuzwidah and Abdel-Aty, 2018). The layout of toll channels and the vehicle speed limit may differ greatly, hence resulting in different initial motion states of ETC vehicles and MTC vehicles after completing payments. Consequently, the limited road space and the overlapping driving decisions of drivers will cause frequent speed and lane changes in the downstream area of the toll station (Abuzwidah, 2019; Saad et al., 2019). It will cause the disorder of traffic flow and the reduction of traffic capacity, further resulting in high accident risks (Li et al., 2016).

Traditional traffic safety studies have predominantly relied on crash records, which only capture incidents after they occur and may not reflect the full spectrum of risky traffic dynamics. With the latest development of video surveillance and image recognition technology, vehicle trajectory data has been widely used in road safety (Jackson et al., 2013; Oh and Kim, 2010; Wu et al., 2018). By utilizing UAV technology to collect high-resolution vehicle trajectory data, a proactive approach can be offered to identifying and analyzing potential safety issues before they escalate into accidents in this area. UAVs can collect detailed micro-dynamics information for multiple vehicles simultaneously, such as speed and acceleration. Moreover, the flexibility of UAV operations, coupled with their minimal demand on human and material

* Corresponding author.
E-mail addresses: 22037469r@connect.polyu.hk (Q. Ren), hejie@seu.edu.cn (J. He), lzyseu6257@outlook.com (Z. Liu), min.m.xu@polyu.edu.hk (M. Xu).

https://doi.org/10.1016/j.eastsj.2024.100138
Received 3 August 2023; Received in revised form 25 April 2024; Accepted 13 May 2024
Available online 17 May 2024

resources, allows for the rapid acquisition of extensive vehicle motion trajectory data and characteristics (Xing et al., 2019). Therefore, this research employs UAV video data collection in the downstream area of toll stations. For example, Li et al. (2017) studied the influence of acceleration control parameter setting on expressway rear-end collision by using vehicle micro trajectory data. Xing et al. (2019) extracted the vehicle trajectory data of typical toll plaza in Nanjing based on OpenCV automatic analysis system and then compared the different effects of payment methods and target lanes on the collision risk. Chen et al. (2021) used the NGSIM vehicle trajectory data set to study the influencing factors of potential collision risk in the lane-changing process considering vehicle groups and discussed the unobserved heterogeneity of individual lane-changing strategies.

For the study on traffic characteristics and traffic safety of toll stations, numerous research explored the traffic features of hybrid toll stations by constructing micro traffic simulation models. For example, Hajiseyedjavadi et al. (2015) used VISSIM for micro traffic flow simulation and further evaluated the impact of lane configurations on the traffic safety of toll stations. Regarding the traffic safety research of toll stations based on accident data, Abuzwidah and Abdel-Aty (2018) collected the accident data of multiple toll stations to explore the impact of various factors on the safety of toll stations, including the type, the gradient rate, and the longitudinal slope of the entrance section.

Considering the merging and interweaving characteristics, the downstream area of the toll station is regarded as the bottleneck of the expressway and could significantly influence the traffic capacity and safety (Jiang et al., 2021). Realizing the organic connection at this area is of great value to improve the expressways' transportation efficiency and road service levels (Chamoso et al., 2018). However, to the best of our knowledge, the historical record data in the downstream area of the expressway toll station is lacking in previous studies, which makes it difficult to analyze the characteristics of traffic flow and traffic conflicts. Therefore, it is necessary to collect dynamic data of vehicles to fill the gap. Analyzing the traffic operation status and safety features of the downstream area could provide references for proposing effective management strategy. Therefore, the objective of this paper is to analyze the traffic characteristics and traffic safety in the downstream area of the expressway toll station based on video trajectory data extraction technology.

2. Data

2.1. Data collection

The Huai'an South Toll Station, located in Jiangsu, China, is selected as the research area. The geographic location and satellite map of this toll station are shown in Fig. 1. This toll station is a mixed mainline toll station with 8 toll lanes, including 3 ETC lanes and 5 MTC lanes, and the detailed layout and labels are shown in Fig. 2. The DJI Mavic 2 Pro Unmanned Aerial Vehicle (UAV) was used to collect the vehicle trajectory video data under fair weather during 14:30–14:42, December 3rd, 2020. This specific timing was chosen for the following reasons: (1) Mitigation of peak hour queueing influence: To minimize the influence of vehicle queuing during peak traffic periods on the diverging behaviour of vehicles in the downstream area, this study specifically targeted the off-peak hours for video recording. By doing so, we aimed to capture the traffic flow under conditions that are not skewed by the extreme congestion and potential atypical driving patterns that often occur during peak hours. Besides, the moderate traffic volume during the selected timeframe is more representative of the typical driving behaviour and allows for a clearer analysis of the natural traffic flow characteristics at the toll station. (2) Lighting and weather conditions: The selected timeframe benefits from optimal lighting conditions for aerial video recording, ensuring high-quality data capture. Additionally, conducting the survey under fair weather conditions minimizes the impact of environmental factors on vehicle behaviour and traffic flow, allowing for a more accurate analysis of the inherent characteristics of the toll station's traffic dynamics.

The collected video provides 4 K ultra high definition and 30 frames per second (fps) data vertically above the downstream area. The micro

Fig. 1. Satellite map of Huai'an South Toll Station.

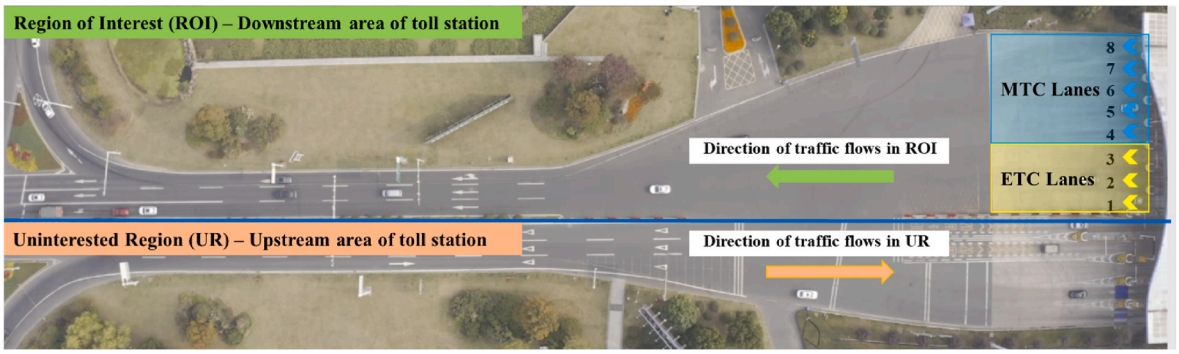


Fig. 2. The layout of downstream area of Huai'an South Toll Station.

vehicle trajectory contains continuous vehicle motion information (Li et al., 2021), which could intuitively reflect the actual vehicle movement. As a result, the on-site counted traffic flow volume of 862 vehicles/h was used to explore the impact of traffic volume on traffic safety, which corresponds to the moderate traffic flow volume and congestion conditions.

2.2. Spatial division of data in ROI

Given that different regions with different geometric characteristics or specific road configurations could potentially affect driver behavior, vehicle dynamics, and the likelihood of accidents (Xing et al., 2019, 2020), the downstream area is divided into two major parts to capture variations in traffic characteristics and conflict patterns in the region of interest (ROI): the downstream confluence area (CA) of the toll station and the connecting section (CS) between the toll station and the signalized intersection, as shown in Fig. 3.

Considering different geometric characteristics, the CA of the toll station further includes an unmarked straight confluence area (USCA) and an unmarked gradient confluence area (UGCA). The UGCA is the contraction gradient change section behind the toll station. There are no markings between two lanes and vehicles can merge freely. Likewise, the CS between the toll station and signalized intersection can be divided into a marked straight connecting section before the indicator (MSCS-BF) and a marked straight connecting section behind the indicator (MSCS-BH) based on the location of the lane guidance sign (LGS). In the MSCS-BF, vehicles would start changing lanes to meet the driving requirements according to the ground marking and lane indication sign. When entering the MSCS-BH, vehicle drivers will complete the lane-changing to meet the safety driving demands and then adjust the position flexibly according to real-time traffic flow situations, such as the queue length and surrounding vehicle speed.

3. Methodology

The methodology used in this study includes the micro vehicle trajectory data extraction framework and surrogate indicators for safety evaluation, the detailed introduction is as follows.

3.1. Micro vehicle trajectory data extraction framework

A novel vehicle trajectory extraction framework in Fig. 4 proposed by Liu et al. (2022), YSKT algorithm framework, was utilized in this research to extract complete and continuous vehicle micro trajectory data from the UAV traffic video. Specifically, the YOLOv4 algorithm was used to detect different types of vehicles, the simple online and real-time tracking (SORT) algorithm was used to realize the vehicle tracking, and then the initial trajectory data was cleaned, transformed, and reconstructed to generate the final data set using the KD-Tree trajectory data reconstruction algorithm. Readers may refer to Liu et al. (2022) for detail.

In our study, the location and type of each vehicle in the frame photos were marked, then the data set was split as training set (80%) and test set (20%). Based on the weight file trained by the official COCO data set, the weight file adapted to the vehicle characteristics of the traffic video collected in this paper was trained by the transfer learning method based on the YOLOv4 framework (Umair et al., 2021). After that, each vehicle's position, type, and confidence level in each frame of the collected traffic video can be output.

The SORT algorithm was then used to find the optimal vehicle matching in consecutive frames and extract the initial trajectory data of each vehicle. In this paper, vehicle tracking based on SORT algorithm can be divided into two processes (Hou et al., 2019; Hua and Anastasiu, 2019): (a) The Hungarian algorithm was used to find the optimal vehicle matchings among the vehicle detection results in the first and second frames to initialize the Kalman filter; (b) Using the Kalman filter to predict the vehicle position in the third frame, the Hungarian algorithm

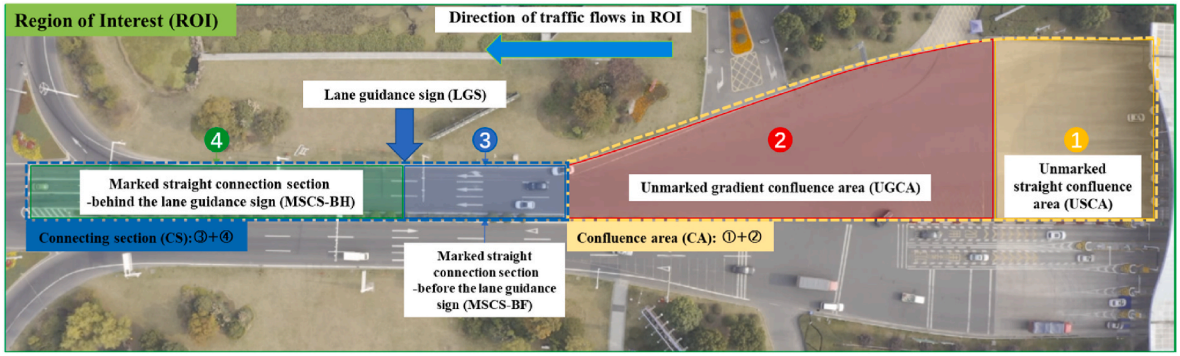


Fig. 3. Division of downstream area of the toll station.

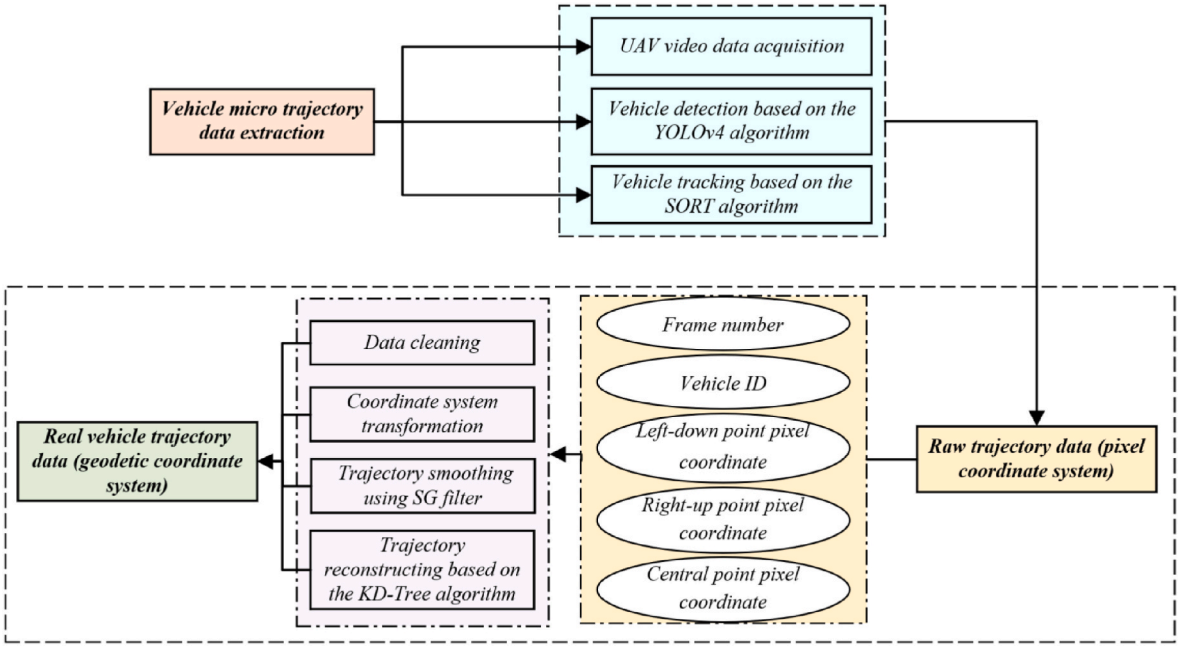


Fig. 4. Micro vehicle trajectory data extraction process.

could be used to find the optimal location matching pairs among the prediction results and detection results in the third frame, then the parameters in the Kalman filters would be updated. Repeating this step in later frames, the initial vehicle trajectories data could be obtained.

The initial vehicle trajectory data set includes the information of frame number, vehicle ID, the left-down point pixel coordinate, the right-up point pixel coordinate, and central point pixel coordinate of the tracking bounding box. The vehicle coordinates obtained by the video recognition system are pixel coordinates, which are inconsistent with the geodetic coordinate system in the road. Hence, fixed points should be selected from the two coordinate systems respectively to calibrate the coordinate system and integrate the pixel coordinates with the geodetic coordinate system. As shown in Fig. 5, the point on the first lane marking on the straight connecting section was taken as the origin *O* of the geodetic coordinate system, the extension line of the lane marking was taken as the *x* axis, and the connecting line of the endpoints of the two adjacent lane lines was taken as the *y* axis. Four points A, B, C, D in Fig. 5 were selected as the reference points. The geodetic coordinates of them were A (0, 0), B (6,300, 0), C (6,300, 3500), D (0, 3500), while the corresponding pixel coordinates were A' (849, 313), B' (892, 313), C' (892, 289), D' (849, 289), respectively. In order to deal with the slight fluctuation of trajectory data, the transmission coordinates of the same vehicle ID in two adjacent frames were compared, and then the trajectory data was smoothed. In this study, the Savitzky-Golay filter based on the least square principle was employed to smooth the center point coordinates twice (Seleznov et al., 2020).

The KD-Tree (K-Dimensional Tree) algorithm and the cubic spline interpolation method were finally used to reconstruct trajectory data (Feng et al., 2021). The KD-Tree algorithm was employed for searching the nearest neighbor of key data in multi-dimensional space and the subsequent broken track data with the highest feature similarity with the previous broken trajectory data. The successfully matched broken trajectory data were connected by cubic spline interpolation to obtain new complete trajectory data.

In terms of vehicle type distribution, cars, trucks, and buses constitute 78.6%, 9.6%, and 11.8% of the average traffic volume, respectively. As for the classification based on toll collection method, ETC and MTC vehicles account for 68.14% and 31.86% of the total traffic volume, respectively. The detailed distribution of trajectory data points can be seen in Table 1. A total of 482,005 raw data points were generated in the initial vehicle trajectory extraction, and the data points used for the final analysis in the ROI were 189,810. The reasons of excluding in excess of 290,000 data points include: (1) A total of 274,602 raw data points in uninterested region (UR) were excluded after the spatial division and data verification; and (2) A total of 207,403–189,810 = 17,593 raw data points in ROI were excluded after the trajectory reconstruction and smoothing. In this regard, 113 complete vehicle trajectories were reconstructed. Then 50 complete vehicle tracks were randomly selected

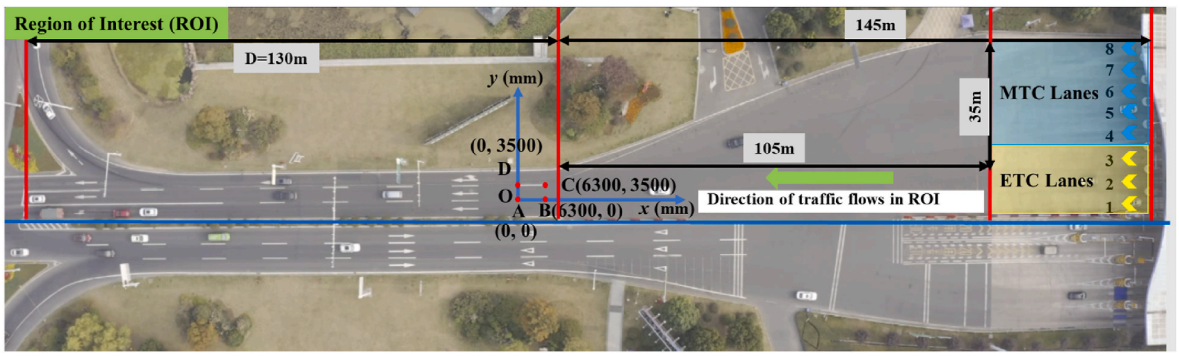


Fig. 5. Schematic diagram of *Oxy* geodetic coordinate system.

Table 1
The detailed distribution of trajectory data points.

Time	Total number of collected raw data points	Total number of raw data points in UR	Total number of raw data points in ROI	Number of data points in complete trajectories in ROI
14:30–14:35	199,668	122,462	77,206	70,206
14:36–14:42	282,337	152,140	130,197	119,604
14:30–14:42	482,005	274,602	207,403	189,810

for visualization, as shown in Fig. 6, in which the x-y coordinates were the converted trajectory geodetic coordinates.

3.2. Extended Time-To-Collision (ETTC) and Post Encroachment Time (PET)

As a widely recognized traffic safety measure of non-accident statistics, surrogate safety indicators can identify potential traffic accidents and effectively improve the reliability of the traditional accident-based analysis. In this paper, two typical surrogate safety analysis parameters: Extended Time-to-Collision (ETTC) (Xing et al., 2019) and Post Encroachment Time (PET) (Cooper, 1984) were employed to evaluate the potential traffic conflict risks. The ETTC and PET with trajectory data can be calculated by Equations (1) and (2), respectively, as follows:

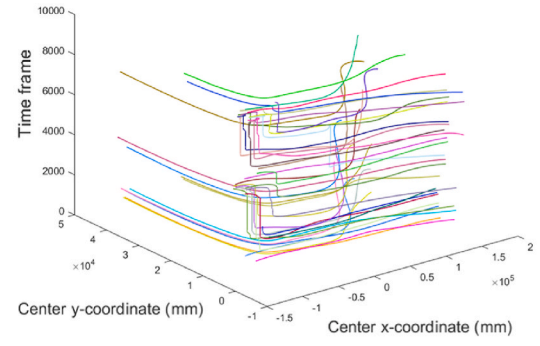
$$ETTC = -\frac{\sqrt{(\boldsymbol{O_i}-\boldsymbol{O_j})^T(\boldsymbol{O_i}-\boldsymbol{O_j})} - 0.5L_i - 0.5L_j}{\frac{1}{\sqrt{(\boldsymbol{o_i}-\boldsymbol{o_j})^T(\boldsymbol{o_i}-\boldsymbol{o_j})}}(\boldsymbol{O_i}-\boldsymbol{O_j})^T(V_i - V_j)} \quad (1)$$

$$PET = \frac{\sqrt{(\boldsymbol{O_i}-\boldsymbol{O_j})^T(\boldsymbol{O_i}-\boldsymbol{O_j})} - 0.5L_i - 0.5L_j}{V_j} \quad (2)$$

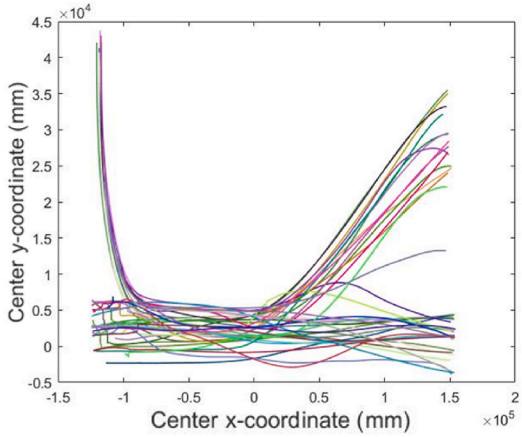
where $\boldsymbol{O_i}$ and $\boldsymbol{O_j}$ denote center coordinates of the front vehicle and the following vehicle; L_i and L_j are the lengths of the front vehicle and the following vehicle, which could be obtained with the angular point coordinates of tracking bounding boxes; and V_i and V_j denote velocities of the front vehicle and the following vehicle. According to the definition, a larger ETTC or PET implies that drivers have more time to avoid collision, which results in smaller probability of traffic collisions (Li et al., 2016).

4. Results and discussion

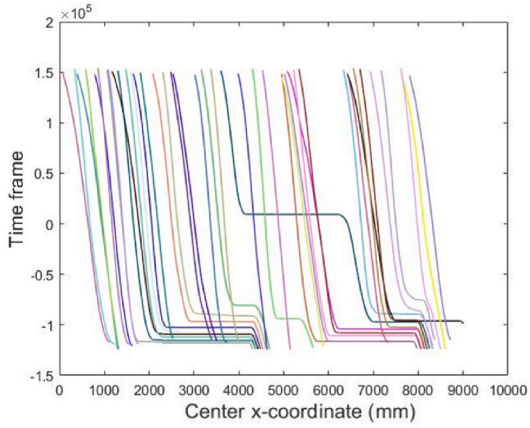
The complete trajectory samples were selected to explore the different speed characteristics and safety features in four road segments (USCA, UGCA, MSCS-BF, and MSCS-BH) caused by distinct external factors such as the road space, marked lines and intersections in this section.



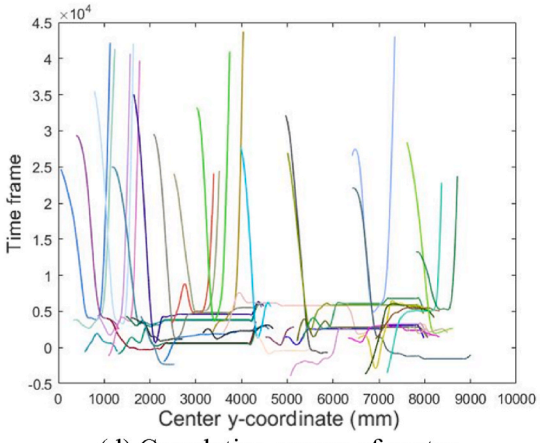
(a) Correlation curves of center *x-y* coordinates and time frames



(b) Correlation curves of center *x* coordinates and center *y* coordinates



(c) Correlation curves of center *x* coordinates and time frames



(d) Correlation curves of center *y* coordinates and time frames

Fig. 6. Space-time diagram of 50 complete vehicle trajectories.

4.1. Traffic flow characteristics in ROI

4.1.1. Vehicle speed characteristics

(1) Longitudinal velocity and longitudinal acceleration

As shown in Table 2, the overall longitudinal velocity of vehicles in ROI was at a relatively low value with the trend of accelerating first and decelerating later. The average speed in MSCS-BF was the highest (41.34 km/h), while the average speed in MSCS-BH was 19.15 km/h, which was the lowest compared with other research sections. Meanwhile, the standard deviation of vehicle speed in these two areas were twice as big as that in USCA and UGCA. This may be attributed to the fact that vehicles needed to stop and wait due to the existence of intersections and signal lights. In USCA, the average speed of ETC vehicles was 31.20 km/h, which was greater than that of MTC vehicles (20.34 km/h). It may be because ETC vehicles had an initial high speed because of non-stop payment. Subsequently, both types of vehicles accelerate in UGCA. We further illustrate the average longitudinal speed distribution of individual vehicles in Fig. 7. It can be seen that the distribution of average speed in the whole process had apparent periodicity relative to the arrival order of vehicles, which gradually increased from 5 km/h to 45 km/h in a traffic light circle. It reached the minimum when it needed to wait for the whole red light in MSCS-BH.

As shown in Table 3, vehicles accelerated longitudinally after entering the USCA and UGCA. Under the influence of varied road space and lane-changing behaviors in UGCA, the average longitudinal acceleration was only 0.29 m/s^2, which was smaller than that of USCA (0.74 m/s^2). In MSCS-BF, drivers needed to choose different lanes according to the roadside lane indication signs, and there might be overlapping decisions and traffic conflicts with surrounding vehicles. Therefore, the vehicles decelerated in the MSCS-BF and the average longitudinal deceleration was −0.16 m/s^2. Since vehicles needed to avoid collision with other low-speed vehicles and queue up before the signal light in MSCS-BH, the overall deceleration was 0.59 m/s^2. We further show the average longitudinal acceleration distribution of individual vehicles in Fig. 8. There were negative values (deceleration) and positive values (acceleration) of lateral acceleration simultaneously, indicating that there was heterogeneity in the speed change of the vehicle within the merging area. The change of speed was directly affected by surrounding vehicles, road space, lane queuing, etc.

In USCA, ETC vehicles conducted non-stop payment, so it was easier for them to accelerate. In contrast, MTC vehicles needed to stop and start again after paying. Therefore, the longitudinal average acceleration of ETC vehicles (0.84 m/s^2) was greater than that of MTC vehicles (0.61 m/s^2). In UGCA, the ETC vehicles had a high moving speed after the acceleration in previous stage and nearly reached the maximum speed for driving demands, while the MTC vehicles needed to accelerate to meet the safe driving requirements. Consequently, the average longitudinal acceleration of ETC vehicles was 0.18 m/s^2, which was lower than that of the MTC vehicle (0.52 m/s^2).

(2) Lateral velocity and lateral acceleration

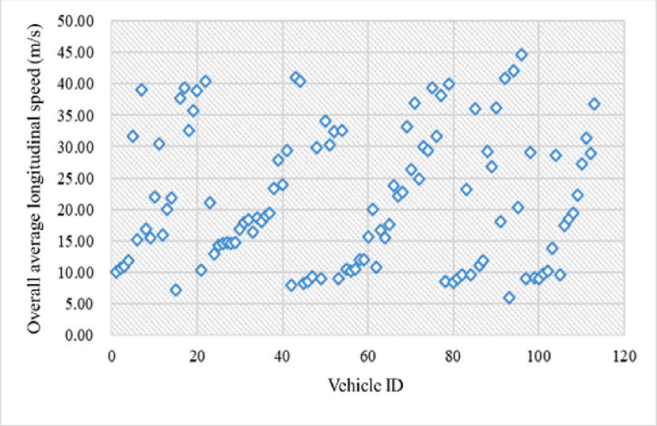


Fig. 7. Average longitudinal speed distribution of individual vehicles.

Table 4 shows the average lateral speed distribution of vehicles in ROI. It can be seen that the overall average lateral speed of vehicles in the downstream area was −0.25 m/s, indicating that the vehicles accelerated to the inside of the road. Due to the layout of the downstream toll area, vehicles needed to merge from the outside to the inside of the road when entering the CS. The lateral average speed reached the maximum in UGCA (−0.74 m/s), indicating that vehicles change lanes more frequently affected by the alignment of transition section. ETC vehicles tended to change lanes to the outside, while MTC vehicles tended to change lanes to the inside. Likewise, the movement of vehicles in USCA was the same as UGCA. Vehicles might need to change lanes to reach the designated route due to the existence of intersection and lane indication signs. Therefore, the average lateral speed in MSCS-BH was the smallest compared with the other three areas, which was −1.01 m/s.

The fluctuation of the average lateral speed of different vehicles in ROI was shown in Fig. 9. The relationship between the average lateral speed and vehicle ID reflects the lane-changing direction and lane-changing times of vehicles. A positive value of the average lateral speed indicates that the vehicles changed lanes to the outside in ROI, while a negative value indicates that the vehicles changed lanes to the inside in ROI. At the same time, the absolute value could reflect the number of lane-changing times to a certain extent.

The fluctuation of the average lateral acceleration of different vehicles in ROI is shown in Table 5 and Fig. 10. It can be seen that the average lateral acceleration in the USCA was 0.115 m/s^2, which was smaller than its longitudinal acceleration of 0.74 m/s^2 in Table 3. This may be attributed to the fact that the drivers tended to drive outside the road affected by the narrow space. Meanwhile, the average lateral acceleration of MTC vehicles was greater than that of ETC vehicles in the UGCA. The layout of the lanes forces vehicles to accelerate to the inside of the road when entering the CS. The average decelerations were −0.061 m/s^2 and –0.075 m/s^2 in the UGCA and MSCS-BF, respectively. Since MTC lanes were outside the road, vehicles needed to travel more distance to change to the target lanes inside the road. Consequently, the

Table 2
Average longitudinal speed distribution in ROI.

ROI		Vehicle type	Average speed (km/h)	Number of observations	S.D. (km/h)	Min (km/h)	Max (km/h)
CA	USCA	All vehicles	27.61	15,072	8.47	5.17	48.98
		ETC vehicles	31.20	8528	7.38	5.17	48.98
		MTC vehicles	20.34	7456	5.24	7.50	36.57
	UGCA	All vehicles	39.12	30,376	9.63	5.87	65.55
		ETC vehicles	41.82	22,861	9.26	21.45	65.55
		MTC vehicles	33.54	7515	7.45	5.87	57.11
CS	MSCS-BF	All vehicles	41.34	13,982	16.25	0	64.02
	MSCS-BH	All vehicles	19.15	130,380	16.05	0	56.71
Total			28.05	189,810	17.25	0	65.55

Table 3
Average longitudinal acceleration distribution in ROI.

ROI		Vehicle type	Average acceleration (m/s^2)	Number of observations	S.D. (m/s^2)	Min (m/s^2)	Max (m/s^2)
CA	USCA	All vehicles	0.74	15,072	0.62	−4.45	7.67
		ETC vehicles	0.84	8528	0.67	−2.91	7.67
		MTC vehicles	0.61	7456	0.70	−4.45	5.98
	UGCA	All vehicles	0.29	30,376	0.67	−5.06	5.44
		ETC vehicles	0.18	22,861	0.67	−5.06	2.83
		MTC vehicles	0.52	7515	0.64	−4.28	5.44
CS	MSCS-BF	All vehicles	−0.16	13,982	0.78	−5.31	3.94
	MSCS-BH	All vehicles	−0.59	130,380	0.76	−7.59	6.66
Total			0.05	189,810	0.78	−7.59	7.67

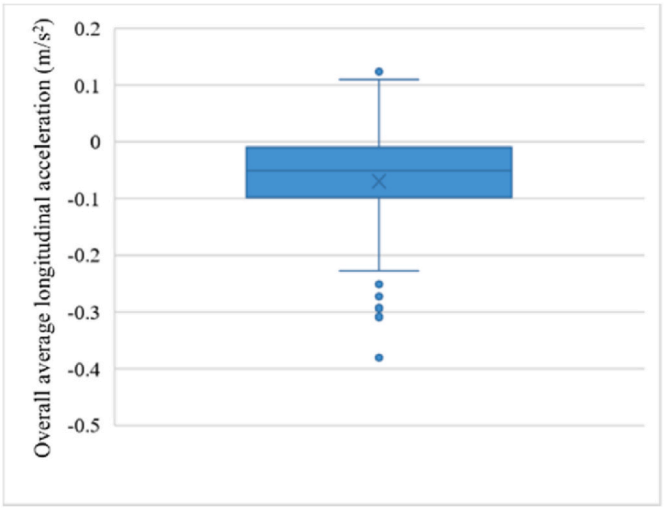


Fig. 8. Average longitudinal acceleration distribution of individual vehicles.

average lateral acceleration of MTC vehicles was −0.081 m/s^2, which was greater than that of ETC vehicles (−0.055 m/s^2). In the MSCS-BH, vehicles needed to drive to the lanes that could avoid potential collisions with other low-speed vehicles and queue up for the signal light. The overall lateral acceleration was −0.203 m/s^2, indicating that vehicles' lane-changing behaviors were more intense in the MSCS-BH.

4.1.2. Lane-changing and travel time characteristics

(1) Lane-changing points

To eliminate the inclination of shooting angle caused by wind and obtain more accurate target detection results, the geodetic coordinate systems during 14:30–14:35 and during 14:36–14:42 were calibrated respectively. The spatial distribution of vehicles' lane-changing points in two time periods is shown in Fig. 11, and the statistical results are listed in Table 6. The initial lanes of vehicles with more lane changes were mainly located outside the toll station (6, 7, and 8 MTC lanes), while the initial lanes of vehicles with less lane changes were mainly located inside the toll station (1, 2 and 3 ETC lanes). It can also be found that vehicles on both sides of the lane tended to change lanes in advance in the USCA compared to the middle. In the UGCA, the lane-changing behaviors of vehicles were more frequent on lanes 5, 6, 7 and 8 to enter the straight-line connecting sections. In the MSCS-BF, vehicles that did not enter the designated lane in the previous stage continued to change lanes. Due to the different queue length on lanes 2 and 3, drivers continued to change lanes in the MSCS-BH to shorten the queuing time and pass the intersection as soon as possible.

(2) Lane-changing times

To analyze the lane-changing behavior of vehicles in the downstream area of the toll station, this paper quantitatively classified the vehicles based on total lane-changing times starting from vehicles entering the merging area to leaving the connecting section. The distribution of lane-changing times is shown in Table 7 and Fig. 12. It can be seen that the distribution of lane-changing times was uneven. The average number of lane-changing times per vehicle was 2.56 and 61.9% of the vehicles changed lanes no more than twice. Only two vehicles changed lanes three times in the whole process, and the number of vehicles changed lanes for 4, 5, 6, and 7 times were 11, 10, 14, and 6, respectively. The

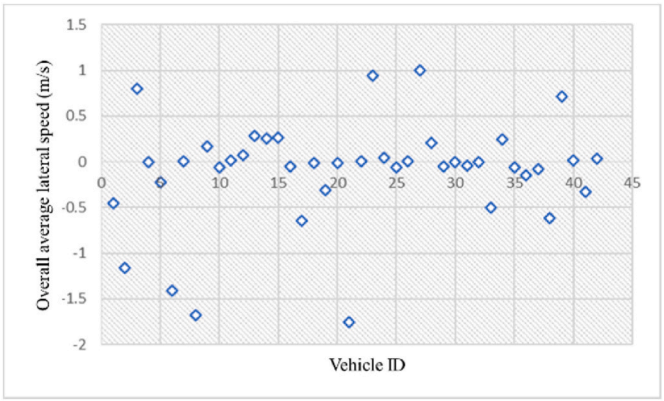


Fig. 9. Average lateral speed distribution of individual vehicles.

Table 4
Average lateral speed distribution in ROI.

ROI		Vehicle type	Average speed (km/h)	Number of observations	S.D. (km/h)	Min (km/h)	Max (km/h)
CA	USCA	All vehicles	−1.20	1572	0.53	−3.80	8.51
		ETC vehicles	0.57	8528	0.97	−3.80	1.89
		MTC vehicles	−3.08	7456	2.05	−2.86	8.51
	UGCA	All vehicles	−0.74	30,376	0.95	−3.25	12.60
		ETC vehicles	0.99	22,861	2.59	−3.25	11.43
		MTC vehicles	−7.13	7515	2.45	1.57	12.60
CS	MSCS-BF	All vehicles	0.10	13,982	0.23	−0.38	1.22
	MSCS-BH	All vehicles	−1.01	130,380	2.15	−10.32	0.70
Total			−0.25	189,810	1.79	−10.32	12.60

Table 5
Average lateral acceleration distribution in ROI.

ROI		Vehicle type	Average acceleration (m/s^2)	Number of observations	S.D. (m/s^2)	Min (m/s^2)	Max (m/s^2)
CA	USCA	All vehicles	0.115	15,072	0.103	−0.216	0.661
		ETC vehicles	0.020	8528	0.066	−0.216	0.200
		MTC vehicles	0.209	7456	0.136	−0.023	0.661
	UGCA	All vehicles	−0.061	30,376	0.186	−1.007	1.511
		ETC vehicles	−0.055	22,861	0.192	−1.007	0.533
		MTC vehicles	−0.081	7515	0.167	−0.591	1.511
CS	MSCS-BF	All vehicles	−0.075	13,982	0.161	−0.667	0.264
	MSCS-BH	All vehicles	−0.203	130,380	0.456	−2.650	0.789
Total			−0.100	189,810	0.358	−2.650	0.789

Fig. 10. Average lateral acceleration distribution of individual vehicles.

vehicles' lane-changing width was obtained by multiplying the total number of lane-changing times generated in ROI by the measured road width 3.5 m. A bigger lane-changing time corresponds to a larger lane-changing width, which was more likely to cause the traffic flow disorder. It indicated that the layout of toll lanes had a direct impact on the lane-changing process and indirectly affected the collision risk in the confluence area.

(3) Travel time characteristics

The time from entering the specific area of the downstream area of the expressway toll station to leaving this section is defined as the travel time. A statistical analysis of the travel time of vehicles was conducted in Table 8. The results showed that vehicles' average travel time in the CA was 14.28 s, with a standard deviation of 3.39 s. Affected by the traffic signal control and the reduction of lanes, the vehicles in the CS needed a longer time to pass through the merging area. Therefore, the average driving time (45.43 s) in the CS was higher than that in the CA, while the standard deviation in the CS was 31.49 s, which was much greater than that in the CA. The average travel time in ROI was 59.70 s with maximum 165.56 s and minimum 22.02 s.

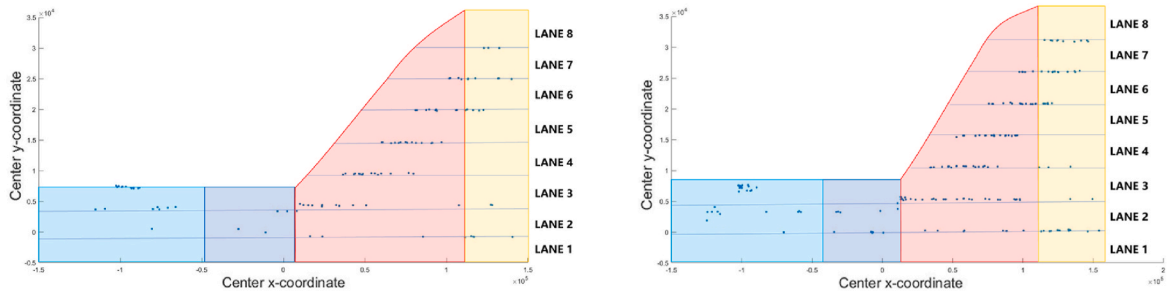
Considering the sample data size will affect the distribution hypothesis test results, the P–P (Probability Plot) diagram and Q-Q (Quantile-Quantile) diagram were used to judge the data distribution intuitively. Figs. 13 and 14 show the lognormal P–P diagram and detrending lognormal P–P diagram of the travel time of vehicle samples in the CS. In the normal P–P diagram, the actual distribution of samples was close to the theoretical distribution. Although the residuals fluctuated up and down, most of the differences were less than 0.10 in the detrending lognormal P–P diagram. Therefore, it can be considered that

Table 6
Statistics of location distribution of lane-changing points.

ROI		Number of lane-changing points during 14:30–14:35	Number of lane-changing points during 14:36–14:42
CA	USCA	17	39
	UGCA	61	106
CS	MSCS-BF	5	8
	MSCS-BH	22	31

Table 7
Statistics of lane-changing times.

Lane-changing times	Lane-changing width (m)	Frequency	Percentage	Cumulative percentage
0	0	25	22.1	22.1
1	3.5	27	23.9	46.0
2	7.0	18	15.9	61.9
3	10.5	2	1.8	63.7
4	14.0	11	9.7	73.5
5	17.5	10	8.8	82.3
6	21.0	14	12.4	94.7
7	24.5	6	5.3	100.0
Total	–	113	100.0	–



(a) Lane-changing points during 14:30-14:35 (b) Lane-changing points during 14:36-14:42

Fig. 11. The distribution of lane-changing points in ROI.

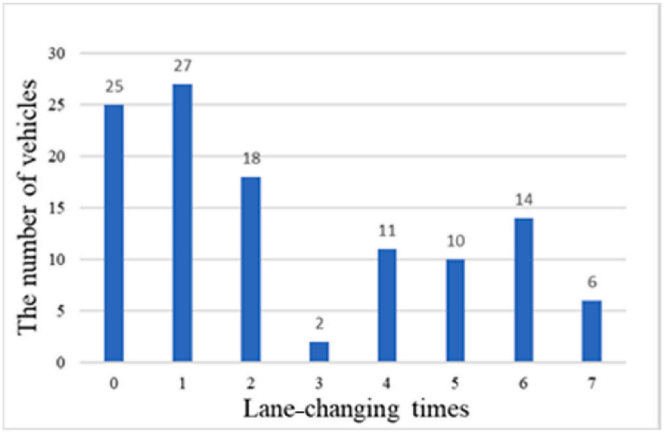


Fig. 12. The distribution of vehicle lane-changing times.

Table 8
The distribution of travel time in ROI.

	Travel time (s)				
	Number of vehicles	Average	S.D.	Min	Max
CA	113	14.28	3.39	9.43	30.56
CS		45.43	31.49	10.03	147.53
ROI		59.70	32.13	22.02	165.56

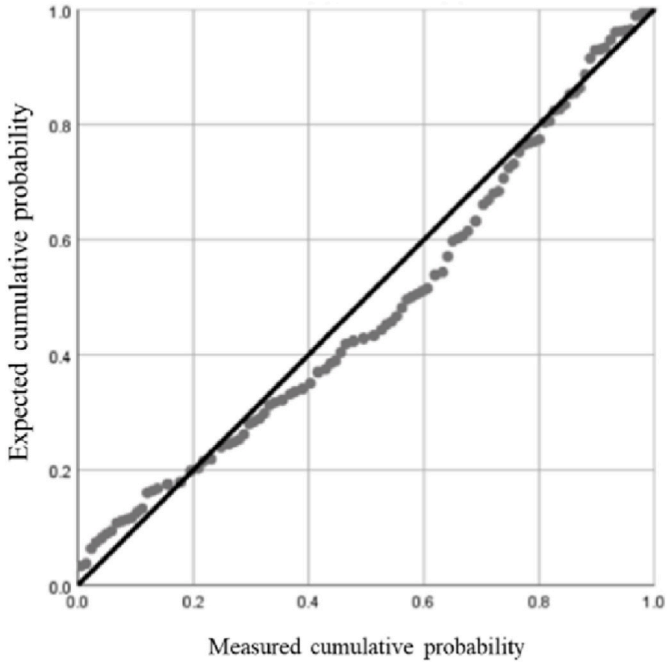


Fig. 13. Lognormal P–P diagram of the travel time.

the sample raw data approximately obey lognormal distribution in 90% confidence level.

4.2. Traffic conflict analysis

To reveal the safety characteristics in the downstream area of the expressway toll station, the traffic conflict surrogate indicators were utilized. According to previous studies (Rahman et al., 2019; Xing et al., 2020), the thresholds of ETTC and PET in this study were set as 3 s and 1.5 s, respectively. If both the ETTC and PET calculated were less than

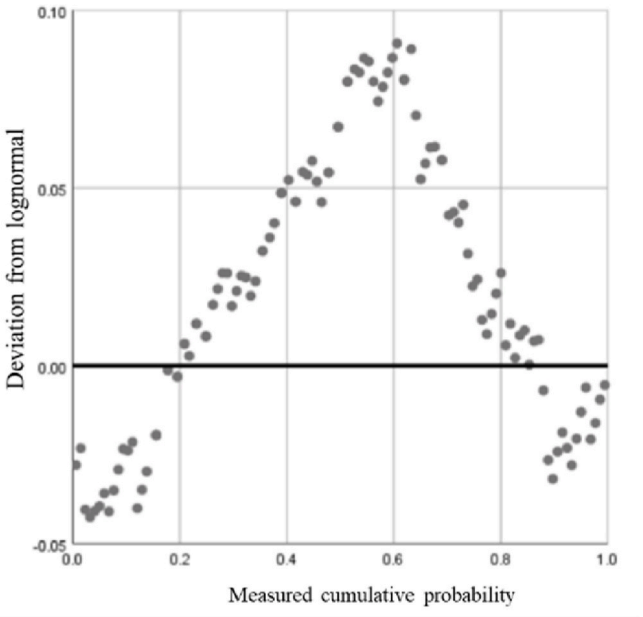


Fig. 14. Detrending lognormal P–P diagram of the travel time.

their thresholds, it was marked as a traffic conflict case.

4.2.1. Spatial distribution and severity characteristics of traffic conflicts

The spatial distribution of identified conflict samples is shown in Figs. 15 and 16. Note that the conflict samples were derived from data point pairs of vehicles sampled at a rate of 30 frames per second. This micro-level segmentation can offer us a more detailed, high-precision depiction of traffic conflict distribution, thus yielding an accurate representation as indicated by past research (Hussain et al., 2023; Nazir et al., 2023; Zheng and Sayed, 2020). A total of 4653 conflict points were found in the complete vehicle trajectory samples in the downstream area of the toll station, including 850 conflict points in the CA and 3803 conflict points on the CS. It can be attributed to the fact that the space of the CS was narrower than that of CA. Meanwhile, drivers may face more decision-making challenges due to the coexistence of right turn ramps, intersections, and traffic signs. The result indicates that the conflicts between MTC vehicles on the outer lanes and ETC vehicles on the inner lanes were mixed before the merging area and then interleaved in the merging area. As the vehicles completed the merging, the conflict points gradually gathered in the rear section of the CS.

Table 9 summarizes the traffic conflict severity characteristics. The average of ETTC was 1.87 s, among which the ETTC values of MTC and ETC vehicles were 1.83 s and 1.88 s, respectively. It shows that the gap between two types of vehicles was small, which may attribute to the fact that the CS directly connected with the ETC toll lanes and the lane-changing behavior can be completed with fewer lane changes. As shown in Fig. 17, the ETTC ranged from 0 s to 3 s and the number of conflicts with high risk (0 s–1 s) was relatively low. In the high-risk observations, the number of ETC vehicles was greater than MTC vehicles. Meanwhile, the average PET value of all conflict samples was 0.73 s and the PET values of MTC and ETC vehicles were 0.85 s and 0.69 s respectively, indicating that ETC vehicles had a higher risk of traffic accidents in the CA of the toll station. As shown in Fig. 18, the PET was concentrated in the range of 0.3 s–1.2 s. In the conflicts with high risk (0 s–0.9 s), the number of conflict samples of ETC vehicles was more than that of MTC vehicles and the proportion of ETC vehicles keeps increasing with the increase of risk level, indicating more serious traffic conflicts of ETC vehicles, which can be explained by their larger initial speed.

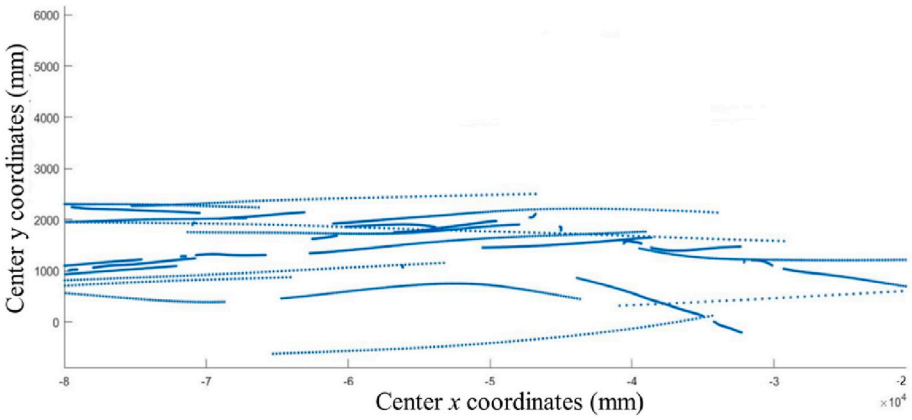


Fig. 15. The distribution of traffic conflict points in CA.

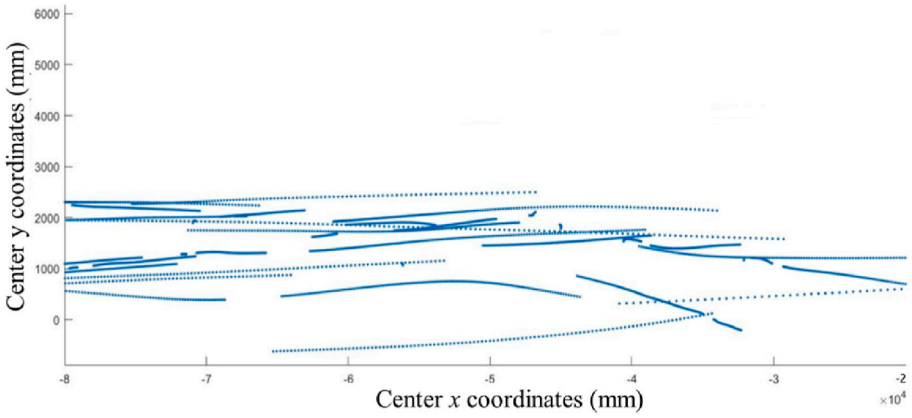


Fig. 16. The distribution of traffic conflict points in CS.

Table 9
Traffic conflict severity characteristics of conflict samples.

Indicator		ETTC (s)				PET (s)			
Vehicle type	Number	Average	S.D.	Min	Max	Average	S.D.	Min	Max
MTC vehicles	456	1.83	0.67	0.28	2.99	0.85	0.27	0.28	1.50
ETC vehicles	1226	1.88	0.61	0.58	3.00	0.69	0.30	0.25	1.49
All vehicles	1682	1.87	0.63	0.28	3.00	0.73	0.30	0.25	1.50

4.2.2. The impact of initial toll lane and termination lane selection on traffic conflicts

Table 10 shows the distribution of conflict samples by initial toll lane. Vehicles that chose ETC toll lanes 1 and 2 were more likely to have serious traffic conflicts. Besides, the possibility of traffic conflict showed a decreasing trend from the outside to the inside of the toll channel, but there existed heterogeneity on lanes 6 and 7. Meanwhile, the vehicles choosing lane 7 were more dangerous than those on its adjacent outer lanes 6 and 8. The uneven initial toll lane selection also leaded to the varying frequency of traffic conflicts. The number of conflicts of vehicles choosing inner ETC lanes 1–3 was larger than that of vehicles choosing outer MTC lanes 4–8. The ETTC and PET values of the conflict samples on lanes 2 and 7 were both smaller than that of other toll lanes, which meant that vehicles that chose these two toll channels would suffer higher risks. Therefore, in the confluence area of the toll station, the choice of channels at the starting point could directly determine the lane-changing behavior and possible conflicts. For vehicles selecting the inner straight lanes, they had less lane-changing demand in the vertical road direction than those that selected outer toll channels. It was noteworthy that the conflicts of ETC vehicles may increase since their initial speed were not zero and were prone to potential risks.

Table 11 displays the statistics of safety indicators on different termination lanes. The results demonstrated that the number of conflicts on termination lanes 3 and 4 were lower than that of lanes 1 and 2, indicating that vehicles to lanes 1 and 2 were more likely to have accident risks in the confluence area. The distribution characteristics of the termination lanes selected by vehicles affected the distribution of traffic conflicts. At the same time, there were differences in the termination lane selection between ETC vehicles and MTC vehicles. The average ETTC and PET of ETC vehicles were smaller than the conflict indicators of MTC vehicles in total conflict samples. This was because the higher average speed of ETC vehicles caused by non-stop payment. Fig. 19 shows the distribution of conflict samples of MTC and ETC vehicles on each termination lane. The proportion of different vehicles with the same termination lane showed an imbalance. Among the conflict samples, the proportion of ETC vehicles to the termination lane 2 was the highest, while the numbers of conflicts between MTC vehicles and ETC vehicles on termination lanes 3 and 4 were similar.

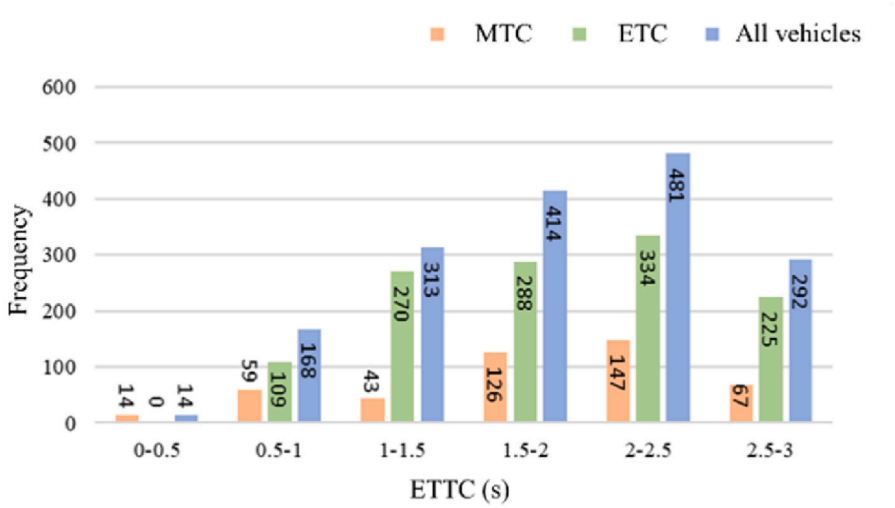


Fig. 17. The frequency distribution of ETTC.

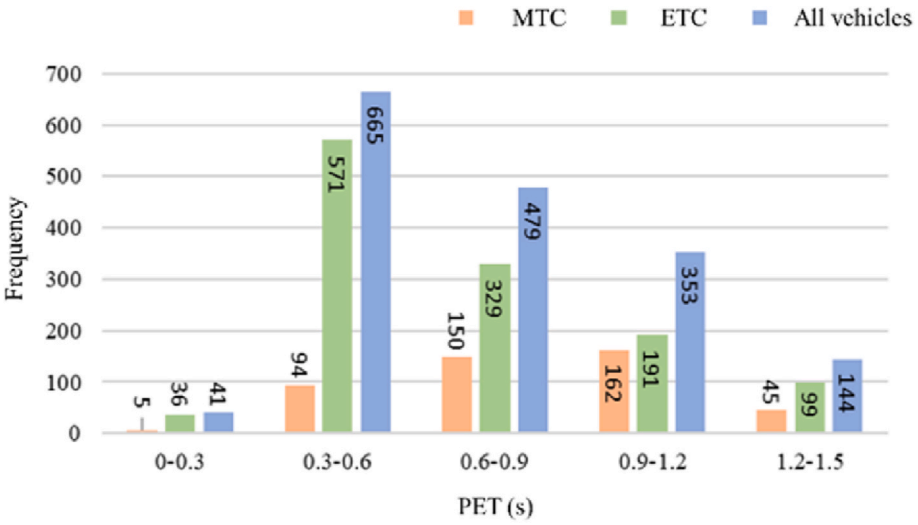


Fig. 18. The frequency distribution of PET.

Table 10
The distribution of conflict samples by initial toll lane.

Initial toll lane	Number of samples	ETTC (s)		PET (s)	
		Average	S.D.	Average	S.D.
1	1851	1.53	0.62	0.87	0.33
2	808	1.46	0.70	0.81	0.35
3	684	1.73	0.64	0.76	0.33
4	129	1.82	0.43	0.81	0.15
6	565	1.74	0.71	0.86	0.35
7	223	1.49	0.79	0.75	0.33
8	394	1.91	0.72	0.97	0.37
Total	4653	1.61	0.68	0.84	0.34

5. Conclusions

The purpose of this study is to explore the traffic flow characteristics and depicts the traffic safety features in the downstream area of the expressway toll station. Based on the aerial video of the downstream area of the Huai'an South Toll Station, the YSKT vehicle trajectory data extraction framework was employed to obtain the micro vehicle trajectory raw data. The data was then used to deeply mine the characteristics of traffic flow in the downstream area of the toll station including driving speed characteristics, lane-changing characteristics, and travel time. The characteristic differences of ETC vehicles and MTC vehicles were also compared to understand the mechanism and distribution of different types of vehicle traffic flow. Moreover, the characteristics of traffic conflict including spatial distribution, severity and lane selection in ROI were further analyzed by utilizing the traffic conflict indicators ETTC and PET, which indicated that the ETC vehicles had a higher risk of traffic accidents compared to MTC vehicles since their greater initial speed.

The results of this paper could provide theoretical guidance and empirical references for the road design and traffic management in the expressway toll stations. For example, the current study indicates that lane-changing points are primarily concentrated in the UGCA area. Considering the gradient slope of this region could significantly affect the status of lateral traffic flow, it would be reasonable to further calculate the appropriate slope to improve safety levels in the future. Besides, the findings demonstrate that vehicles in MSCS-BF exhibit higher average longitudinal speed and standard deviation. And within the MSCS-BH area, vehicles experience greater deceleration and standard deviation in both longitudinal and lateral directions due to traffic signal lights and lane-changing demands. The traffic conditions in this area are more chaotic compared to other regions. Considering the

Table 11
The distribution of conflict samples by termination lane.

Termination lane	Vehicle type	Number of Samples	ETTC (s)		PET (s)	
			Average	S.D.	Average	S.D.
1	MTC vehicles	515	1.72	0.66	0.82	0.32
	ETC vehicles	919	1.69	0.60	0.83	0.33
	All vehicles	1435	1.70	0.62	0.83	0.33
2	MTC vehicles	346	1.50	0.73	0.72	0.33
	ETC vehicles	2038	1.48	0.67	0.80	0.35
	All vehicles	2384	1.48	0.68	0.79	0.34
3	MTC vehicles	92	1.86	0.89	1.13	0.30
	ETC vehicles	100	1.81	0.58	1.08	0.28
	All vehicles	192	1.83	0.75	1.10	0.29
4	MTC vehicles	357	2.03	0.63	1.01	0.32
	ETC vehicles	286	1.55	0.60	0.99	0.30
	All vehicles	642	1.82	0.66	1.00	0.31
Total	MTC vehicles	1310	1.76	0.72	0.87	0.35
	ETC vehicles	3343	1.55	0.65	0.83	0.34
	All vehicles	4653	1.61	0.68	0.84	0.34

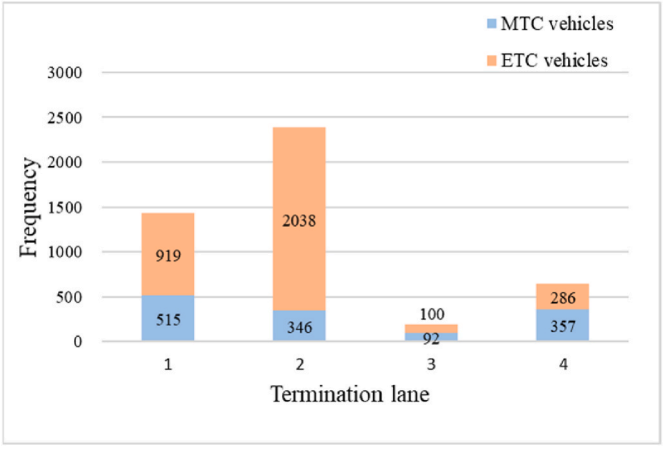


Fig. 19. The distribution of conflict samples of MTC and ETC vehicles by termination lane.

crucial role of this node in connecting city roads and highways, a well-designed linking section length should be explored further to improve safety characteristics. Given the distinct traffic features and conflicts in the four regions, it would be beneficial to enhance driver safety awareness for such nodes and validate the effectiveness of existing warning signs. Moreover, it is recommended to implement differentiated management approaches for ETC vehicles and MTC vehicles, such as setting different warning signs at toll gates, to ensure safety in the toll station. Regarding the upstream traffic demand management, the information service can be provided for the vehicles through the traffic guidance facilities to induce the drivers to choose the exit and get off the expressway with good traffic conditions when the connecting part is paralyzed.

This study can also provide a reference for assessing traffic safety at other complex road nodes, such as intersections and overpasses. However, it is important to note that our findings and conclusions are derived from the trajectory data collected in the downstream area of expressway toll station during off-peak hours. Therefore, the absence of data from other time periods and road sections may limit the generalizability of our results. In the future, more representative time periods (such as morning and night), similar road sections (such as upstream area of toll station), and longer data collection time should be considered to provide a comprehensive understanding of the temporal variations and spatial distribution patterns. Additionally, the trajectory data and historical crash data can be combined to further analyze the correlation between crashes and identified conflicts among such complex road sections and provide more effective traffic management strategies for expressway toll stations. Finally, this study lays the groundwork for various other studies, such as traffic conflict assessment, signal timing optimization, vehicle lane-changing safety studies, and traffic organization optimization.

CRediT authorship contribution statement

Qiaoqiao Ren: Conceptualization, Methodology, Writing – original draft. **Jie He:** Conceptualization, Data curation, Writing – review & editing. **Ziyang Liu:** Conceptualization, Data curation, Writing – review & editing. **Min Xu:** Conceptualization, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (Grant No. 51778141 and 52072069), the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. PolyU 15210620), and the Hong Kong Polytechnic University (G-UARN).

References

Abuzwidah, M.A., 2019. Evaluation and Modeling of the Safety of Open Road Tolling System, vol. 2011, pp. 2004–2019.
Abuzwidah, M., Abdel-Aty, M., 2018. Crash risk analysis of different designs of toll plazas. Saf. Sci. 107, 77–84.
Chamoso, P., González-briones, A., Rodríguez, S., Corchado, J.M., 2018. Tendencies of technologies and platforms in smart cities: a state-of-the-art review. Wireless Commun. Mobile Comput. 2018, 3086854.
Chen, H., You, Y., Zhou, J., Wang, L., Shang, Z., 2013. A simplified approach to estimate the urban expressway capacity after traffic accidents using a micro-simulation model. Adv. Mech. Eng. 5, 656345.
Chen, Q., Gu, R., Huang, H., Lee, J., Zhai, X., Li, Y., 2021. Using vehicular trajectory data to explore risky factors and unobserved heterogeneity during lane-changing. Accid. Anal. Prev. 151, 105871.
Cooper, P., 1984. Experience with traffic conflicts in Canada with emphasis on post encroachment time techniques. International Calibration Study of Traffic Conflicts. NATA ASI Series F5, 75–96.
Feng, H., Ren, X., Li, L., Zhang, X., Chen, H., Chai, Z., Chen, X., 2021. A novel feature-guided trajectory generation method based on point cloud for robotic grinding of freeform welds. Int. J. Adv. Manuf. Technol. 115 (5), 1763–1781.
Hajiseyedjavadi, F., McKinnon, I., Fitzpatrick, C., Knodler Jr, M.A., 2015. Application of microsimulation to model the safety of varied lane configurations at toll plazas. Transportation Research Board 94th Annual Meeting 15–3480. Washington DC, United States, January 11-15.
Hou, X., Wang, Y., Chau, L.-P., 2019. Vehicle tracking using deep sort with low confidence track filtering. In: 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Taipei, Taiwan, pp. 1–6. September 18-21.
Hua, S., Anastasiu, D.C., 2019. Effective vehicle tracking algorithm for smart traffic networks. In: 2019 IEEE International Conference on Service-Oriented System Engineering (SOSE), San Francisco, CA, USA, pp. 67–6709. April 04-09.

Hussain, F., Ali, Y., Li, Y., Haque, M.M., 2023. Real-time crash risk forecasting using Artificial-Intelligence based video analytics: a unified framework of generalised extreme value theory and autoregressive integrated moving average model. Anal. Methods Accid. Res. 40, 100302.
Jackson, S., Miranda-Moreno, L.F., St-Aubin, P., Saunier, N., 2013. Flexible, mobile video camera system and open source video analysis software for road safety and behavioral analysis. Transport. Res. Rec. 2365 (1), 90–98.
Jiang, Z., Yu, D., Zhou, H., Luan, S., 2021. A trajectory optimization strategy for connected and automated vehicles at junction of freeway and urban road. Sustainability 13 (17), 9933.
Li, Y., Li, Z., Wang, H., Wang, W., Xing, L., 2017. Evaluating the safety impact of adaptive cruise control in traffic oscillations on freeways. Accid. Anal. Prev. 104, 137–145.
Li, Y., Wang, H., Wang, W., Liu, S., Xiang, Y., 2016. Reducing the risk of rear-end collisions with infrastructure-to-vehicle (I2V) integration of variable speed limit control and adaptive cruise control system. Traffic Inj. Prev. 17 (6), 597–603.
Li, Y., Wu, D., Chen, Q., Lee, J., Long, K., 2021. Exploring transition durations of rear-end collisions based on vehicle trajectory data: a survival modeling approach. Accid. Anal. Prev. 159, 106271.
Liu, Z., He, J., Zhang, C., Yan, X., Wang, C., Qiao, B., 2022. Vehicle trajectory extraction at the exit areas of urban freeways based on a novel composite algorithms framework. J. Intell. Transport. Syst. 1–19.
Ministry of Transport of China, 2021. National Toll Road Statistics Bulletin of 2020.
Nazir, F., Ali, Y., Sharma, A., Zheng, Z., Haque, M.M., 2023. Car-following crash risk analysis in a connected environment: a Bayesian non-stationary generalised extreme value model. Anal. Methods Accid. Res. 39, 100278.
Oh, C., Kim, T., 2010. Estimation of rear-end crash potential using vehicle trajectory data. Accid. Anal. Prev. 42 (6), 1888–1893.
Rahman, M.S., Abdel-Aty, M., Lee, J., Rahman, M.H., 2019. Safety benefits of arterials' crash risk under connected and automated vehicles. Transport. Res. C Emerg. Technol. 100, 354–371.
Saad, M., Abdel-Aty, M., Lee, J., 2019. Analysis of driving behavior at expressway toll plazas. Transport. Res. F Traffic Psychol. Behav. 61, 163–177.
Seleznov, I., Popov, A., Kikuchi, K., Kolosova, E., Kolomiiets, B., Nakata, A., Kaneko, M., Kiyono, K., 2020. Detection of oriented fractal scaling components in anisotropic two-dimensional trajectories. Sci. Rep. 10 (1), 1–11.
Umair, M., Farooq, M.U., Raza, R.H., Chen, Q., Abdulhai, B., 2021. Efficient video-based vehicle queue length estimation using computer vision and deep learning for an urban traffic scenario. Processes 9 (10), 1786.
Vahidi, A., Sciarretta, A., 2018. Energy saving potentials of connected and automated vehicles. Transport. Res. C Emerg. Technol. 95, 822–843.
Wu, J., Xu, H., Zheng, Y., Tian, Z., 2018. A novel method of vehicle-pedestrian near-crash identification with roadside LiDAR data. Accid. Anal. Prev. 121, 238–249.
Xing, L., He, J., Abdel-aty, M., Cai, Q., Li, Y., Zheng, O., 2019. Examining traffic conflicts of up stream toll plaza area using vehicles' trajectory data. Accid. Anal. Prev. 125, 174–187.
Xing, L., He, J., Li, Y., Wu, Y., Yuan, J., Gu, X., 2020. Comparison of different models for evaluating vehicle collision risks at upstream diverging area of toll plaza. Accid. Anal. Prev. 135, 105343.
Zheng, L., Sayed, T., 2020. A bivariate Bayesian hierarchical extreme value model for traffic conflict-based crash estimation. Anal. Methods Accid. Res. 25, 100111.